



**PERBANDINGAN REGRESI *ZERO INFLATED POISSON* (ZIP) DAN  
REGRESI *ZERO-INFLATED NEGATIVE BINOMIAL* (ZINB) PADA  
KASUS TANAH LONGSOR DI KOTA SEMARANG**

**JURNAL ILMIAH**

Diajukan sebagai salah satu syarat memperoleh gelar Sarjana Statistika

Oleh :

**M Salman Al Farisyi**

**B2A016031**

**PROGRAM STUDI S1 STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS MUHAMMADIYAH SEMARANG  
2020**

**HALAMAN PENGESAHAN**

Skripsi dengan judul “PERBANDINGAN REGRESI *ZERO INFLATED POISSON* (ZIP) DAN REGRESI *ZERO-INFLATED NEGATIVE BINOMIAL* (ZINB) PADA KASUS TANAH LONGSOR DI KOTA SEMARANG” yang disusun oleh :

Nama : M Salman Al Farisyi

NIM : B2A016031

Program Studi : S1 STATISTIKA

Telah disetujui oleh dosen pembimbing pada 7 Mei 2020

Pembimbing Utama



Indah Manfaati Nur, M.Si  
NIK. 28.6.1026.221

Pembimbing Pendamping



M Al Haris, M.Si  
NIP.CP. 1026.070



Mengetahui  
Ketua Program Studi

Indah Manfaati Nur, M.Si  
NIK. 28.6.1026.221

**SURAT PERNYATAAN  
PUBLIKASI KARYA ILMIAH**

Yang bertanda tangan di bawah ini, saya :

Nama : M Salman Al Farisyi  
NIM : B2A016031  
Fakultas/Jurusan : FMIPA/S1 Statistika  
Jenis Penelitian : Skripsi  
Judul : PERBANDINGAN REGRESI *ZERO INFLATED POISSON* (ZIP) DAN REGRESI *ZERO-INFLATED NEGATIVE BINOMIAL* (ZINB) PADA KASUS TANAH LONGSOR DI KOTA SEMARANG  
Email : [farisysalman98@gmail.com](mailto:farisysalman98@gmail.com)

Dengan ini menyatakan bahwa saya menyetujui untuk :

1. Memberikan hak bebas royalti kepada Perpustakaan UNIMUS atas penulisan karya ilmiah saya, demi pengembangan ilmu pengetahuan
2. Memberikan hak menyimpan, mengalih mediakan/mengasihkan formatkan, mengelola dalam bentuk pangkalan data (database), mendistribusikannya, serta menampilkannya dalam bentuk *softcopy* untuk kepentingan akademis kepada Perpustakaan UNIMUS, tanpa perlu izin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta
3. Bersedia dan menjamin untuk menanggung secara pribadi tanpa melibatkan pihak Perpustakaan UNIMUS dari semua bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini.

Demikian pernyataan ini saya buat dengan sesungguhnya dan semoga dapat digunakan sebagaimana mestinya.

Semarang, 7 Mei 2020  
Yang membuat pernyataan,



M Salman Al Farisyi  
(B2A016031)

# PERBANDINGAN REGRESI ZERO- INFLATED POISSON (ZIP) DAN REGRESI ZERO-INFLATED NEGATIVE BINOMIAL (ZINB) PADA KASUS TANAH LONGSOR DI KOTA SEMARANG

M Salman Al Farisyi <sup>(1)</sup>, Indah Manfaaati Nur<sup>2)</sup> dan M Al Haris <sup>(3)</sup>

<sup>(1,2,3)</sup> Program Studi Statistika Universitas Muhammadiyah Semarang

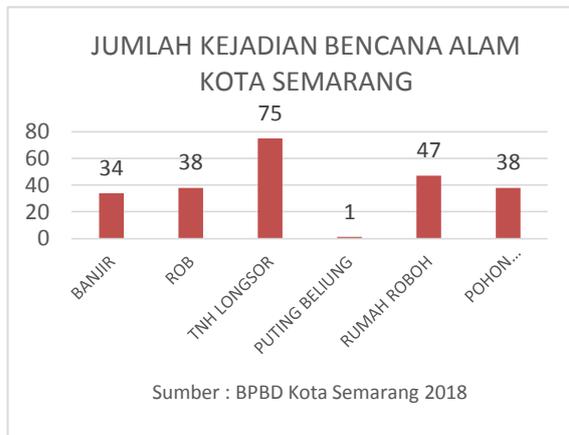
Email : [farisysalman98@gmail.com](mailto:farisysalman98@gmail.com)

Article history	Abstract
Submission :	<i>Poisson regression is a nonlinear regression which is often used to count data and has equidispersion assumption (variance value equal to mean value). However in practice, the assumption of equidispersion is often not violated. One of the violations is overdispersion (variance value equal to mean value). The cause of overdispersion is the number of zero values in the response variabel (excess zeros). There are many methods to handle the problem of overdispersion because of excess zeros. Two of them are Zero-Inflated Poisson (ZIP) regression and Zero-Inflated Negative Binomial (ZINB) regression. The purpose of this research is to determine which regression model is better for handling the overdispersion data. Data that can be analyzed using the ZIP and ZINB regression is landslide case data in Semarang City. Landslide rate data has a proportion of zero values of more than 50% on the response variabel. In this research estimates the maximum likelihood method, followed by testing the suitability of the model using the likelihood ratio test and testing the significance of the parameters with the Wald Test and selecting the best model using the Akaike Information Criterion (AIC) by taking the smallest AIC value. The software used in this study is the R 3.6.1 program. Based on the analysis obtained the best model results for the case of landslides in the city of Semarang, namely the ZIP regression model and the significant variable is the height of the region. With the AIC value resulting in a value of 63,715%</i>
Revised :	
Accepted :	
<b>Keyword:</b> <i>Overdispersion, Poisson Regression, ZIP Regression, ZINB Regression, Landslide Cases in Semarang City</i>	

## PENDAHULUAN

Bencana alam adalah bencana yang diakibatkan oleh peristiwa atau serangkaian peristiwa yang disebabkan oleh alam antara lain berupa gempa bumi, tsunami, gunung meletus, banjir, kekeringan, angin topan, dan tanah longsor. Bencana alam yang terjadi di Jawa Tengah khususnya di Kota Semarang pada tahun 2018 ini begitu memprihatinkan dan mengundang perhatian dari seluruh masyarakat Kota Semarang. Kejadian bencana alam yang

pernah terjadi di Kota Semarang dapat dilihat pada gambar 1.1 berikut:



Pada grafik tersebut memberikan informasi bahwa Kota Semarang merupakan daerah yang rentan terhadap bencana, salah satunya yang tertinggi adalah bencana tanah longsor. Secara geologi Kota Semarang memiliki 2 topografi yaitu dataran rendah dan dataran tinggi. Dimana sebagian besar bencana longsor terjadi pada dataran tinggi dengan daerah perbukitan yang bergelombang.

Hasil Penelitian terdahulu oleh (Nugrogo dkk dalam Faizana, 2015) menunjukkan bahwa faktor ketinggian wilayah, kepadatan penduduk, penggunaan lahan dan curah hujan berhubungan dengan tingkat kerawanan kejadian tanah longsor. Keterkaitan faktor-faktor yang mempengaruhi terjadinya tanah longsor dapat didekati dengan analisis regresi. Hubungan antara variabel respon dan variabel bebas dalam persamaan regresi dapat berbentuk hubungan linier dan non-linier. Hubungan linear dinyatakan dalam persamaan regresi linier.

Persamaan regresi linear digunakan untuk menganalisis variabel respons yang berupa peubah acak kontinu dan mengikuti distribusi normal, namun banyak ditemukan variabel respons yang tidak berdistribusi normal dan tidak linear dalam parameter. Untuk mengatasi hal tersebut dikembangkan Generalized Linear.

Model (GLM) digunakan sebagai perluasan model regresi umum dengan variabel responsnya berdistribusi keluarga eksponensial, meliputi distribusi normal, binomial, Poisson, binomial negatif, eksponensial, gamma, dan invers normal. Jika variabel respons yang digunakan merupakan peubah acak diskret yang berdistribusi Poisson, maka dapat digunakan model regresi Poisson untuk pembentukan model regresi. Analisis data menggunakan regresi Poisson harus memenuhi asumsi seperti nilai varians dan rata-rata dari variabel respons

tersebut sama atau *equidispersi* [1] Pada kenyataannya tidak sepenuhnya asumsi tersebut terpenuhi, seperti nilai varian lebih besar dari nilai rata – ratanya yang disebut *overdispersi*. *Overdispersi* yang disebabkan oleh banyaknya nilai nol yang berlebih pada variabel respon (*excess zeros*) pada dasarnya tetap dapat disetiasi menggunakan regresi poisson. Tetapi untuk data yang terdapat banyak nilai nol perlu adanya metode tertentu untuk mengatasinya, diantaranya adalah metode regresi *Zero-Inflated Poisson* (ZIP) dan metode regresi *Zero-Inflated Negative Binomial* (ZINB).

Studi kasus yang digunakan pada penelitian ini adalah jumlah Kasus Kejadian Tanah Longsor di Kota Semarang tahun 2018. Bencana longsor merupakan salah satu bencana alam yang sering mengakibatkan kerugian harta benda maupun korban jiwa dan menimbulkan kerusakan sarana dan prasarana lainnya yang bisa berdampak negatif pada kondisi ekonomi dan sosial. Peristiwa tanah longsor atau dikenal dengan gerakan massa tanah, batuan atau kombinasinya, sering terjadi pada lereng alami atau lereng non alami dan sebenarnya merupakan fenomena alam, yaitu alam mencari keseimbangan baru akibat adanya gangguan atau faktor yang mempengaruhi dan menyebabkan terjadinya pengurangan kuat geser serta peningkatan tegangan geser tanah. Data menunjukkan terjadi peristiwa tanah longsor tahun 2018 sebanyak 75 kasus di 16 Kecamatan. Proporsi nilai nol pada variabel responsnya lebih besar dibandingkan nilai lainnya. Angka tersebut menunjukkan bahwa kecamatan yang tidak terjadi kasus tanah longsor banyak atau nol kasus.

## TINJAUAN PUSTAKA

### Tanah Longsor

Tanah longsor adalah, Gerakan massa tanah pada lereng-lereng alam. Bisa dikatakan bencana apabila terjadi pada daerah yang dihuni manusia.(Rudiyanto,2016).

### Ketinggian Wilayah

Ketinggian suatu daerah dapat menentukan besarnya kekuatan tanah yang terjatuh apabila terjadi longsor. Semakin tinggi suatu tempat semakin besar kekuatan tanah yang terjatuh karena adanya pengaruh gravitasi (Noor,2005)

## Penggunaan lahan

Penggunaan lahan mempunyai pengaruh besar terhadap kondisi air tanah, hal ini akan mempengaruhi kondisi tanah dan batuan yang pada akhirnya juga akan mempengaruhi keseimbangan lereng. (Rudiyanto, 2016)

## Kepadatan Penduduk

Kepadatan penduduk adalah perbandingan antara jumlah penduduk dengan luas wilayah yang didiami penduduk ( $\text{Jiwa}/\text{km}^2$ ). Peraturan Kepala BNPB Nomor 02 (2012 : 29) (Ni'mah, 2107).

## Curah Hujan

Semakin tinggi nilai curah hujannya, maka sudah dapat dipastikan bahwa wilayah tersebut merupakan wilayah yang mempunyai potensi tertinggi terjadi bencana tanah longsor. (Rudiyanto, 2016)

## Multikolinearitas

$$VIF_j = \frac{\text{Var}(\hat{\beta}_j)}{\sigma^2} = r^{jj}$$

Nilai  $VIF_j$  yang bernilai 1 menunjukkan bahwa variabel-variabel independent tidak saling berkorelasi, jika nilainya lebih dari 10 menunjukkan adanya kolinieritas antara variabel-variabel independen. (Hocking, 1996).

## Regresi Poisson

Regresi poisson merupakan analisis regresi nonlinier dari distribusi poisson, dimana analisis ini sangat cocok digunakan dalam menganalisis data diskrit (count).  $\mu_i$  adalah rata-rata jumlah kejadian yang terjadi dalam interval waktu tertentu. Dalam regresi poisson terdapat asumsi yang harus dipenuhi yaitu variabel respon (Y) diskrit dan asumsi *equidispersi*. *Equidispersi* yaitu nilai rata-rata sama dengan nilai varian atau  $V(Y) = E(Y) = \mu$ . Tetapi pada kenyataannya sering terjadi pelanggaran asumsi *equidispersi*. Yaitu terjadinya *Overdispersi* (varians > rata-rata) atau *Underdispersi* (varians < rata-rata).

## Overdispersi

Nilai varians > Mean. Penyebabnya terdapat Excess Zeros ( nilai nol berlebih pada variable Y). Ada atau tidaknya overdispersi dapat dilihat dari nilai Deviance atau Pearson Chi-square yang dibagi dengan derajat bebasnya. Apabila > 1 maka menunjukkan nilai variansi

yang lebih besar daripada rata-rata.

## Model Regresi ZINB

Model yang terbentuk dari distribusi campuran poisson dan gamma. Model regresi ZINB dibagi menjadi dua komponen yaitu model data *diskret* untuk  $\mu_i$  dan model *zero-inflation* untuk  $p_i$  yaitu

$$\ln(\mu_i) = x_i^T \beta, \quad \mu_i \geq 0, i = 1, \dots, n$$

$$\text{logit}(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) = z_i^T \gamma, \quad 0 \leq p_i \leq 1, i = 1, \dots, n$$

## Estimasi Parameter

Penaksiran parameter model regresi ZIP dan ZINB dilakukan dengan metode penaksir kemungkinan maksimum (Jansakul & Hinde, 2002). Penaksiran parameter dengan metode penaksir kemungkinan maksimum dihitung dengan memaksimalkan fungsi loglikelihoodnya (MLE). Penjumlahan fungsi log likelihood tidak dapat diselesaikan dengan metode numerik biasa, oleh karena itu dapat digunakan algoritma EM (*Expectation Maximization*) (Garay & Hashimoto, 2011).

## Uji Kesesuaian Model

Pengujian kesesuaian model regresi ZIP dan ZINB dapat menggunakan nilai Likelihood Ratio (LR) (Hilbe, 2011). Hipotesis uji rasio likelihood adalah:

$$G = -2 \ln \left[ \frac{L(y; \hat{\omega})}{L(y; \hat{\Omega})} \right]$$

Kriteria Uji : Tolak  $H_0$  pada taraf signifikansi  $\alpha$  jika  $G_{\text{hitung}} > \chi^2_{(v,\alpha)}$ , dimana  $v$  adalah derajat bebas.

## Uji Signifikansi Parameter

Pengujian parameter regresi ZIP dan ZINB secara parsial dapat dilalui dengan dua cara, yaitu pengujian parameter model log dan pengujian parameter model logit. Statistik uji yang digunakan adalah dengan menggunakan uji Wald (Myers et al., 2010) sebagai berikut:

$$W_i = \left( \frac{\hat{\beta}_i}{\text{SE}(\hat{\beta}_i)} \right)^2 \quad W_i = \left( \frac{\hat{\gamma}_i}{\text{SE}(\hat{\gamma}_i)} \right)^2$$

Tolak  $H_0$  pada taraf signifikansi  $\alpha$  jika  $W_i > X_{(a,1)}^2$  pada taraf alfa.

#### Pemilihan Model Terbaik

Pemilihan model terbaik pada regresi ZIP dan ZINB dapat dilihat nilai AIC (Akaike's Information Criterion) (Sharma & Landge,(2013) :

$$AIC = -2 \ln L(\hat{\theta}) + 2k$$

dimana  $L(\hat{\theta})$  adalah nilai likelihood, dan  $k$  adalah jumlah parameter. Model terbaik Regresi ZIP dan ZINB adalah model dengan nilai AIC terkecil.

### METODE PENELITIAN

#### Sumber Data

Data yang digunakan adalah data sekunder. Data bersumber dari BPBD Kota Semarang, BAPPEDA Kota Semarang, BPS Kota Semarang, dan BMKG kota Semarang tahun 2018.

#### Variabel Penelitian dan Struktur Data

Variabel penelitian yang digunakan dalam penelitian ini terdiri dari variabel respon (Y) yaitu jumlah kejadian bencana tanah longsor. Sedangkan untuk variabel prediktor (X) dipilih sebanyak empat variabel yaitu ketinggian wilayah (X1), kepadatan penduduk (X2), persentase penggunaan lahan (X3), dan curah hujan (X4).

#### Struktur Data Penelitian

Kecamatan	Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
1	Y <sub>1</sub>	X <sub>11</sub>	X <sub>21</sub>	X <sub>31</sub>	X <sub>41</sub>
2	Y <sub>2</sub>	X <sub>12</sub>	X <sub>22</sub>	X <sub>32</sub>	X <sub>42</sub>
:	:	:	:	:	:
16	Y <sub>16</sub>	X <sub>116</sub>	X <sub>216</sub>	X <sub>316</sub>	X <sub>416</sub>

#### Langkah-langkah Penelitian

Metode analisis yang digunakan dalam penelitian ini adalah *Zero-Inflated Negative Binomial (ZINB)* dan *Zero-Inflated Poisson (ZIP)*. Secara skematik, langkah-langkah yang dilakukan adalah sebagai berikut :

1. Menguji kolinieritas

Uji kolinieritas dapat menggunakan nilai VIF dan korelasi

2. Melakukan analisis regresi *Poisson*
3. Menentukan *Overdispersion*
4. Perbandingan Model Regresi ZIP dan ZINB menggunakan Estimasi
5. Melakukan uji kesesuaian model regresi ZIP dan ZINB menggunakan *Likelihood Ratio (LR)* test.
6. Melakukan uji signifikansi pada model regresi ZIP dan ZINB secara parsial menggunakan uji wald.
7. Penentuan metode terbaik dengan nilai AIC terkecil
8. Selesai

### HASIL PENELITIAN DAN PEMBAHASAN

#### Analisis Deskriptif

Berikut adalah analisis deskriptif variable-variabel yang diduga berpengaruh terhadap jumlah kasus kejadian tanah longsor apabila disajikan dalam tabel.

Variabel	Mean	Varians	St. Deviasi	Minimum	Maksimum
Y	2.81	8.696	2.949	0	9
X1	76.625	14742.550	121.419	0	311
X2	6971.25	16126114.20	4015.733	1053	12307
X3	61.022	732.048	27.0564	16.97	98.75
X4	236.863	1394.923	37.349	153.3	313.2

terlihat bahwa variabel kepadatan penduduk varians nya sangat tinggi, dapat dilihat dari nilai maksimum variabel kepadatan penduduk yang sangat tinggi dibandingkan dengan nilai variabel lainnya, dapat dikatakan untuk variabel kepadatan penduduk memiliki nilai varians tertinggi diantara variabel lainnya, artinya variabel kepadatan penduduk mempengaruhi jumlah kasus kejadian Tanah longsor. Nilai variansi yang tinggi bisa di artikan bahwa data yang digunakan untuk variable kepadatan penduduk sangat fluktuatif.

Pemeriksaan Kolinieritas Pada Variabel Independen

Variabel ke-j	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
VIF	1.725	8.078	8.263	1.104

KET: Y: Kasus kejadian tanah longsor  
 X1 : Ketinggian wilayah  
 X2 : Kepadatan Penduduk  
 X3 : Persentase Penggunaan Lahan  
 X4 : Curah Hujan

Jika nilai VIF<sub>j</sub> bernilai 1 maka variabel-variabel independent tidak saling berkorelasi. Nilai VIF<sub>j</sub> yang lebih besar dari 10 menunjukkan adanya kolinieritas antara variabel-variabel independent. Berdasarkan tabel diatas dapat disimpulkan bahwa tidak adanya kolinieritas antar variabel independent karena nilai VIF kurang dari 10.

Pengujian Distribusi Poisson pada Variabel Dependen

Jumlah Sampel	Nilai Asymp. Sig
16	0.084

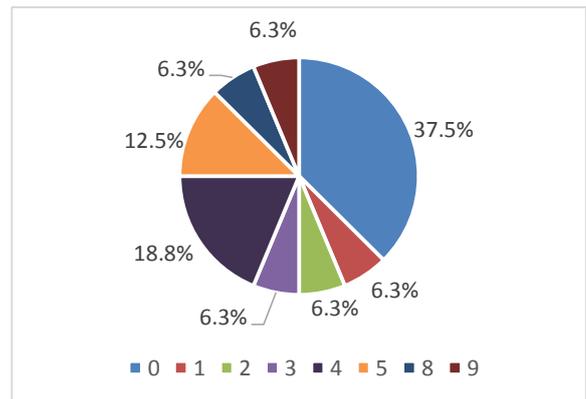
Kriteria pengujian dalam uji Kolmogorof-Smirnov adalah tolak H<sub>0</sub> jika α pada taraf signifikansi atau nilai signifikansi. Nilai T untuk uji Kolmogorof-Smirnov adalah 0.084 sehingga gagal tolak H<sub>0</sub> yang artinya variabel dependen atau variabel respon berdistribusi Poisson.

Overdispersion Regresi Poisson

Kriteria	Db	Nilai	Nilai/db
Deviance	11	44.162	4.0147
Pearson Chi-Square	11	49.8619	4.5329

Terlihat bahwa nilai Deviance/db sebesar 4.0147 dan nilai Pearson Chi-Square/db sebesar 4.5329, kedua nilai ini lebih besar dari 1 sehingga dapat disimpulkan bahwa pada data respon (Y) terjadi kasus overdispersion.

Pengujian Excess Zeros pada variable Dependen



Berdasarkan diagram diatas dapat dilihat bahwa nilai nol yang ada pada variabel respon memiliki proporsi tertinggi dibandingkan dengan nilai lainnya. Presentasi nilai nol yang ada memiliki nilai sebesar 37.5 %. Sehingga dapat disimpulkan bahwa variabel respon mengalami excess zero

Pemodelan Regresi ZIP

Salah satu metode yang dapat digunakan untuk mengatasi *overdispersi* pada Regresi Poisson adalah Regresi ZIP, karena Regresi ZIP mampu mengendalikan *overdispersi* dalam distribusi Poisson dan inflasi nilai 0 sehingga akurasi estimasi parameter dapat terjamin (Kartiningrum, E dan Saidah, N,2014).

	Estimate	Std. Error	z	Pr(> z )
(Intercept)	1.618	0.165	9.799	0.000
X <sub>1</sub>	0.302	0.169	1.785	0.074*
X <sub>2</sub>	0.542	0.483	1.124	0.261
X <sub>3</sub>	0.037	0.474	0.078	0.938
X <sub>4</sub>	0.118	0.144	0.816	0.415

Model Diskrit

$$\mu_i = 1.618 + 0.302x_1 + 0.542x_2 + 0.037x_3 + 0.118x_4, \mu_i \geq 0, i = 1,2$$

Model *Zero-Inflation* untuk  $\omega$  yaitu

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-4.971	8.740	-0.569	0.570
X <sub>1</sub>	-5.346	14.193	-0.377	0.706
X <sub>2</sub>	1.175	2.025	0.580	0.562
X <sub>3</sub>	2.548	2.728	0.934	0.350
X <sub>4</sub>	2.187	2.243	0.975	0.330

$$\omega = \frac{-4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4}{1 - 4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4}$$

Dimana  $p_i$  = peluang resiko tidak terjadinya kejadian tanah longsor  
 Model logit regresi ZIP menjelaskan bahwa peluang variable dependen bernilai nol dipengaruhi oleh Ketinggian wilayah

Kesesuaian Model ZIP

Pengujian yang digunakan menggunakan *Likelihood Ratio* (LR) dengan prosedur pengujiannya adalah sebagai berikut :

$$G = -2((-32.787) - (-21.857)) = 21.86$$

Dengan nilai p-value sebesar 0.0059

Hasil analisis diatas didapatkan nilai G hitung = 21.86. Sedangkan tabel chi- squares  $X^2(\alpha;db) = X^2(0.1;4)$  diperoleh nilai = 7.779. Karena nilai G hitung >  $X^2(\alpha;db)$  (21.86 > 7.779) atau p-value <  $\alpha$  (0.0059 < 0.1) maka tolak H<sub>0</sub>. Dapat disimpulkan bahwa setidaknya ada satu satu  $\beta_j \neq 0$  atau  $\gamma_j \neq 0$ , sehingga metode regresi *Zero- Inflated Poisson* dapat digunakan.

Pengujian signifikansi parameter regresi ZIP secara Individu

Pengujian parameter individu pada model diskrit dan *Zero-Inflation* dilakukan dengan uji Wald, dengan prosedurnya sebagai berikut :

- Hipotesis :  
 $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$  (Koefisien tidak signifikan)  
 $H_1: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 \neq 0$  (Koefisien Signifikan)

• Statistik Uji

$$W_i = \left( \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \right)^2 \quad W_i = \left( \frac{\hat{\gamma}_i}{SE(\hat{\gamma}_i)} \right)^2$$

Dengan kriteria uji : Tolak H<sub>0</sub> jika  $w_i > X^2_{\alpha;1}$  atau p-value <  $\alpha$  (0.1). Dengan nilai tabel Chi-square  $X^2_{\alpha;1}$  sebesar 2.705 pada taraf signifikansi  $\alpha$  (0.1).

Pengujian signifikansi parameter  $\beta$  (model diskrit)

	Estimate	Z <sub>i</sub>	Z <sub>i</sub> <sup>2</sup> Z <sub>i</sub> = w <sub>i</sub>	P-value	Keterangan
Intercept	1.618				
$\beta_1$	0.302	1.785	3.186	0.074*	Signifikan
$\beta_2$	0.542	0.078	0.006	0.938	Tidak Signifikan
$\beta_3$	0.037	1.124	1.263	0.261	Tidak Signifikan
$\beta_4$	0.118	0.816	0.666	0.415	Tidak Signifikan

didapatkan nilai nilai  $w_1 = 3.186$  dengan p-value = 0.074. karena nilai  $w_1 = 3.186 > 2.705$  ( $X^2_{\alpha;1}$ ) atau nilai p-value 0.074 < 0.10 ( $\alpha$ ), maka dapat disimpulkan bahwa parameter  $\beta_1$  signifikan terhadap model. Sedangkan pada parameter  $\beta_2, \beta_3,$  dan  $\beta_4$  didapatkan nilai  $w_2 = 0.006, w_3 = 1.263, w_4 = 0.666$  dengan p-value masing-masing 0.938, 0.261 dan 0.415 > 0.10, maka dapat disimpulkan bahwa parameter  $\beta_2, \beta_3,$  dan  $\beta_4$  tidak signifikan terhadap model. Sehingga model diskrit yang dihasilkan dari ZIP yaitu :

$$\mu_i = 1.618 + 0.302x_1 + 0.542x_2 + 0.037x_3 + 0.118x_4, \mu_i \geq 0, i = 1, 2$$

Berdasarkan model diskrit, rata-rata banyaknya kasus kejadian tanah longsor akan bertambah sebesar  $\exp(0.302)=1.352$  kali dari banyaknya kasus semula untuk setiap penambahan satu persen ketinggian wilayah (X<sub>1</sub>), jika variabel lain dianggap konstan

Pengujian signifikansi parameter  $\gamma$  (*Zero-Inflation*)

	Estimate	Z <sub>i</sub>	Z <sub>i</sub> <sup>2</sup> Z <sub>i</sub> = w <sub>i</sub>	P-value	Keterangan
Intercept	-4.971				
$\gamma_1$	-5.346	-0.569	0.324	0.706	Tidak Signifikan
$\gamma_2$	1.175	-0.377	0.142	0.562	Tidak Signifikan
$\gamma_3$	2.548	0.580	0.336	0.350	Tidak Signifikan
$\gamma_4$	2.187	0.975	0.951	0.330	Tidak Signifikan

Berdasarkan tabel diatas, pada parameter  $\gamma$  didapatkan nilai nilai  $w_1 = 0.324$ ,  $w_2 = 0.142$ ,  $w_3 = 0.336$ ,  $w_4 = 0.951$  bernilai  $< 2.705$  ( $X^2\alpha:1$ ) atau nilai p-value masing-masing  $0.706, 0.562, 0.350, 0.330 > 0.10$  maka dapat disimpulkan bahwa parameter  $\gamma_1, \gamma_2, \gamma_3, \gamma_4$  tidak signifikan terhadap model. Model *Zero-Inflation* yang dihasilkan yaitu :

$$\omega = \frac{-4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4}{1 - 4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4}$$

$$0 \leq p \leq 1, i = 1, \dots, n$$

Dimana  $pi$  = peluang resiko tidak terjadinya kejadian tanah longsor

Interpretasi yang diperoleh setiap penambahan 1 persen ketinggian suatu wilayah wilayah maka akan meningkatkan peluang jumlah kasus kejadian tanah longsor sebesar  $\exp(5.346)=209.78$  kali dari kasus tanah longsor semula, jika variabel bernilai konstan.

Hasil kesimpulan menunjukkan bahwa keseluruhan pendugaan parameter  $\gamma$  yaitu model *Zero-Inflation* state pada tabel diatas tidak ada variabel predictor yang berpengaruh signifikan terhadap kasus kejadian tanah longsor di Kota Semarang pada tahun 2018. Sehingga model *Zero-Inflation* pada kasus kejadian tanah longsor menunjukkan hasil yang tidak baik dan tidak diinterpretasikan pada metode ZIP. Hal ini mengidentifikasikan bahwa dalam model *Zero-Inflation* pada ZIP belum mampu menangani nilai nol yang ada pada kasus kejadian tanah longsor di Kota Semarang tahun 2018.

#### Pemodelan Regresi ZINB

##### Model Diskrit

$$\log(\mu) = \exp(1.618 + 0.302x_1 + 0.542x_2 + 0.037x_3 + 0.118x_4)$$

##### Model *Zero-Inflation*

$$pi = \frac{\exp(-4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4)}{1 + \exp(-4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4)}$$

#### Kesesuaian Model ZINB

$$G = -2((-32.611) - (-21.857))$$

$$= 21.508$$

Dengan nilai p-value sebesar 0.0059

Hasil analisis diatas didapatkan nilai G hitung = 21.508. Sedangkan tabel chi- squares

$X^2(\alpha;db) = X^2(0.1;4)$  diperoleh nilai = 7.779. Karena nilai G hitung  $> X^2(\alpha;db)$  ( $21.508 > 7.779$ ) atau p-value  $< \alpha$  ( $0.0059 < 0.1$ ) maka tolak  $H_0$ . Dapat disimpulkan bahwa setidaknya ada satu satu  $\beta_j \neq 0$  atau  $\gamma_j \neq 0$ , sehingga metode regresi *Zero- Inflated Negative Binomial* dapat digunakan.

#### Pengujian signifikansi parameter $\beta$ (model diskrit)

Berdasarkan tabel diatas, didapatkan

	Estimate	$Z_i$	$Z_i^2 = w_i$	P-value	Keterangan
Intercept	1.618				
$\beta_1$	0.302	1.785	3.186	0.074*	Signifikan
$\beta_2$	0.542	0.078	0.006	0.938	Tidak Signifikan
$\beta_3$	0.037	1.124	1.263	0.261	Tidak Signifikan
$\beta_4$	0.118	0.816	0.666	0.415	Tidak Signifikan

nilai nilai  $w_1 = 3.186$  dengan p-value = 0.074. karena nilai  $w_1 = 3.186 > 2.705$  ( $X^2\alpha:1$ ) atau nilai p-value  $0.074 < 0.10$  ( $\alpha$ ), maka dapat disimpulkan bahwa parameter  $\beta_1$  signifikan terhadap model. Sedangkan pada parameter  $\beta_2, \beta_3,$  dan  $\beta_4$  didapatkan nilai  $w_2 = 0.006, w_3 = 1.263, w_4 = 0.666$  dengan p-value masing-masing 0.938, 0.261 dan 0.415  $> 0.10$ , maka dapat disimpulkan bahwa parameter  $\beta_2, \beta_3,$  dan  $\beta_4$  tidak signifikan terhadap model.

Sehingga model diskrit yang dihasilkan dari ZINB yaitu :

$$\mu_i = \exp(1.618 + 0.302x_1 + 0.542x_2 + 0.037x_3 + 0.118x_4)$$

Dimana  $\mu_i$  = nilai harapan dari banyaknya kejadian tanah longsor ke-i

Berdasarkan model diskrit, rata-rata banyaknya kasus kejadian tanah longsor akan bertambah sebesar  $\exp(0.302)=1.352$  kali dari banyaknya kasus semula untuk setiap penambahan satu persen ketinggian wilayah ( $X_1$ ), jika variabel lain dianggap konstan

#### Pengujian signifikansi parameter $\gamma$ (*Zero-Inflation*)

	Estimate	$Z_i$	$Z_i^2 = w_i$	P-value	Keterangan
Intercept	-4.971				
$\gamma_1$	-5.346	-0.569	0.324	0.706	Tidak Signifikan
$\gamma_2$	1.175	-0.377	0.142	0.562	Tidak Signifikan
$\gamma_3$	2.548	0.580	0.336	0.350	Tidak Signifikan
$\gamma_4$	2.187	0.975	0.951	0.330	Tidak Signifikan

Pada parameter  $\gamma$  didapatkan nilai nilai  $w_1 = 0.324$ ,  $w_2 = 0.142$ ,  $w_3 = 0.336$ ,  $w_4 = 0.951$  bernilai  $< 2.705$  ( $X^2\alpha:1$ ) atau nilai p-value masing-masing  $0.706, 0.562, 0.350, 0.330 > 0.10$  maka dapat disimpulkan bahwa parameter  $\gamma_1, \gamma_2, \gamma_3, \gamma_4$  tidak signifikan terhadap model.

Model *Zero-Inflation* yang dihasilkan yaitu :

$$p_i = \frac{\exp(-4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4)}{1 + \exp(-4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4)}$$

$$0 \leq p \leq 1, i = 1, \dots, n$$

Dimana  $p_i$  = peluang resiko tidak terjadinya kejadian tanah longsor

Interpretasi yang diperoleh setiap penambahan 1 persen ketinggian suatu wilayah wilayah maka akan meningkatkan peluang jumlah kasus kejadian tanah longsor sebesar  $\exp(5.346) = 209.78$  kali dari kasus tanah longsor semula, jika variabel bernilai konstan.

Hasil kesimpulan menunjukkan bahwa keseluruhan pendugaan parameter  $\gamma$  yaitu model *Zero-Inflation* state pada tabel diatas tidak ada variabel prediktor yang berpengaruh signifikan terhadap kasus kejadian tanah longsor di Kota Semarang pada tahun 2018. Sehingga model *Zero-Inflation* pada kasus kejadian tanah longsor menunjukkan hasil yang tidak baik dan tidak diinterpretasikan pada metode ZINB. Hal ini mengidentifikasi bahwa dalam model *Zero-Inflation* pada ZINB belum mampu menangani nilai nol yang ada pada kasus kejadian tanah longsor di Kota Semarang tahun 2018.

Perbandingan Model Regresi ZIP dan ZINB

Menurut Akaike (1978), nilai AIC (*Akaike Information Criterion*) dapat digunakan untuk pemilihan model terbaik. Untuk melihat kebaikan model, pada penelitian ini akan dilakukan perbandingan antara nilai AIC model regresi *Zero Inflated Poisson* (ZIP) dan nilai AIC model regresi *Zero Inflated Negative Binomial* (ZINB).

Dispersi Statistik	1.99	2.39
AIC	63.7148	65.7149

Hasil estimasi parameter model Regresi ZIP dan ZINB diperoleh model yang terbaik dengan melihat AIC. Semakin kecil nilai yang diperoleh maka akan semakin baik modelnya. Diperoleh nilai AIC terkecil adalah model ZIP sebesar 63.7148 artinya dari kedua model Regresi ZIP dan ZINB, model yang terbaik untuk kasus kejadian tanah longsor di Kota Semarang adalah model *Zero-Inflated Poisson*. Selain itu bisa dilihat dari nilai dispersi statistiknya dimana nilai yang mendekati 1 sebagaimana dijelaskan pada penentuan overdispersi diatas adalah model ZIP .

## KESIMPULAN DAN SARAN

### Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan maka dapat ditarik kesimpulan sebagai berikut :

1. Model kasus tanah longsor di Kota Semarang dengan pemilihan model terbaik *Zero-Inflated Poisson* sebagai berikut:

Model diskrit

$$\ln(\mu) = 1.618 + 0.302x_1 + 0.542x_2 + 0.037x_3 + 0.118x_4$$

Model *Zero Inflation*

$$\omega = \frac{-4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4}{1 - 4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4}$$

2. Model kasus tanah longsor di Kota Semarang dengan pemilihan model terbaik *Zero-Inflated Negative Binomial* sebagai berikut:

Model Diskrit

$$\log(\mu) = \exp(1.618 + 0.302x_1 + 0.542x_2 + 0.037x_3 + 0.118x_4)$$

Model *Zero-Inflation*

$$p_i = \frac{\exp(-4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4)}{1 + \exp(-4.971 - 5.346X_1 + 1.175X_2 + 2.548X_3 + 2.187X_4)}$$

3. Model ZIP dan ZINB dengan nilai AIC terkecil yaitu model *Zero-Inflated Poisson* (ZIP) pada kasus tanah longsor di Kota Semarang

Saran

Saran yang dapat diberikan untuk penelitian selanjutnya adalah sebagai berikut :

1. Penelitian selanjutnya mengkaji lebih dalam tentang metode Regresi Zero-Inflated Poisson dan Zero Inflated Negative Binomial, yaitu dengan menambahkan unsur spasial agar lebih mengetahui persebarannya.
2. Adanya pengembangan lebih lanjut dengan menggunakan variable yang lebih banyak dan tepat terhadap penyebab terjadinya tanah longsor ,
3. Penelitian selanjutnya untuk menyelesaikan kasus overdispersi dikarenakan excess zero pada regresi poisson dapat menggunakan model Zero-Inflated Generalized Poisson (ZIGP) atau Zero-Altered Poisson (ZAP)

#### Daftar Pustaka

Akaike, H. 1978. A Bayesian Analysis of The Minimum AIC Prosedure. *Annals of The Institute Statistical Mathematics, Part A* Page 9-14.

Ariawan, Bayu, Suparti Suparti, and Sudarno Sudarno. *Pemodelan Regresi Zero-inflated Negative Binomial (ZINB) untuk Data Respon Diskrit dengan Excess Zeros*. *Jurnal Gaussian* 1.1 2012: 55-64.

Badan Pusat Statistik. 2018. *Kota Semarang Dalam Angka*. Semarang: BPS

BAPEDDA Semarang. 2010. *Rencana Tata Ruang Wilayah Kota Semarang 2010-2030*. Semarang

Bencana, Badan Nasional Penanggulangan. "Definisi dan Jenis bencana." Diunduh dari <http://www.bnpd.go.id/>, diakses 25 2012.

BPBD. 2018. *Data Rekapitulasi Bencana Kota Semarang*. Semarang. BPBD

Camelia, P. S., Nur, I. M., & Darsyah, M. Y. 2016. *PEMODELAN KEMATIAN BALITA MALNUTRISI DENGAN PENDEKATAN ZERO-INFLATED POISSON (ZIP) REGRESSION DI PROVINSI JAWA*

*TENGAH*. *Jurnal Statistika Universitas Muhammadiyah Semarang*, 4(2).

Dewanti, N. P. P., Susilawati, M., & Srinadi, I. G. A. M. Perbandingan Regresi *Zero Inflated Poisson* (Zip) dan Regresi *Zero Inflated Negative Binomial* (Zinb) Pada Data *Overdispersion* (Studi Kasus: Angka Kematian Ibu di Provinsi Bali) 2016. *E-Jurnal Matematika*, 5(4), 133-138.

Faizana, Fina, Arief Laila Nugraha, and Bambang Darmo Yuwono. "Pemetaan Risiko Bencana Tanah Longsor Kota Semarang." *Jurnal Geodesi Undip* 4.1 2015: 223-234.

Fotheringham, A.S., Brundson, C. and Charlton, M. 2002. *Geographically Weighted Regression*. John Wiley and Sons, Chichester, UK.

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

Hocking, R. R., 1996, *Methods and Application Log Linear Models*, John Wiley and Sons, Inc., New York.

Ismail, N. & A.A. Jemain. 2005. *Generalized Poisson Regression: An Alternative For Risk Classification*. *Jurnal Teknologi*. Vol. 43, No.1, 39-50.

Istiqomah, Rofikoh. 2018 *Pemodelan Zero-Inflated Negative Binomial (ZINB) Pada Kasus Difteri Di Provinsi Jawa Tengah*.

Jansakul, N. & J.P. Hinde. 2002. *Score Tests for Zero-Inflated Poisson Models*. *Computational Statistics & Data Analysis*, 40 : 75-96.

Khoshgoftaar T.M., K. Gao, & R.M. Szabo. 2004. Comparing Software Fault Predictions of Pure and Zero-Inflated Poisson Regression Models. *International Journal of System Science* 36(11): 705-715.

Lambert, D. 1992. *Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing*. *Technometrics*, 34(1) : 1-14.

Lestari A, 2008. *Pemodelan Regresi Zero Inflated Poisson (Aplikasi Pada Data Pekerja Seks Komersial di Klinik Reproduksi*

Putat Jaya Surabaya), [Tesis ] Surabaya : Program Studi Magister, Jurusan Statistika, Fakultas MIPA, Institut Teknologi Surabaya

*Regression against Zero-Inflated Generalized Poisson Alternatives. Journal of Applied Statistic*, 40(9) : 2056-2068.

Ni'mah, F. 2017. *Tingkat Kerentanan Longsor Di Kecamatan Munjungan Kabupaten Trenggalek Tahun 2016. Geo Educasia-S1*, 2(2), 286-294.

Nusantara, R. K. P., & Purhadi, P. 2015. *Pemodelan Jumlah Kasus Penyakit Tetanus Neonatorum di Jawa Timur Tahun 2012 dengan Geographically Weighted Zero-Inflated Poisson Regression (GWZIPR). Jurnal Sains dan Seni ITS*, 4(1), D79-D84.

Purhadi, P., & Rachmah, N. F. 2014. *Pemodelan Jumlah Kematian Ibu dan Jumlah Kematian Bayi di Provinsi Jawa Timur Menggunakan Bivariate Poisson Regression. Jurnal Sains dan Seni ITS*, 3(2).

Rifana, R. 2018. *Perbandingan Model Regresi Poisson Dengan Model Zero-Inflated Poisson (ZIP) Untuk Menganalisis Penyakit Kusta Di Indonesia.*

Rizaldi, A., Dixit, V., Pande, A., & Junirman, R. A. 2017. *Predicting casualty-accident count by highway design standards compliance. International journal of transportation science and technology*, 6(3), 174-183.

Rudiyanto, R. 2010. *Analisis Potensi Bahaya Tanah Longsor Menggunakan Sistem Informasi Geografis (SIG) di Kecamatan Selo, Kabupaten Boyolali (Doctoral dissertation, Universitas Muhammadiyah Surakarta).*

Ruliana, R., Hendikawati, P., & Agoestanto, A. 2016. *Pemodelan Generalized Poisson Regression (GPR) Untuk Mengatasi Pelanggaran Equidispersi Pada Regresi Poisson Kasus Campak di Kota Semarang Tahun 2013. Unnes Journal of Mathematics*, 5(1), 39-46.

Rusdiana, R. Y. 2017. *Pemodelan Regresi Hurdle Negative Binomial dengan Variabel Dependen Tersensor Kanan Pada Kasus Tetanus Neonatorum di Indonesia (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).*

Zamani, H. & N. Ismail. 2013. *Score Test for Testing Zero-Inflated Poisson*