Malaysian Journal of Mathematical Sciences 11(S) February: 115–141 (2017)
Special Issue: Conference on Agriculture Statistics 2015 (CAS 2015)

MALAYSIAN JOURNAL OF MATHEMATICAL SCIENCES

Journal homepage: http://einspem.upm.edu.my/journal

On Generalization of Additive Main Effect and Multiplicative Interaction (AMMI) Models: An Approach of Row Column Interaction Models for Counting Data

Hadi, A. F. *[1], Sa'diyah, H. [2], and Iswanto, R. [3]

[1]*Statistical Laboratory, Department of Mathematics, University of Jember, Indonesia*
[2]*Biometrics Laboratory, Department of Agronomy, University of Jember, Indonesia*
[3]*Indonesian Legumes and Tuber Crops Research Institute, Indonesia*

E-mail: afhadi@unej.ac.id
**Corresponding author*

ABSTRACT

Additive Main Effect and Multiplicative Interaction (AMMI) model was commonly used to analyze Genotype Environment × Interaction with normal response variables, now it had been generalized for categorical or other non-normal response variables, called GAMMI model. This development was conducted by introducing multiplicative terms to the Generalized Linear Model (GLM). This research round up our previous work on developing an approach of Row Column Interaction Models (RCIMs) comprise to GAMMI model and focus to get more generalized for counting data with overdispersed and zeros problems. A few interesting things here are (i) an issue of distribution on GLM sense and (ii) an issue of model's complexity that is the number of multiplicative terms to fit the interaction effect more properly. On the distribution issue of counting data, we will focus on Poisson, Negative Binomial (NB), and zero inflated problems with Zero Inflated Poisson (ZIP) and Zero Inflated NB (ZINB)

distribution. A simulation conducted by adding outlier(s) on a Poisson counting data for overdispersed condition, and adding zeros observation on the data for illustrating the zero problems. We propose the NB model for overdispersed data and model of ZIP or ZINB for data with both, overdispersed and zero problem. In the case of both illnesses conditions happened simultaneously, the mean square error of NB and ZINB will increase slightly. But the ZINB was resulting the simplest model of RCIM with less number of interaction terms.

Keywords: Multiplicative Models, Negative Binomial, Overdisperse, Poisson, Zero Inflated.

1. Introduction

AMMI model is commonly used to analyze stability and adaptability on the Genotype × Environments interaction (GEI) studies. AMMI provide an additive model for main effects of genotype and environment plus a complete multiplicative terms for the interaction effects. Basically, the interaction terms was modeled by a statistical technique of reduction dimension called Singular Value Decomposition (SVD). With SVD, the interaction terms will have complete parameters, a paramater for every single cell of the two ways table. SVD will visualize the interaction terms by Biplot and makes the GEI analysis become easier. With this feature of Biplot, AMMI said to be most powerful model for the GEI (Hadi et al. (2010)).

The advantages of AMMI model for the GEI analysis, together with it's limitation on normality assumption, inspire many statisticians to develop AMMI to be more generalized by introducing GLM sense to AMMI model. Van Eeuwijk (1995) propose the Generalized AMMI model which introduce multiplicative term to GLM. GAMMI also keep the feature of Biplot visualization of GEI. Nowadays, GAMMI model had been broadly applied to a counting data response as in Hadi et al. (2010). Another overlapping methodology is the RCIM of Yee and Hadi (2014). RCIM is designed for many kind of interaction model with various response including GAMMI for Poisson count. In case of Poisson response in two ways table, formula of RCIM look identical to GAMMI model with log-link function. RCIM was an approach built up on Reduce Rank Regression (RRR) for GLM, it's called RR-VGLM (Yee and Hastie (2003)) while GAMMI used a criss-cross regression (Van Eeuwijk (1995)). Computation of parameter estimate of these two models used the same type of algorithm of alternating regression, it was confirmed by Turner and Firth (2015) and also by Yee and Hadi (2014). In spite of this similarity, they are different in model parameterization and constraint, thus it still allows them to give different result.

Starting with a framework of statistical development of RCIM as an alternative model to GAMMI for two ways table of counting data, this paper want to deliver a wrap up review of our previous work on application of RCIM to GEI analysis, and to get more generalized of GAMMI model by an approach of RCIM model. We focus on the generalization for counting data related to the limitation problematic of equidispersion assumption of Poisson distribution (that the conditional means of Poisson data distributed is equal to it's variance). With this strict assumption of Poisson, any extreme values of observation may cause a violation of the pure Poisson distribution. The main violation here is the overdispersion, when a counting data has variance greater

than the mean is called overdispersed. A Poisson data distributed with large mean value will also has a large variance, thus any large extreme value (right outlier) will be interesting here. On the other hand, any extreme value(s) here is including a zero valued observation(s). Poisson with a low mean valued may have extra-zero observations, then here Poisson will mimic a problem called an extra-zero or a zero-inflation problem.

We found a few interesting issues. The first issue is the data distribution, about the canonical link function applied on GLM to fit the data properly. Here, we focus on Poisson and NB distribution, including problems of overdispersion by outliers and also zero-inflation with ZIP and ZINB distribution. The second issue is the model's complexity. It is about the number of multiplicative terms involved in the model (represented by rank of model) to describe the interaction effects. A simple scheme of simulation was conducted to present an overdispersion and zero-inflated condition into a Poisson counting data. Then we introduce RCIM with NB distribution for handling overdispersed data count by outliers, also ZIP and ZINB for zero-inflated problem.

2. Row Column Interaction Models (RCIM)

This section will discuss a framework of developing model of RCIM (Figure 1). RCIM will fit data count of two ways table with every single cell containing row and column effects plus some interaction effects as reduce rank regression and visualize the interaction terms of rank = 2 by biplot. Here RCIM was very similar to GAMMI model, where GAMMI was decomposing the interaction effects by Singular Value Decomposition (SVD) and also visualize it through biplot for the first two singular vectors. Both approaches give similar results, differing only in numerical computation aspect and no statistically essential. Yee and Hadi (2014) said that RCIM is developed from RR-VGLM which is applied to a matrix data of row-column containing interaction effects by reduce rank regression. RR-VGLM it self is a variant of Vector GLM (VGLM), as clearly described in Yee and Hastie (2003). For further reading, there were some early discussions of RR-VGLM in RCIM context such as Yee and Hadi (2014) and Hadi and Sa'diyah (2016), we also can find for more wider class of modeling in Yee (2015). Something important here, that is the compliance of parameter setting in RR-VGLM. We turn it to the SVD parameterization, and we will depart to develop RCIM for two way table that comprise to GAMMI model for the GEI. The RCIMs approach for GEI analysis has introduce an applicative biplot visualization of the interaction effects in case of Poisson counting data response with zero-inflated problems (Hadi and Sa'diyah (2014)). Later, Hadi and Sa'diyah (2016) was supplementing RCIM as alternative way to GAMMI

model with the deviance analysis for determining the number of multiplicative terms needed for the interaction analysis or even for handling overdispersion, both at once. Furthermore, we now will develop RCIM to be more generalized with Negative Binomial and also Zero Inflated Negative Binomial.

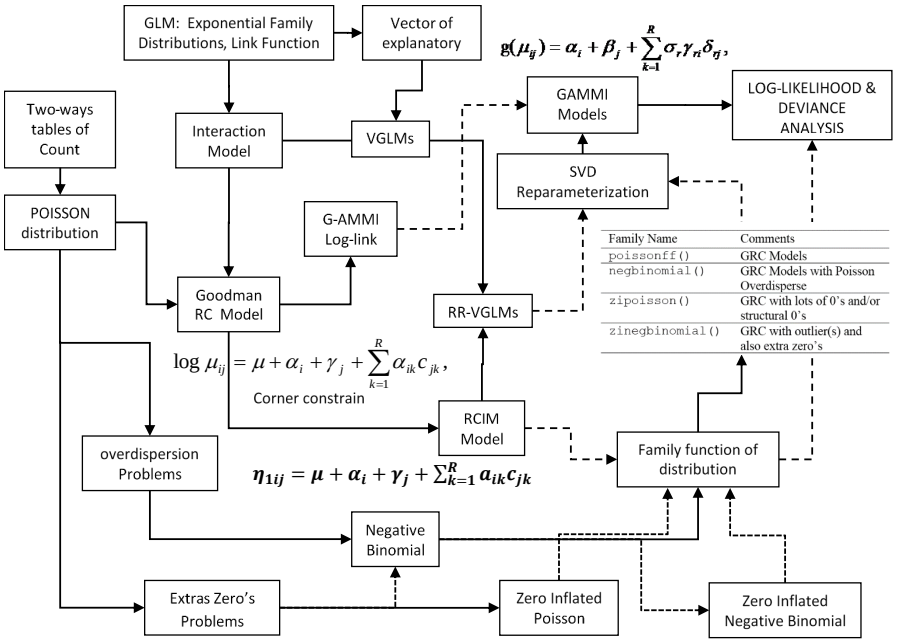

Figure 1: Statistical framework of The RCIM

2.1 VGLMs and RR-VGLMs

RCIM was built upon VGLMs and RR-VGLMs (see Figure 1). We now generally, will talk about VGLMs and RR-VGLMs, for more detail the reader directed to see Yee and Hastie (2003), Yee (2014) and also Yee (2015). Let the observed response y is a q-dimensional vector. VGLMs are defined as a model where

$$f(y|x;\mathbf{B}) \;=\; h(y,\eta_1,\ldots,\eta_M) \tag{1}$$

for some known function $h(\cdot)$, $\mathbf{B} = (\beta_1\, \beta_2\, \cdots\, \beta_M)$ is a $p \times M$ matrix of unknown regression coefficients and $\mathbf{x}$ is explanatory. The jth linear predictor is

$$\eta_j \;=\; \beta_j^T \mathbf{x} \;=\; \sum_{k=1}^{p} \beta_{(j)k}\, x_k, \qquad j = 1, \ldots, M, \tag{2}$$

where $\mathbf{x} = (x_1, \ldots, x_p)^T$ with $x_1 = 1$ for an intercept.

GLMs only have single linear predictor of η for the mean, but VGLMs may have more, each may be applied to a certain parameters of a distribution. For example, a univariate distribution has two parameters of the location parameter a and the scale parameter b. Then we might take two linear predictors of VGLM here, η_1 is for a and η_2 is for b. In general, $\eta_j = g_j(\theta_j)$ for some certain link function g_j and parameter θ_j. **VGAM** offers many link functions, that can be assigned to any parameters, ensuring maximum flexibility.

Most VGLMs have a log-likelihood which is maximized. Let x_i denote the explanatory vector for the ith observation, for $i = 1, \ldots, n$. Then we can write the equation 2 as

$$\eta_i \;=\; \begin{pmatrix} \eta_1(x_i) \\ \vdots \\ \eta_M(x_i) \end{pmatrix} \;=\; \mathbf{B}^T x_i \;=\; \begin{pmatrix} \beta_1^T x_i \\ \vdots \\ \beta_M^T x_i \end{pmatrix}. \tag{3}$$

The IRLS algorithm behind **VGAM** almost always implements Fisher scoring based on the expected information matrix (EIM) at the individual i level.

In practice we may wish to constrain the effect of a covariate to be the same for some of the η_j and to have no effect for others. For example,

$$\begin{aligned} \eta_1 &= \beta^*_{(1)1} + \beta^*_{(1)2}\, x_2 + \beta^*_{(1)3}\, x_3, \\ \eta_2 &= \beta^*_{(2)1} + \beta^*_{(1)2}\, x_2, \end{aligned}$$

so that $\beta_{(1)2} \equiv \beta_{(2)2}$ and $\beta_{(2)3} \equiv 0$. The *star* superscript denote regression parameters that are actually estimated. For VGLMs, we can represent these models using

$$\eta(x) \;=\; \sum_{k=1}^{p} \beta_{(k)}\, x_k \;=\; \sum_{k=1}^{p} \mathbf{H}_k\, \beta^*_{(k)}\, x_k \tag{4}$$

where $\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_p$ are known full-column rank *constraint matrices*, $\beta^*_{(k)}$ is a vector containing a possibly reduced set of regression coefficients. With no constraint at all, all $\mathbf{H}_k = \mathbf{I}_M$ and $\beta^*_{(k)} = \beta_{(k)}$. Then

$$\mathbf{B}^T \;=\; \begin{pmatrix} \mathbf{H}_1\beta^*_{(1)} & \mathbf{H}_2\beta^*_{(2)} & \cdots & \mathbf{H}_p\beta^*_{(p)} \end{pmatrix}. \tag{5}$$

2.1.1 The RR-VGLMs

Represent the VGLMs of (1) and its linear predictor of (2), we now turn to partition x into $(x_1^T, x_2^T)^T$ and $\mathbf{B} = (\mathbf{B}_1^T\ \mathbf{B}_2^T)^T$. In general, $\mathbf{B}$ is a full rank matrix of $\min(M, p)$. There are $M \times p$ regression coefficients to be estimated. In some cases, it would be a problem here like degree of freedom deficiencies or other problems regarding that is too many parameters to be estimated. Now we need a method of dimension reduction here. That is to replace $\mathbf{B}_2$ by an RRR of $\mathbf{B}_2 = \mathbf{A}\,\mathbf{C}^T$ with lower rank of $R \ \leq \ \min(M, p)$ matrices of $\mathbf{A}$ and $\mathbf{C}$. This reduction of the number of regression coefficients will done efficiently by put R as low, e.g., 0 or 1 or 2. Something grab our attention for the next feature of our model. That is a fact that by taking $R = 2$, the $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{C}}$ may be biplotted. The reduced-rank regression is applied to $\mathbf{B}_2$ because we want to make provision the variables in x_1 to be left alone for the intercepts.

Now we have the RR-VGLMs of the form

$$\eta \ = \ \mathbf{B}_1^T x_1 + \mathbf{B}_2^T x_2 \ = \ \mathbf{B}_1^T x_1 + \mathbf{A}\,\mathbf{C}^T x_2 \ = \ \mathbf{B}_1^T x_1 + \mathbf{A}\,\nu \tag{6}$$

where $\mathbf{C} = (c_{(1)} \cdots c_{(R)})$ is $p_2 \times R$, $\mathbf{A} = (a_{(1)} \cdots a_{(R)}) = (a_1, \ldots, a_M)^T$ is $M \times R$. Both $\mathbf{A}$ and $\mathbf{C}$ are of full-column rank. Of course, $R \leq \min(M, p_2)$ but ideally we want $R \ll \min(M, p_2)$. One can think of (6) as a reduced-rank regression of the coefficients of x_2 after having adjusted for the variables in x_1 (commonly x_1 is left as the intecept of μ).

In order to make the parameter being unique, we may enforce identifiability constraint to restrict $\mathbf{A}$ to the form

$$\mathbf{A} \ = \ \begin{pmatrix} \mathbf{I}_R \\ \widetilde{\mathbf{A}} \end{pmatrix}, \quad \text{say,} \tag{7}$$

called *corner* constraints. Actually, it may be necessary to represent $\mathbf{I}_R$ in rows other than the first R; this is controlled by the argument `Index.corner` which has value `1:Rank` as default. It transpires that RR-VGLMs are VGLMs where the constraint matrices are estimated. An alternating algorithm is used which toggles between estimating $\mathbf{A}$ and $\mathbf{C}$ one at a time based on the current estimate of the other.

2.2 Row-Column Interaction Model for Data count in the RR-VGLM

2.2.1 Goodman's Row-Column association model

Hadi and Sa'diyah (2014) use an association model of Goodman's Row-Column (GRC) of Goodman (1981) to describe the RCIM model in the RR-VGLMs framework, by firstly assuming that $\mathbf{Y} = [(y_{ij})]$ is a $n \times M$ matrix of counts and Y_{ij} has a Poisson distribution, $E(Y_{ij}) = \mu_{ij}$ is the mean of the i-j cell. Goodman's RC(R) association model fits a reduced-rank type model to $\mathbf{Y}$, and the linear predictor is

$$\log \mu_{ij} \quad = \quad \mu + \alpha_i + \gamma_j + \sum_{r=1}^{R} c_{ir}\, a_{jr}, \tag{8}$$

where $i = 1, \ldots, n,\ j = 1, \ldots, M$. Note that (8) is saturated when $R = \min(n, M)$.

In (8) the parameters α_i and γ_j are called the *row* and *column scores* (or *effects*) respectively. Identifiability constraints are needed for these, such as corner constraints, e.g., $\alpha_1 = \gamma_1 = 0$. The parameters a_{ir} and c_{jr} also need constraints, e.g., $a_{1r} = c_{1r} = 0$ for $r = 1, \ldots, R$.

We can write (8) as

$$\log \mu_{ij} \quad = \quad \mu + \alpha_i + \gamma_j + \delta_{ij},$$

where the $n \times M$ matrix $\Delta = [(\delta_{ij})]$ of interaction terms is approximated by the reduced rank quantity

$$\delta_{ij} \quad = \quad \sum_{r=1}^{R} c_{ir}\, a_{jr}. \tag{9}$$

The GRC association model fits within the VGLM framework of (6) by letting

$$\eta_i = \log \mu_i \tag{10}$$

where μ_i^T is the mean of the ith row of Y. Then the GRC model will fit the matrix $(\eta_1, \ldots, \eta_n)^T$ using RRR by setting up the indicator variables in $\mathbf{B}_1^T x_{1i}$. The reader directed to Yee and Hadi (2014) and Yee (2015) for further reading about how to get the appropriate indicator variable setting. Similarly, $\mathbf{B}_2$ is

approximated by $\mathbf{C}\mathbf{A}^T$, the Δ is approximated by

$$\begin{pmatrix} x_{21} \\ \vdots \\ x_{2n} \end{pmatrix} \mathbf{C}\,\mathbf{A}^T$$

The desired reduced-rank approximation of Δ can be obtained if $x_{2i} = e_i$ so that $\mathbf{I}_{p_2}\mathbf{C}\,\mathbf{A}^T = \mathbf{C}\,\mathbf{A}^T$. Note that

$$\begin{aligned} \Delta &= \begin{pmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & \widetilde{\Delta} \end{pmatrix} \approx \mathbf{C}\,\mathbf{A}^T \\ &= \begin{pmatrix} \mathbf{0}^T \\ \mathbf{C}_{(-1)} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \left(\mathbf{A}_{(-1)}\right)^T \end{pmatrix}, \end{aligned} \tag{11}$$

that is, the first row of $\mathbf{A}$ consists of structural zeros which are 'omitted' from the reduced rank regression of Δ (Yee and Hastie (2003)).

2.2.2 The RCIMs, GAMMI models, and SVD reparametriation

Finally, we define RCIMs as a RR-VGLM with

$$\eta_{1ij} = \mu + \alpha_i + \gamma_j + \sum_{r=1}^{R} c_{ir}\, a_{jr}, \tag{12}$$

Note that (12) applied to the *first* linear/additive predictor; for models with $M > 1$ one can leave $\eta_2, \ldots, \eta_M$ unchanged. Of course, choosing η_1 for (12) is only for convenience and is the default.

GAMMI model of Van Eeuwijk (1995) as described in Turner and Firth (2015), use the singular value to factor out a measurement of the strength of interaction between the row and column scores corresponding to each multiplicative component. It is indicating the importance of the component, or axis. For cell means μ_{ij} a GAMMI-R model has the form

$$g(\mu_{ij}) = \alpha_i + \beta_j + \sum_{k=1}^{R} \sigma_k \gamma_{ki} \delta_{kj} \tag{13}$$

Based on (13) GAMMI model appear to be identical to RCIMs. Here GAMMI apply a SVD to the $\mathbf{A}\mathbf{C}^T$, and also some constraints of $\sum_{\forall i} \alpha_i = \sum_{\forall i} \gamma_i = 0$, the parameters a_{ir} and c_{jr} use constraints of $\sum_{\forall i} a_{1r} = \sum_{\forall i} c_{1r} = 0$ for $r = 1, \ldots, R$ (Van Eeuwijk (1995)). While in RCIMs, the interaction term uses corner constraints. The advantage of RCIMs is that it should work

for any VGAM family functions, thus the family size is much bigger (Yee and Hadi (2014), Yee (2010), Yee (2008)). It is easy to perform some post-transformations such as applying a function of `svd()` to the VGAM output to obtain the SVD parameterization for GAMMI model (Yee and Hadi (2014)). Now we can see that GAMMI is an RCIMs with some other parameterization of SVD related to what described in Yee and Hastie (2003).

3. Material and Methods

This research use three datasets which are originally obtained from the experimental trial conducted by Indonesian Legumes and Tuber Crops Research Institute (ILETRI), Malang, Indonesia. The *first* dataset comes from the experimental trial of study the endurance of five genotypes of soybean to 5 types of its leaf pests. The *second* dataset comes from a study of leaf disease attack on mung bean. The experimental trial involved 12 genotypes (varieties) of mung bean which planted in 5 different environments at Probolinggo, Jombang, Jember, Rasanae, and Bolo. The *third* dataset obtained from a study of soybean in ILETRI. This experiment uses 15 types of soybean lines grown at 8 locations with a number of soybean pods. This is a counting data without outlier neither zero observation, which is presented in the form of a matrix with a size of 15×8.

3.1 Methodology

We will use the 1^{st} dataset of Poisson distributed to summarize our development of deviance analysis feature of RCIM to determine the rank of model for analyzing the interaction terms as provided by GAMMI. A biplot of the interaction analysis was also provided by RCIM. A study of outlier, overdispersion, and the NB model will be discussed by conducted a scheme of simulation on the 3^{rd} dataset to make an illness condition of overdispersion by outliers. We impose the outliers of about 20 percent cells of 15×8 cells data matrix. With this simulated dataset, we investigate the influence outliers to the estimated value of dispersion parameter in a standard Poisson model. We also discuss about the use of NB distributional with its canonical link-function to overcome the overdispersion compared to the standard Poisson model with more interaction terms in the model. We compare the log-likelihood and also the MSE of the Poisson and NB models. In addition we investigate the influence of the percentages of outliers to the MSE of Poisson and NB models by setting increment of 0.8%, 1.7%, 2.5%, ... , 19.2% of outliers in 15×8 cells data matrix.

We discuss the zero problem in RCIM by firstly summarize our previous development of introduce the ZIP distribution to our RCIM model, and apply ZIP RCIM to the 2^{nd} dataset also comparing to standard Poisson one by it's log-likelihood value. Next, we introduce a ZINB distribution on RCIM, and conduct a scheme of simulation to add three zeros into the 3^{rd} dataset by replacing the three smallest value observations at every column by zeros. Here, we got a lot of zeros on the 3^{rd} dataset and we compare the MSE of the NB and the ZINB. For both illnesses condition of outliers and zeros in one dataset, we replace the maximum value by the outlier, and the smallest three values by zeros simultaneously into the 3^{rd} dataset. Again, we compare the MSE of the NB and ZINB model. Last, we compare the NB and ZINB for a data with structural zero as we got in 2^{nd} dataset. We also add outlier(s) by replacing the maximum value observation at every column by a value of max(column) + 3 × stdev(column). So we compare the MSE of NB and ZINB to the data with structural zeros and also outliers at once.

4. Result and Discussion

4.1 An Application of RCIM for GAMMI Models: The Deviance Analysis and Biplot of RCIM

The *first* dataset Table (1) contain the population count of 5 types of leaf pest on four soybean genotypes. It was originally analyzed by Hadi et al. (2010) on the Poisson distribution with the GAMMI model proposed by Van Eeuwijk (1995).

Table 1: The 1^{st} dataset: Count of population of Leaf Pests on some Soybean Genotypes

	Leaf Pests				
Genotype	Bemisia tabacci	Empooascan sp.	Agromyza phaseoli	Lamprosema indicata	Longitarsus suturellinus
AC100	2	7	9	2	7
IAC80	12	11	4	7	13
W80	14	12	5	8	8
Wilis	16	12	4	7	16

The deviance for a model of μ is defined as the ratio the likelihood of the saturated model $L(y; y)$ denumerated by the likelihood of the particular model $L(\mu; y)$ (Pawitan (2001)):

$$D = 2 \log \frac{L(y; y)}{L(\mu; y)} \tag{14}$$

It measures the distance between a particular model μ and the observed data y. Deviance is also commonly used to compare the two nested models with different rank. Suppose we have two models, model A with μ_A have $X_1\beta_1$ of rank p and model B with μ_B have $X_1\beta_1 + X_2\beta_2$ with rank of q, for p less than q. The difference in the observed deviance

$$D(y, \widehat{\mu_A}) - D(y, \widehat{\mu_B}) = 2 \log \frac{L(y; y)}{L(\mu; y)} \tag{15}$$

is the usual likelihood ratio test for the hypothesis $H_0 : \beta_2 = 0$.

Table 2: The Deviance Analysis of RCIM Models for testing the Rank=2

Source	df	Deviance	Mean Deviance	Ratio of Mean Deviance	p-value
Leaf Pests (column)	4	16.7380	4.1845	78.38	0.01283
Genotype (row)	3	11.3434	3.7812	70.83	0.01423
GAMMI1 (rank=1)	6	14.6836	2.4473	45.84	0.02172
GAMMI2 (rank=2)	4	3.7908	0.9477	17.75	0.05482
Residual	2	0.1068	0.0534		
Total	19	46.6626	2.4560		

Here we used RCIM model by VGAM Package with the function of `rcim`. From (15), one can provide analysis of deviance that is commonly used in the GAMMI model of Van Eeuwijk (1995), as provided in Table 2. We obtained the deviance of additive models (with no interaction term) by subtracting the residual deviance of the null model by residual deviance in each model. The deviance of GAMMI1 model obtained from a subtracting the residual deviance of rank = 0 model by its of rank = 1 model, with corresponding degree of freedom, then we continue for GAMMI2, GAMMI3 models and so on. For more details of this calculation, please see Hadi and Sa'diyah (2016). With this analysis of deviance, Hadi and Sa'diyah (2016) concluded that the interaction analysis was best fitted by RCIM with rank of 2 (or GAMMI2 model) with log-link. Then the visualization of the interaction effects was done by Biplot of RCIM with rank of 2. This Biplot based on RCIM approach (Figure 2) was verified statistically, that there is no clearly difference to the Biplot from GAMMI model of Van Eeuwijk (1995) as figure out and well described in Hadi et al. (2010) .

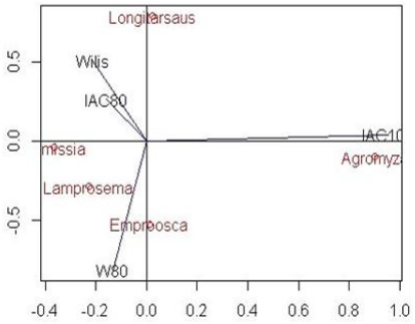

Figure 2: The Biplot of RCIM rank=2 for endurance of varieties of soybean to some leaf pests.

4.2 RCIM with Overdispersion problem on Counting Data: A Negative Binomial Distribution

Poisson models assume a strict relationship between the mean and variance that may not appropriate for some counting data. Practically, a common cause of overdispersion is an additional variation to the mean or heterogeneity, particularly may caused by outlier(s) (Hadi and Sa'diyah (2016)). Nevertheless, overdispersion can occur mathematically, if the conditional mean of an outcome Y_μ was Poisson random variable with mean μ, and the μ is also random variable with mean E_μ and variance σ^2.

For example, plants vary in the propensity to their leaves to be infected by insect of leaf pest, eventhough the number of infected leaf per individual is a Poisson distribution, the marginal distribution of Y_μ has mean and variance respectively as follows:

$$\begin{aligned} E(Y_\mu) =& E[E(Y_\mu|\mu)] \\ =& E\mu \end{aligned}$$

$$\begin{aligned} var(Y) &= E[var(Y_\mu|\mu)] + var[E(Y_\mu|\mu)] \\ &= E\mu + var(\mu) \\ &= E\mu + \sigma^2 \end{aligned}$$

The mean and the variance above are indicating an extra variability to the pure Poisson model. If μ was a gamma distributed random variable with parameter of integer α we will get the marginal probability as negative binomial distribution (Pawitan (2001)).

This section will discuss the problem of overdispersion in RCIM modeling, starting with investigation of the influence of outlier(s) in Poisson count data from an overdispersed simulated data and then we propose the RCIM with NB distribution for this illness condition of overdispersion. A dataset mainly used here was the response of a number of non-empty soybean's pods from an experiment involving 15 types of soybean lines in 8 locations. There was no outlier neither zero observation. A simulation then carried out by adding outlier(s) to learn whether it will shift the estimated value of the dispersion parameter getting larger than it should be. The imposing of outlier in to the data was completed by adding a tripled standard deviation of each row (column) to the cell containing the maximum value of its row (column). We added up to 20 outliers into the rows and columns observation of data matrix of the 3^{rd} dataset, we have simulated data matrix that contains up to 19.2 % cells of outliers.

4.2.1 Overdisperse in Poisson Count Data: Outlier and The Dispersion Parameter

Here we briefly summarize the magnification in estimated value of dispersion parameter influenced by outlier in Poisson model of RCIM, also straighten out some less informed about the log-likelihood comparison on our previous study on Hadi and Sa'diyah (2016). Table 3 showed that outliers made a suffer illness of overdispersion.

Table 3: Estimated value of Dispersion Parameter for standard Poisson Model of RCIM

Model	Overdispersed with no outlier			Overdispersed with Outliers		
	Deviance	df	Estimated Dispersion	Deviance	df	Estimated Dispersion
Rank = 1	170.615	20	8.531	201.868	20	10.093
Rank = 2	99.284	18	5.516	119.123	18	6.618
Rank = 3	61.391	16	3.837	69.151	16	4.322
Rank = 4	30.348	14	2.168	35.177	14	2.513
Rank = 5	14.192	12	1.183	16.680	12	1.390
Rank = 6	4.869	10	0.487	5.617	10	0.562

In lower rank of RCIM, Poisson with log-link failed to fit the overdispersed Poisson counting data with or without outlier. Since the dispersion parameter larger than 1, it may cause a problem in hypothesis testing of parameter models determining best fit ones. However, in higher rank model of RCIM, the Poisson model may overcome the overdispersion, this is shown by the estimate value of dispersion parameter of rank = 5, less than 1.25 for Poisson data without outlier. But for overdispersed data by outliers, the estimated value of dispersion

parameter is still larger than 1.25 at rank = 5 model of RCIM. As Hilbe (2011) suggested using Poisson regression, if the dispersion value of less than or equal to 1.25, we should worry to use Poisson model of RCIM for this data containing outlier, unless we use full model of rank = 6.

4.2.2 Overdisperse in RCIM: Canonical Link Function and Multiplicative Term

With the same datasets as previous section (4.2.1), we now try to do RCIM with other distribution function of NB in spite of usual Poisson to model an overdispersed counting data. In case of there was no outlier in overdispersed counting data, the NB model of RCIM could overcome the overdispersion better than Poisson, generally for all rank of RCIM. Poisson can do it by the rank = 2 model or by model with more complex interaction terms to get equal log-likelihood value. See Table 4 for rank = 2 (or more) of Poisson RCIM had the same log likelihood value.

In case of there was a suffer overdispersion by outliers, NB model provide similar information of overcoming the overdispersion problem. But here, Poisson need one more rank of 3 to do it with equal log-likelihood value as the NB model.

Table 4: The Log-Likelihood of RCIM models (with canonical link function) affected by outliers

Model	Overdispersed no outlier		Overdispersed with Outlier(s)	
	Poisson Regression	NB Regression	Poisson Regression	NB Regression
Null	-2198.594	-631.832	-2281.542	-634.415
Rank = 0	-547.002	-496.746	-588.291	-508.626
Rank = 1	-454.704	-451.278	-471.524	-462.886
Rank = 2	-419.038	-419.038	-430.151	-430.076
Rank = 3	-400.092	-400.092	-405.165	-405.165
Rank = 4	-384.570	-384.570	-388.178	-388.178
Rank = 5	-376.492	-376.492	-378.929	-378.930
Rank = 6	-371.831	-371.831	-373.398	-373.398
Rank = 7	-369.397	-369.397	-370.590	-370.590

Table 5: The MSE RCIM Model with Poisson and Negative Binomial Distribution

Data	Model	Poisson	Negative Binomial
Overdispersed data with no outlier	RCIM 1	0.021562830	0.020838680
	RCIM 2	0.012077778	0.012077740
	RCIM 3	0.008437611	0.008434989
	RCIM 4	0.005427515	0.005427515
	RCIM 5	0.002552477	0.002552321
	RCIM 6	0.000600769	0.000600608
Overdispersed data with outliers	RCIM 1	0.025835524	0.024468820
	RCIM 2	0.014287709	0.014146090
	RCIM 3	0.009145594	0.009145374
	RCIM 4	0.005893773	0.005893766
	RCIM 5	0.002428332	0.002428258
	RCIM 6	0.001028197	0.001028055

Table 5 contain the MSE that shows how close the predicted value of the model to its actual observation data. Here we got similar information that in general, (1) the outliers will affect the model to get the larger MSE, (2) the NB give a better fit to observation than Poisson with smaller MSE. It was also confirmed here that Poisson with rank = 4 of RCIM model had fitted the overdispersed data with no outlier as good as the NB model by exactly the same value of MSE. But for overdispersed data containing outliers, there is none of the rank of the Poisson model that can fit the outliers data as good as the NB model.

We now turn to discuss the influence of the percentages of outliers in the data by setting increment of 0.8%, 1.7%, 2.5%, . . , 19.2% of outliers in the 15 $\times$ 8 data matrix. Figure 3 shows that on a multiplicative model with lowest rank (rank = 1 and rank = 2) NB model of RCIM perform better than Poisson model, by smaller MSE. But in the more complex model, Poisson model give an exactly equal MSE to NB model. Here we can say that Poisson model can handle the overdispersion by involving more interaction terms in its model. The more complex the model, the more severe overdispersed can be handled.

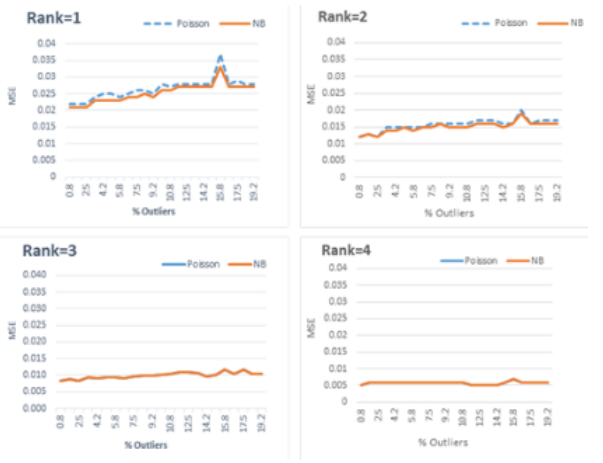

Figure 3: The MSE of RCIM rank=1, 2, 3 and 4 for Poisson and Negative Binomal (NB) Distributions on Simulated Data containing 0.8% - 19.2% outliers

4.3 RCIM with Excess Zero Problems on Counting Data

Another problem of Poisson counting data that will be discussed here is the excess zero observations. Again, this problem was coming out with well-known assumptions of Poisson, that is equidispersion. A Poisson count with small value of mean should also have small value of variance. In this case, Poisson potentially to have a number of zero observations might be more than its expected. With this zero-inflation, the underlying distributional assumption of Poisson may not be met. We also need to consider here is the possibility of overdispersion. The zero-inflation and overdispersion may occur simultaneously in a data set. Zero inflated models are often discussed in additive modeling, but less in multiplicative model, like GAMMI or even more RCIM.

Practically, in the study of GEI, the response variable may be an observation of counting data measuring the lower the better. Researcher expected to get a genotype with more zero count of attack at more environments. There would be inflation of zero. With this kind of zero observation, there were some other approaches of distributional data to model zero-inflated count data, that concern to classify the zero observation into two groups, ie. a zero-modified distributions (zero-altered, zero-truncated, distribution with added zero) or a mixture distribution (zero-inflated distribution).

In this mixture distribution of zero-inflated, the zero will be classified into two groups, a group of positive (discrete) count distribution (Poisson or NB) occur with probability of $1-\omega$; the other represent the 'extra' zero, occur with probability of ω.

4.3.1 Zero-inflated Poisson in the RR-VGLM

ZIP model is powerful in dealing with counting data with excess zeros than the usual Poisson distribution, partly it is because the ZIP model also handles overdispersion. To see that, we will write down the probability mass function (p.m.f.) of the ZIP in two stages with two-components mixture distribution.

$$f(Y|\theta,\omega) = \begin{cases} \frac{(1-\omega)e^{-\theta}\theta^y}{y!}, & \text{for } y = 1,2,3,... \\ \omega + (1-\omega)e^{-\theta}, & \text{for } y = 0; \text{with } 0 \leq \omega < 1 \end{cases} \tag{16}$$

Hadi and Sa'diyah (2014) write down the expectation of the p. m. f. to get the mean and variance of ZIP as $E(Y) = \mu$ and $var(Y) = \mu + (\omega/(1-\omega))\mu^2$. For positive ω, the conditional distribution shows an overdipersion and the ZIP will turn to a standard Poisson if $\omega = 0$. The log-likelihood function of a vector of random sample ZIP distributed as $l(\theta,\omega;\mathbf{y})$, please see Hadi and Sa'diyah (2014) for more detail and also to get the joint model for ω and θ as

$$\begin{aligned} log\ (\tfrac{\omega}{1-\omega}) &= \mathbf{G}\gamma \quad \text{and} \\ log\ (\theta) &= \mathbf{XB} \end{aligned} \tag{17}$$

And in the linear predictors of RCIM model as (12) we now write η_1 and η_2 as

$$\eta = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} logit\ \omega \\ log\ \mu \end{pmatrix} \tag{18}$$

There are two processes of how the data occurs, the first data is zero and the second is Poisson count data. Both processes are modeled respectively by η_1 and η_2.

Which in the fact now, can be seen simply that this is a dimension reduction regression models ZIP or reduced-rank zero-inflated Poisson model (RR-ZIP). RR-ZIP is given by

$$logit\ \omega = \eta_1 = \beta_{(1)1} + \alpha_{(1)1}.\eta_2 \tag{19}$$

$$log\ \mu = \eta_2 = \beta_2^T X \tag{20}$$

with $\beta_{(1)1}$ and $\alpha_{(1)1}$ are coefficients who want predictable.

With (19) and (20) the RR-ZIP model of rank=1 has $H_1 = I_2$ and $H_2 = \ldots = H_p = \left(\alpha_{(1)1}\right)^T$. There is a trivial complication that the constraint angle (can use other constraints) imposed on parameters that are used instead of the first two. This can be simplified if the order parameter exchanged.

4.3.2 An Application of ZIP of RCIM for analysis of the GEI

Table 6 is the dataset of Hadi and Sa'diyah (2014). In this table we focused on the endurance to leaf rust disease. The cells is the number of crops attacked by leaf-rust observed in three replications. A genotype with large numbers indicate the most vulnerability, and in vice verse, the smaller number the better endurance. The zeros on the observation in Probolinggo sometimes called *escape* observation, where all the columns on this row are zero. The ZIP model relies on the assumption that zero are as structural and random ones. The ZIP model will provide us the probability of the cell to be zero, and the fitted value for Poisson count, as well.

Table 6: The 2^{nd} data set: Count of Leaf Rust Disease Attacks on Mung Bean

Genotype	Environments				
	Proboliggo	Jember	Jombang	Bolo	Rasanae
MLG1002	0	167	100	150	150
MLG1004	0	217	250	233	250
MLG1021	0	200	217	183	217
MMC74dkp1	0	133	200	183	133
MMC71dkp2	0	200	200	233	367
MMC157dkp1	0	133	150	167	150
MMC203dkp5	0	50	100	67	83
MMC205e	0	50	67	100	67
MMC100fkp1	0	50	83	83	83
MMC87dkp5	0	83	117	133	83
MURAI	0	0	50	33	33
PERKUTUT	0	67	133	117	117

The data were analyzed using RCIM model, following Turner and Firth (2015) work on the Poisson distribution with the GAMMI model of Van Eeuwijk (1995). Determining the rank=2 model, here we use the deviance analysis rather than the log-likelihood ratio test as previously used in Hadi and Sa'diyah (2014).

Since Table 7 showed that the rank = 3 of ZIP model of RCIM does not fit the data properly (p-value is greater than 0.05), then Table 8 determine that ZIP model of rank = 2 is the best way to explain the structure of the main effects of additive and multiplicative interaction. With this rank = 2 of RCIM-ZIP model, the biplot is presented in Figure 4. It was done by run an RCIM model with rank = 0 and SVD-reparameterization on the working residuals to get the interaction visualization with rank = 2. The biplot variability is shown by the eigenvalues of matrix interaction. The first two eigenvalues, explain the total variability of the Biplot, that is 72.78%. The reader is recommended to see Hadi and Sa'diyah (2014) in order to get more interpretation information about the GEI analysis of the biplot of Figure 4.

Table 7: The Deviance Analysis to test rank = 3 ZIP model of RCIM

Source	df	Deviance	Mean Deviance	Ratio of Mean Deviance	p-value
Main Effects (Rank =0)	14	198.8616	14.2044	29.49009463	2.5013E-05
RCIM Rank = 1	14	64.9406	4.6386	9.63033808	0.001572204
RCIM Rank = 2	12	19.7904	1.6492	3.42394357	0.044636088
RCIM Rank = 3	10	0.3172	0.0317	0.06585450	0.999882848
Error	8	3.8533	0.4817		
Total	58	287.7631			

Table 8: The Deviance Analysis to test rank = 2 ZIP model of RCIM

Source	df	Deviance	Mean Deviance	Ratio of Mean Deviance	p-value
Main Effects (Rank =0)	14	198.8616	14.2044	61.30610867	2.9137E-12
RCIM Rank = 1	14	64.9406	4.6386	20.02023257	3.80108E-08
RCIM Rank = 2	12	19.7904	1.6492	7.11793771	0.000125675
Error	18	4.1705	0.2317		
Total	58	287.7631			

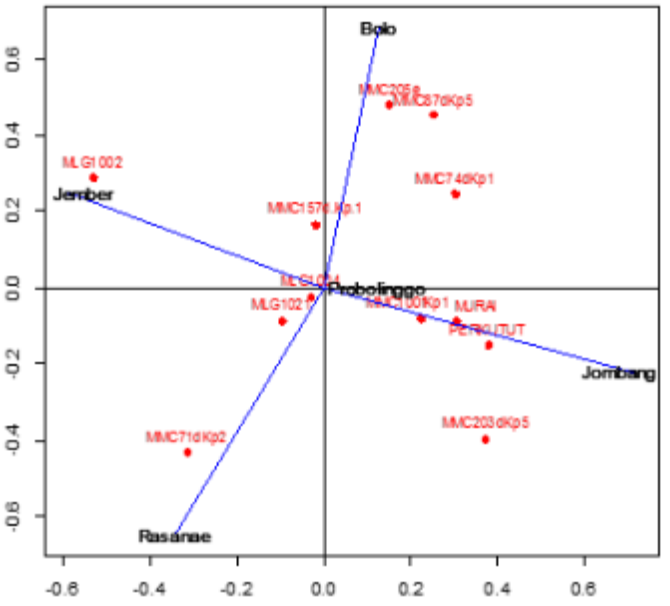

Figure 4: Biplot of the interaction effect on Log-scale of Zero-inflaated Poisson

4.3.3 Comparing a ZIP to a Poisson model in RCIM

Comparing ZIP model with Poisson model was done by Hadi (2012) using dataset with zero-inflated of Table 6. Hadi (2012) said that Poisson model faced difficulties of computing in the estimation model of rank = 3. This failure is expected to occur due to the failure of eigenvalue calculation because there is an infinite calculation. Meanwhile, at a lower rank, the log-likelihood of ZIP model indicates that it is better than Poisson model. Obtaining the best model depends on two things. The first is the link-functions concerning the distribution of data and interpretation of the model, and the second is the decomposition of the interaction, in this case, it is determined by the rank used. The higher the rank used, the more complex model we'll get. In the more complex model, together with the distribution and its canonical link-function, the estimation parameters will paid a longer iteration of computing. Both of these works are inter-related and unseparable, especially when there is overdispersion and/or excess zero. The best fit model can be obtained by the use of distribution and proper link-function, and at the same time, also by the degree of decomposition of the interaction (rank models).

Now we will use the 1^{st} dataset of Table 1 that is modeled by a Poisson model to be compared with the ZIP model for count data with no zero problems. The RCIM ZIP able to model the structure of interaction on Poisson counting data even though it contains no zero observation at all. This capability is indicated by ZIP model, since it has very similar results to Poisson models. Table 9 shows that log-likelihood value of the ZIP model is exactly the same as Poisson model for a counting data with no zero.

Table 9: The Log-likelihood of ZIP and Poisson model of RCIM for data count with no zero

Model	The ZIP	Standard Poisson
RCIM FullRank=3	-39.04556	-39.04556
RCIM Rank=2	-39.09895	-39.09895
RCIM Rank=1	-40.99432	-40.99432
Main Effects(Rank=0)	-48.33612	-48.33612

4.3.4 RCIM with Zero Inflated Negative Binomial Distributions

In this section, we will propose the ZINB for both overdispersion and/or excess zero. Zero inflated Negative Binomial (ZINB) is one of the methods used to deal with problem of overdispersion in a case of excess zero. ZINB formed by Negative Binomial distribution, mixture of the Poisson-Gamma and excess zero. A Negative Binomial distribution has parameters μ and α, the p. m. f. of random variable Y NB distributed can be written as:

$$f_{\mathrm{NB}}(y;\mu,\alpha) = \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_i!\,\Gamma\left(\frac{1}{\alpha}\right)} \left(\frac{\frac{1}{\alpha}}{\frac{1}{\alpha}+\mu_i}\right)^{\frac{1}{\alpha}} \left(\frac{\mu_i}{\frac{1}{\alpha}+\mu_i}\right)^{y_i} \tag{21}$$

A random variable Y of ZINB distribution has a valued of zero with probability of ω and follows the NB distribution with probability of $(1-\omega)$. For $Y=0$ occur with probability of ω, then Y has p. m. f. of the form:

$$\begin{aligned} f_{\text{ZINB}}(y=0;\mu,\alpha,\omega) &= \omega + (1-\omega) f_{\text{NB}}(y=0;\mu,\alpha) \\ &= \omega + (1-\omega)\frac{\Gamma\left(0+\frac{1}{\alpha}\right)}{0!\,\Gamma\left(\frac{1}{\alpha}\right)}\left(\frac{\frac{1}{\alpha}}{\frac{1}{\alpha}+\mu}\right)^{\frac{1}{\alpha}}\left(\frac{\mu}{\frac{1}{\alpha}+\mu}\right)^{0} \\ &= \omega + (1-\omega)\frac{\Gamma\left(\frac{1}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)}\left(\frac{\frac{1}{\alpha}}{\frac{1}{\alpha}+\mu}\right)^{\frac{1}{\alpha}} \\ &= \omega + (1-\omega)\left(\frac{\frac{1}{\alpha}}{\frac{1}{\alpha}+\mu}\right)^{\frac{1}{\alpha}} \end{aligned}$$

The rest occur with probability of $(1-\omega)$, $Y = 1, 2, \ldots$, along with NB distribution:

$$\begin{aligned} f_{\text{ZINB}}(y \neq 0;\mu,\alpha,\omega) &= (1-\omega)\, f_{\text{NB}}(y \neq 0;\mu,\alpha) \\ &= (1-\omega)\frac{\Gamma\left(y+\frac{1}{\alpha}\right)}{y!\,\Gamma\left(\frac{1}{\alpha}\right)}\left(\frac{\frac{1}{\alpha}}{\frac{1}{\alpha}+\mu}\right)^{\frac{1}{\alpha}}\left(\frac{\mu}{\frac{1}{\alpha}+\mu}\right)^{y} \end{aligned}$$

And the mixture distribution of ZINB means for $\omega = 0$, the mean and variance of ZINB will be equal to mean and variance of NB:

$$\begin{aligned} E(Y) &= (1-\omega)\mu \\ E(Y) &= (1-0)\mu \\ E(Y) &= \mu \end{aligned}$$

$$\begin{aligned} Var(Y) &= E(Y)(1+\alpha\mu+\omega\mu) \\ Var(Y) &= (1-0)\mu(1+\alpha\mu+\omega\mu) \\ Var(Y) &= \mu(1+\alpha\mu) \end{aligned}$$

As described in previous section, we will apply the linear predictors of RR-VGLM directly to those three parameters of ZINB distribution. We then compare the MSE of the ZINB versus the NB. Table 10 represents the MSE of the two methods (ZINB vs NB) on an overdispersed count data with zero's problems as we get from our simulation scenario. It shows that the MSE of ZINB

model is better than NB. It is concluded that RCIM ZINB fits excess zeros data better than NB. Now we will move forward to discuss ZINB performance with respect to the presence of outliers and excess zeros at once on the counting data.

Table 11 contains the MSE of NB and ZINB model for data with severe illnesses conditions, which is (i) overdispersed counting data due to outliers and (ii) having zero problems, simultaneously. It seems that the MSE of ZINB model always smaller than NB's at any ranks. This shows that ZINB model fits the data (overdispersed with extra zeros) better than the NB model. Last but not least, we also evaluated the performance of ZINB model to count data containing structural zero of Table 6. Table 12, which represents the MSE of NB and ZINB on counting data with structural zero, shows that although the ZINB can provide smaller MSE than the NB's at low rank of RCIM (rank = 1 and 2), but at higher rank (rank = 3), NB performs better than ZINB one.

Table 10: The MSE of Negative Binomial (NB) and ZINB on overdispersed data with zero

Model	MSE NB	MSE ZINB
RCIM rank = 1	0.20876010	0.07589846
RCIM rank = 2	0.09743042	0.07504630
RCIM rank = 3	0.10276670	0.02441078
RCIM rank = 4	0.05146207	0.01666468
RCIM rank = 5	0.02056970	0.00899084
RCIM rank = 6	0.02681639	0.00535431

Table 11: The MSE of Negative Binomial and ZINB model on data with both outliers and extras zero (3rd scenario)

Model	MSE NB	MSE ZINB
RCIM rank = 1	0.25555470	0.08142671
RCIM rank = 2	0.14247590	0.07504630
RCIM rank = 3	0.07846524	0.05165929
RCIM rank = 4	0.05101145	0.05012614
RCIM rank = 5	0.02153536	0.00899084
RCIM rank = 6	0.01880404	0.00535431

Table 12: The MSE of Negative Binomial and ZINB model on count data with structural zero

Model	MSE NB	MSE ZINB
RCIM rank = 1	0.02549026	0.02143623
RCIM rank = 2	0.01268638	0.01260851
RCIM rank = 3	9.86e-06	0.01209246

Table 13: The MSE of Negative Binomial and ZINB model on data with structural zero and also outliers at once (4th scenario)

Model	MSE NB	MSE ZINB
RCIM rank = 1	0.02187914	0.03839160
RCIM rank = 2	0.02145327	0.05245484
RCIM rank = 3	1.06e-05	0.01474227

Table 13 presents the MSE of ZINB models for data containing both structural zero and outliers at once. It seems that it is similar to Table 12, that ZINB better than NB at the lower rank model of RCIM, but less at higher. It's clear that ZINB can be relied upon to model the data with structural zero with the simplest interaction terms of RCIM. If we take a look at both Table 12 and 13 across rows at the same column, we can see that when outliers come to data with structural zeros, the MSE of both NB and ZINB will increase slightly.

5. Concluding Remark

Here we conclude that in multiplicative modeling, the overdispersion problems of counting data can be handled in two ways. First by choosing the canonical link function of the distributional data. With the same rank of complexity the NB can do better than usual Poisson.

The second, we make model the overdispersion by involving some more multiplicative terms in our model. The standard Poisson model can fit the overdispersed data properly with more complex multiplicative model than the NB one.

The ZIP model may overcome the problem of zero inflated, including the counting data with structural zero. The ZIP give us the same result as Poisson for the data with no zero. We proposed the ZINB model for overdispersed counting data with excess zero, structural zero or containing outliers. The ZINB model result better fitted value for those counting data problems, by the smaller MSE in multiplicative modeling.

Acknowledgement

We thank to Dimas, Graduated School of Mathematics, The University of Jember for LaTeX helping. We also thank to Arif Musaddad, ILETRI, Malang for early communication to get this joint paper writing partnership and thanks

to Thomas W. Yee, Department of Statistics, University of Auckland for the VGAM package and special thanks for his kindness during Hadi's previous visiting research.

References

Goodman, L. A. (1981). Association models and canonical correlation in the analysis of cross-classifications having ordered categories. *Journal of the American Statistical Association*, 76:320–334.

Hadi, A. F. (2012). *The Developing of Robustness on Additive Main Effect - Multiplicative Interaction Models (AMMI)*. PhD thesis, Bogor Agricultural University, Bogor, Indonesia.

Hadi, A. F., Mattjik, A. A., and Sumertajaya, I. M. (2010). Generalized ammi models for assessing the endurance of soybean to leaf pest. *Jurnal Ilmu Dasar*, 11:123–131.

Hadi, A. F. and Sa'diyah, H. (2014). Row-column interaction models for zero-inflated poisson count data in agricultural trial. *Proc. ICCS-13, Bogor, Indonesia*, 27:233 – 244. www.isoss.net/downloads/Prociccs13.pdf.

Hadi, A. F. and Sa'diyah, H. (2016). An approach of row-column interaction models (rcim) for generalized ammi models with deviance analysis. *Agriculture and Agricultural Science Procedia*, 9:134–145. http://dx.doi.org/10.1016/j.aaspro.2016.02.108.

Hilbe, J. M. (2011). *Negative Binomial Regression Second Edition*. Cambridge University Press, New York, NY.

Pawitan, Y. (2001). *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford, Ireland : Clarendon Press.

Turner, H. and Firth, D. (2015). *Generalized nonlinear models in R: An overview of the gnm package*. R package version 1.0-8. http://CRAN.R-project.org/package=gnm.

Van Eeuwijk, F. (1995). Multiplicative interaction in generalized linear models. *Biometrics*, 51:1017–1032.

Yee, T. W. (2008). *VGAM Family Functions for Positive, Zero-altered and Zero-Inflated Discrete Distributions*. University of Auckland, New Zealand, NZ. http://www.stat.auckland.ac.nz/.

Yee, T. W. (2010). The vgam package for categorical data analysis. *Journal of Statistical Software*, 32:1–34. http://www.jstatsoft.org/v32/i10.

Yee, T. W. (2014). Reduce-rank vector generalized linear models with two linear predictors. *Comput Stat Data Anal*, 71:889 – 902.

Yee, T. W. (2015). *Vector Generalized Linear and Additive Models: With an Implementation in R*. Springer, New York, USA.

Yee, T. W. and Hadi, A. F. (2014). Row column interaction models, with an r implemetation. *Computational Statistics*, 29:1427–1445. http://dx.doi.org/10.1007/s00180-014-0499-9.

Yee, T. W. and Hastie, T. J. (2003). Reduced-rank vector generalized linear models. *Statistical Modelling*, 3:15–41.