Global J. Environ. Sci. Manage. 7(1): 117-130, Winter 2021

Global Journal of Environmental Science and Management
(GJESM)

Homepage: https://www.gjesm.net/

CASE STUDY

Using multivariate generalized linear latent variable models to measure the difference in event count for stranded marine animals

R.E. Caraka[1,2], R.C. Chen[1,*], Y. Lee[2], T. Toharudin[3], C. Rahmadi[4], M. Tahmid[5], A.S. Achmadi[4]

[1]*Department of Information Management, College of Informatics, Chaoyang University of Technology, Taiwan*

[2]*Department of Statistics, College of Natural Sciences Seoul National University, Shin Lim-Dong, Kwan Ak Ku, South Korea*

[3]*Department of Statistics, Padjadjaran University, West Java, Indonesia*

[4]*Museum Zoologicum Bogoriense, Research Center for Biology, Indonesian Institute of Sciences, Indonesia*

[5]*Indonesian Agency for Meteorological, Climatological and Geophysics, BMKG, Manado, Indonesia*

ARTICLE INFO

Article History:
Received 23 April 2020
Revised 07 July 2020
Accepted 17 August 2020

Keywords:
Indonesia
Latent
Madden–Julian oscillation (MJO)
Marine species
Multivariate

ABSTRACT

BACKGROUND AND OBJECTIVES: The classification of marine animals as protected species makes data and information on them to be very important. Therefore, this led to the need to retrieve and understand the data on the event counts for stranded marine animals based on location emergence, number of individuals, behavior, and threats to their presence. Whales are generally often stranded in very shallow areas with sloping sea floors and sand. Data were collected in this study on the incidence of stranded marine animals in 20 provinces of Indonesia from 2015 to 2019 with the focus on animals such as *Balaenopteridae, Delphinidae, Lamnidae, Physeteridae*, and *Rhincodontidae*.

METHODS: Multivariate latent generalized linear model was used to compare several distributions to analyze the diversity of event counts. Two optimization models including Laplace and Variational approximations were also applied.

FINDINGS: The best theta parameter in the latent multivariate latent generalized linear latent variable model was found in the Akaike Information Criterion, Akaike Information Criterion Corrected and Bayesian Information Criterion values, and the information obtained was used to create a spatial cluster. Moreover, there was a comprehensive discussion on ocean-atmosphere interaction and the reasons the animals were stranded.

CONCLUSION: The changes in marine ecosystems due to climate change, pollution, overexploitation, changes in sea use, and the existence of invasive alien species deserve serious attention.

DOI: 10.22034/gjesm.2021.01.09

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
41	7	2

*Corresponding Author:
Email: *crching@cyut.edu.tw*
Phone: +886-4-2332-3000
Fax: +886-4-2374-2337

Note: Discussion period for this manuscript open until April 1, 2021 on GJESM website at the “Show Article.

INTRODUCTION

The Indo-West Pacific arguably offers enormous diversity in marine mammal species throughout the world, as observed in representatives of 11 of the 13 families of the Cetacean with more than 40 of the 85 species living in the sea (Rudolph *et al.,* 2009). Furthermore, there are approximately 1,250 species of sharks, rays, and ghost shark in the world. An estimate of 218 species was found in the Indonesian sea, out of which only 26 species have high economic value in the global market (Maryanto, *et al.,* 2008). The different types of sharks existing in the waters include the *Carcharhinidae, Lamnidae, Alopiidae, Sphyrnidae* families as well as the *Mobulidae* rays including the Manta and Mobula which are considered to be the most commonly used groups. The *Rhincodon typus* whale shark is the largest fish species in the world usually found in tropical to subtropical waters (Norman, 2002) as well as oceans and coastal seas including lagoons, coral atolls, and reefs. The total length of a whale shark is more than 18 meters and has been reported to be reproducing through ovoviviparity with the embryo ready to come out of the mother's stomach in sizes ranging between 55 and 64. Moreover, those categorized as adults are usually 7.05-10.26 m or more for males and 12 m or more for females (Stevens, 2007) with both gender found to have a swimming character adapted only to the aquatic environment (Colman, 1997). Whale sharks can live in deep and shallow water near the coast, and their habitat is mostly related to water quality, plankton concentration, temperature, current patterns, weather, and water location. Water quality is a critical factor for whale sharks due to its relation to the availability of nutrients for zooplankton which is their primary food source. These marine animals have a filter system feeder to feed on planktonic and nektonic biota as well as extensive migration abilities and movements thought to be related to the high productivity of zooplankton, changes in water temperature, currents, wind, and other water parameters. Zooplankton obtains food from producers that photosynthesize by phytoplankton and based on different environments considered to be significant for natural life. Meanwhile, several living space models exist with explicit attributes with some using nearness-just information while others use nearness nonattendance or tally information. Apart from the nearness-just models, which do not deal with nonappearance information such as zeros, selecting nearness nonappearance or tally-based models is a difficulty due to the reliance on contemplated species primarily when concentrating on uncommon species based on innate trouble attached to determining the models best suited to countless nonattendances. As previously referenced, uncommon species generally lead to a low number of sightings per unit exertion and this shortage of information makes it hard to determine the best appropriation models. Some studies have, however, addressed the use of models for rare species datasets (Lomba *et al.,* 2010; Mouillot *et al.,* 2013; Demos *et al.,* 2016; Kurniawan, *et al.,* 2018) but reliability and uncertainty associated with the predictions produced by these models remain pending issues (Kurniawan, *et al*., 2018). An option to address these challenges is by testing the maintenance of performance for a species distribution model when there is a decrease in the input data in order to assess its reliability in handling small datasets for rare species. Several events of animal marine species have been stranded in Indonesia in the past few decades, and Chan, *et al*. (2017) have created a database for this occurrence in addition to their mortality. Some factors causing the strandedness have been identified with the most important ones considered to be internal such as illness (Duignan, 2003), malnutrition, natural toxins, and infectious diseases (Wibowo *et al*., 2014). Meanwhile, the interactions between marine mammals and plastic debris have been the focus of studies for many years (Lusher *et al.,* 2018), the ingestion of marine litter such as plastic considered to be leading to whale shark fatality (Abreo *et al.,* 2019). The loud sound emitted during offshore industrial activities has also been reported to have an impact on marine mammals (Verfuss *et al.,* 2018). The data and information on stranded marine animals are very useful and important to understand the patterns and possible cause for the phenomenon in relation to the environment and changes in their habit due to human activities. In the policy context, the study is expected to help determine the areas with the greatest risk of stranded marine animals and the best ways to mitigate its occurrence. This study was, therefore, conducted to measure the incidence of stranded marine animals and differences in the event counts using multivariate generalized linear latent variable models on data collected in 20 Provinces of Indonesia from 2015 to 2019.

MATERIALS AND METHODS

Data collection

The data and information including species and the number of individuals on stranded marine animals were manually collected from media reports from 2015 to 2019 and verified using several sources. This data collected method is limited by the potential omission of information on some remote areas such as small islands beyond media coverage.

Multivariate latent generalized linear models

The linear regression model seeks to establish a linear relationship (Ha *et al.,* 2002) between the response variable and one or more covariates (Crawley, 2012) while the Generalized Linear Model (GLM) is a natural generalization of the Linear Regression model (Jamilatuzzahro *et al.,* 2018). This further allows linking response variables to one or more covariates via the link function (Jamilatuzzahro *et al*., 2019; Rahman *et al*., 2019) in order to explore the distribution and density of the variables. This study used $y_1,\ldots,y_n$ to represent a sequence of independent random variables (Noh *et al*., 2019; Lee *et al.,* 2012; Lee *et al.,* 2001) which were identically distributed by law to belong to the exponential family. The density was compared to the Lebesgue or a counting measure using Eq. 1.

$$f(y_i;\theta_i,\phi)=\exp\frac{y_i\theta_i-b(\theta_i)}{a(\phi)}+c(y_i;\phi) \quad (1)$$

Where, θ_i is the parameter position and ϕ is the dispersion parameter. Moreover, the expectancy and variance of *Y* are provided by $\text{Å}(Y)=b'(\theta)$ and $\text{Var}(Y)=b''(\theta)a(\theta)$, respectively. The latent variable model produces a significant instrument to investigate multivariate information, especially in ecology modeling by offering an applied structure to bring numerous divergent strategies together and serve as a base to create new techniques. Latent models are able to determine the joint dispersion of several arbitrary factors and later change to a latent variable model with a portion of these factors. Herliansyah *et al*. (2018) used these models to measure the diversity of bird species while Caraka *et al.,* (2018) used the negative binomial to determine the diversity in arthropods species counts. Moreover, Rahman *et al*., (2019) analyzed the diversity of Banteng and *Bos javanicus* while Caraka *et al.,* (2020a) used Butterfly diversity on species distribution. Anggraini *et al.,* (2020) also applied a latent factor linear mixed model to Flanders' data. It is, therefore, possible to use the dimension reduction in factor analysis to construct the model with latent variables of $x_1, x_2,\ldots,x_q$ with conditional distribution $g_i(x_i|y),(i=1,2,\ldots,p)$ (Niku *et al.,* 2019a; Niku *et al.,* 2017; Niku *et al.,* 2019b). This equation can be written as shown in Eqs. 2 and 3 (Bartholomew *et al.,* 2011).

$$g_i(x_i|y)=F_i(x_i)G_i(y)\exp\sum_{j=1}^{q}u_{ij}(x_i)\phi_j(y) \quad (2)$$

$$h(y|x)=\frac{h(y)\prod_{i=1}^{p}[F_i(x_i)G_i(y)\exp\sum_{j=1}^{q}X_j\phi_j(y).}{\int h(y)\prod_{i=1}^{p}[F_i(x_i)G_i(y)\exp\sum_{j=1}^{q}X_j\phi_j(y)dy.} \quad (3)$$

The optimization methods used in these models are Laplace (Huber *et al.,* 2004) and Variational approximations (Caraka *et al.,* 2020b). Laplace is a type of multidimensional integral approximation using Eq. 4.

$$\int_{\mathbb{R}^d} b(x)\exp-\lambda h(x)dx \quad (4)$$

Where, λ is a real parameter such that $\lambda>1$ while *h* is supposed to be an overall minimum in *x* which is regular in a neighborhood of *h* to ensure $h'(x)=0;\det[h''(\hat{x})]>0$ and verify the following for all δ> 0. The Laplace method, while considering the integral (1) as the integral of *b,* shows the Gaussian measure for the small variance of order $\frac{1}{\lambda}$. The $h(x)$ in the integrand using the Taylor development (Herliansyah *et al*., 2018; Caraka *et al.,* 2020b) in the order is presented in Eq. 5.

$$\begin{aligned}&\text{Inf}\{h(x)-h(\hat{x}):|x-\hat{x}|>\delta\}>0\\&h(x)\approx h(\hat{x})+\frac{1}{2}(x-\hat{x})^T h''(\hat{x})(x-\hat{x})\\&\int_{\mathbb{R}^d} b(x)\exp-\lambda h(x)dx\approx\int_{\mathbb{R}^d}\\&b(x)\exp-\frac{\lambda}{2}(x-\hat{x})^T h''(\hat{x})(x-\hat{x})dx\end{aligned} \quad (5)$$

The estimation through Laplace shows each of the parameters is simultaneous (Kristensen *et al.,* 2016; Bianconcini *et al.,* 2012) as expressed in Eq. 6.

$$L=\sum_{l=1}^{n}\log f(y_l)=\sum_{l=1}^{n}\log\int_{\mathbb{R}^d} g(y_l\mid x_l)h(z_l)dz_l \quad (6)$$

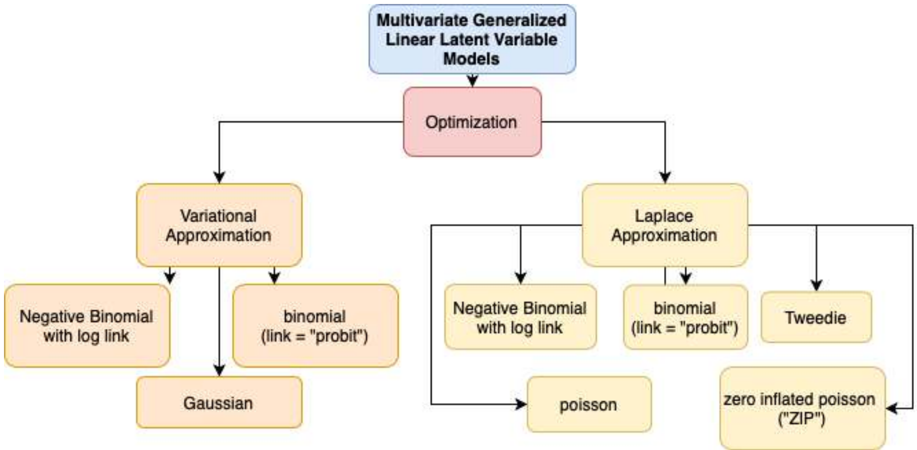

Fig. 1: Type of Optimization and Distribution Applied in Multivariate Generalized Linear Latent Variable Model

Variational approximation, however, shows that a vector $d \ll m$ underlying latent variables, u_{ij} and the parameter vector Ψ assumes responses y_{ij} are obtained from the exponential family of distributions using Eq. 7 (Hui *et al.*, 2017).

$$\ell(\Psi) = \sum_{i=1}^{n} \log\{fy_i, \Psi\} = \sum_{i=1}^{n} \log\left(\int \prod_{j=1}^{m} f\left(y_{ij}|u_i, \Psi\right) f\left(u_i\right) du_i \right) \quad (7)$$

The distribution selection method used in this model was Kolmogorov-Smirnov (Fasano *et al.*, 1987). Meanwhile, the two optimization methods have different advantages and disadvantages based on time performance and distribution. More specifically, it is impossible to use Tweedie distribution in Laplace approximation but it is applicable in Variational approximation even though it requires extended processing time. Moreover, the distributional choice of latent variables, u_i, used in several research papers is a normal distribution with mean zero and constant variance. The types of optimization and distribution used in the multivariate generalized linear latent variable model are, however, shown in Fig. 1.

The distribution functions used for response options include Poisson (link = "log"), "negative. binomial" (with log link), binomial [(link = "probit") and (link="logit") with "LA"], zero-inflated Poisson ("ZIP"), gaussian (link = "identity"), Tweedie ("Tweedie") (with log link only with "LA"), and "ordinal" (only with "VA"). The exponential dispersion in the model was addressed mostly through the distribution of both family of two linear exponential parameters using either dispersion parameter as indicated in Eq. 8.

$$p(y|\theta,\phi) = \alpha(y,\phi)\exp\left(\frac{y\theta - z(\theta)}{\phi}\right) \quad (8)$$

The important information contained in the Tweedie model include normal (*p* = 0), Poisson (*p* = 1), gamma (*p* = 2), and gaussian inverse (*p* = 3). The Tweedie Distribution is principally an exponential family used in dispersing parameters with a variety of functions $V(y) = \mu^p$. Meanwhile, the zero-inflated Poisson is a mixed model between the distribution of Poisson and of events which is excess zero. Eq. 9 explains the random variable Y following ZIP with zero values assumed to be occurring through two scenarios. The first is the probability of π_i which produces zero observations while the second is the probability of $(1-\pi_i)$ which generates data following Poisson(λ).

$$P(y_i|\pi_i,\lambda_i) = \begin{cases} \pi_i + (1-\pi_i)\exp(-\lambda_i), y_i = 0 \\ (1-\pi_i)\dfrac{\exp(-\lambda_i)\lambda_i^{y_i}}{y_i!}, y_i = 1,2,\ldots; 0 \le \pi_i \le 1 \end{cases} \quad (9)$$

Where, *Y* is a Poisson random variable depending on parameters λ which is the value of a random variable Λ and follows Gamma to form a Poisson-Gamma mixture distribution called Negative Binomial as indicated in Eq. 10.

$$P(y|\lambda,\alpha) = \frac{\Gamma(y+\alpha^{-1})}{y!\Gamma(\alpha^{-1})}\left(\frac{\alpha}{1+\alpha\lambda}\right)^{y}\left(\frac{1}{1+\alpha\lambda}\right)^{\alpha^{-1}} \quad (10)$$

Performance Evaluation

Akaike Information Criterion (AIC) in Eq. 11 and Bayesian Information Criterion (BIC) in Eq. 12 have been used in numerous down applications to determine a model or variable (Warton, 2005; Kuha, 2004; Caraka *et al*., 2020c). The model determination criteria are factual devices with the ability to recognize an ideal measurable model from several others with the set typically called a lot of up-and-comer models. Moreover, under normality, thickness capacity strives towards keeping up the properties of $\hat{\theta}$ by assuming the density function of the exact model $g(y,\theta_0)$ belongs to $\mathcal{F}$, and where *k* is the element of the parameter vector ϑ. There are, however, a few expansions of AIC, such as the situation for a small dataset where the Akaike Information Criterion Corrected (AICc) ought to be progressively pertinent with the punishment term rectified to $2k\left[\frac{n}{(n-k-1)}\right]$.

$$AIC = -2\ln f(y,\hat{\theta}) + 2k \tag{11}$$

The Bayesian Information Criterion (BIC) can also be defined using Eq. 12.

$$BIC = -2\ln f(y,\hat{\theta}) + k\ln(n) \tag{12}$$

BIC is an assessment foundation for models which use the most extreme probability strategy and based on the condition that the example size *n* is adequately enormous.

RESULTS AND DISCUSSION

Species counts

Environmental condition is a focal component for the biological specialty of animal types and, by augmentation, its nature. This means species living space connections are one of the premises to explain the high strandedness of marine animals. Information on strandedness in natural surroundings tends to be focused on some species with several theories devoted to their territorial inclinations. It has also been discovered that different categories of animals utilize different natural surroundings for several purposes mostly due to their movement which is associated with relocating, resting, or reproducing. This, therefore, shows their living space is usually defined by both physical and natural attributes which have also been observed to be changing with time. This study determined the varying habitat preferences based on some considered time scales. For the purpose of this study, habitat preference is defined as a positive association with specific environmental conditions to produce random distributions of animals. There was, however, no attempt to relate the observed habitats with associated activities due to the inaccessibility of the information. The preference for habitats depends on species traits, therefore, this research was expected to determine different levels of variability in the ocean preferred by different species based on their life histories. The information presented in Fig. 2 showed one event was recorded in 2015, 4 in 2016, 5 in 2017, 18 in 2018, and 33 events in 2019 while the highest occurrence, 22, was with *Delphinidae*. These data are also presented in Table 1, using the number of inventory or count of individuals to represent the actual stranded population due to the fact that the data were collated manually from media reports and later pre-processed, labeled, and classified based on the type according to the expert system.

Modeling of stranding events

The difference in the diversity of event counts was practically measured using multivariate latent generalized linear models with two optimization methods which are Laplace and Variational approximations. It is also possible to use the model through different distributions such as Tweedie, Negative Binomial (NB), Poisson, Zero-inflated Poisson (ZIP), and Gaussian. The modeling, however, showed the use of Poisson with the smallest AIC value is the appropriate distribution method for Laplace approximation while Gaussian is the best for Variational approximation.

Table 2 is used to represent the performance evaluation with the likelihood being a statistical tool to summarize data evidence of unknown parameters while the log-likelihood value stands for the statistical measures for the models. The summarily means a model with relatively high value is better. It is also possible to consider the Log-Likelihood to be lying between both -Inf and +Inf while the unmitigated appearance at either value does not have the ability to provide any information. Nevertheless, the smallest AIC Laplace approximation for this model was reached at 86.1433 after which the latent value obtained from the best model was used to create the spatial difference in event counts and spatial clusters. Fig. 3,

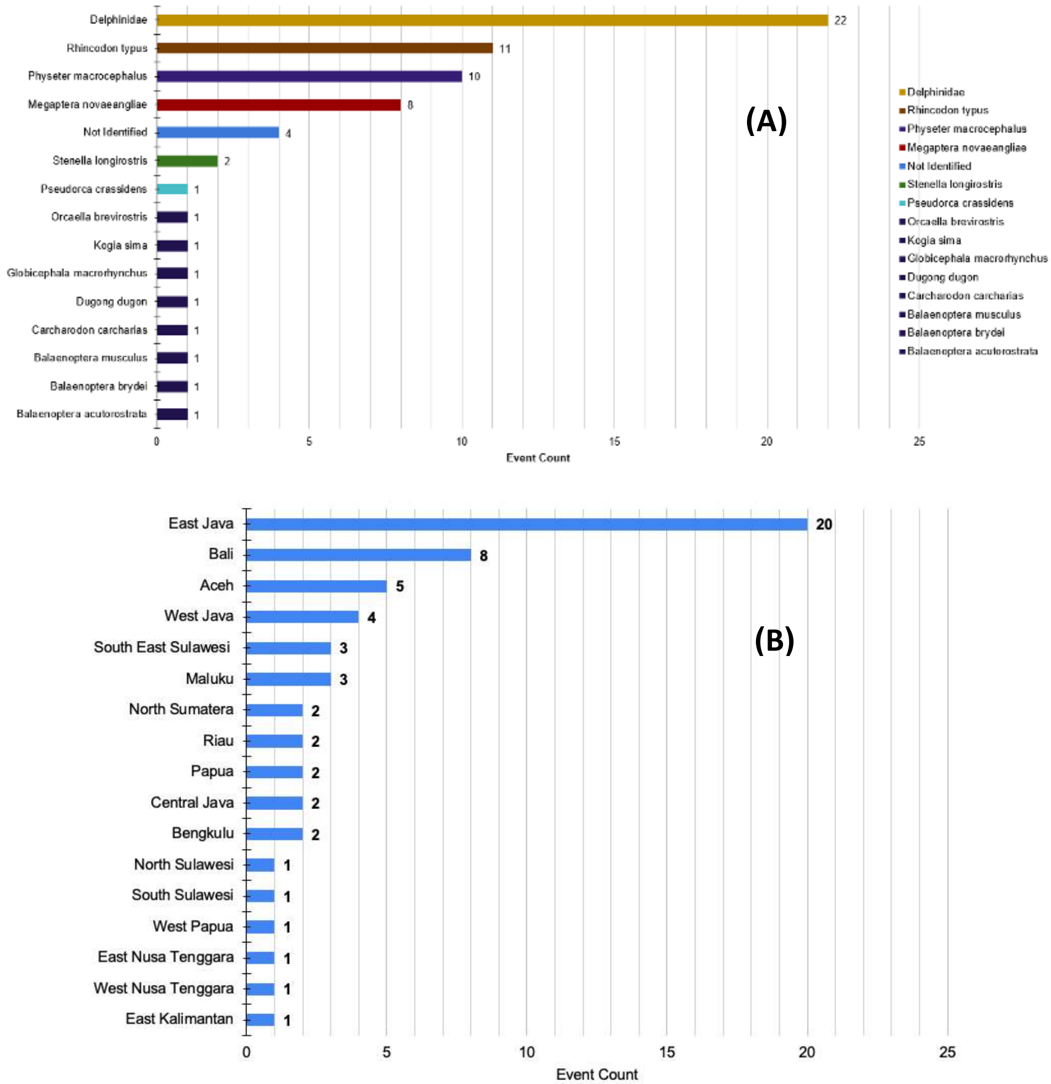

Fig. 2: count of stranded marine animals based on species (A) and Provinces (B)

however, presents the difference in event counts of stranded marine animals in Indonesia from 2015 to 2019 with the highest rate generally recorded around Java Province.

The intercept value models were, therefore, used as the information to create spatial clustering and, based on the three clusters indicated in Fig. 4, the differences between the stranded marine animals based on provinces were more apparent. Group 1 generally covers areas in Central Java, East Java, East Kalimantan, Maluku, Banten, West Kalimantan, North Sumatera, West Sulawesi, East Nusa Tenggara, West Papua, Riau, West Nusa Tenggara, and North Sulawesi. Group 2 covers Bengkulu, Papua, and South East Sulawesi while Group 3 covers the areas of Bali, West Java, and Aceh.

Reasoning

An ocean-atmosphere interaction analysis was conducted on the stranded marine phenomenon in East Java on 15 June 2016, Bali on 10 May 2019, and Bengkulu on 21 March 2018 using a dataset from BMKG Indonesia. Meanwhile, the stranded animal events are assumed to be influenced by climatological conditions and this means further study needs to

Table 1: Event counts stranded marine animals in Indonesia

Event	Year	Province	Beach	Latitude Longitude	*Type*	Size	N	Status
11/12/19	2018	Central Java	nn	-7.764416, 109.519016	*Delphinidae*	nn	1	Died
22/09/19	2019	Bali	nn	-8.634111, 115.293066	*Delphinidae*	nn	1	Died
20/09/19	2019	West Java	nn	-6.797359, 108.787466	*Rhincodon typus*	nn	1	Died
16/09/19	2019	East Java	Pantai Bambang	-8.290448, 113.105731	*Rhincodon typus*	5	1	Died
16/09/19	2019	East Kalimantan	Pantai Corong	-1.403559, 116.653588	*Pseudorca crassidens*	nn	1	Died
15/09/19	2019	East Java	Kepanjen	-8.388036, 113.382127	*Rhincodon typus*	8	1	Died
11/09/19	2019	East Java	PLTU Paiton	-7.711318, 113.583328	*Rhincodon typus*	5	1	Survived
11/09/19	2019	West Java	nn	-6.056518, 107.412216	*Delphinidae*	nn	1	Died
10/09/19	2019	East Java	Pantai Kajaran	-8.289643, 113.103304	*Rhincodon typus*	6	1	Died
02/09/19	2019	Bali	Pantai Serangan	-8.724652, 115.243014	*Kogia sima*	nn	1	Died
01/09/19	2019	East Java	Pantai Klatak	-8.269857, 111.769919	Delphinidae	nn	1	Survived
30/08/19	2019	Maluku	Liliboi	-3.741685, 128.033191	*Megaptera novaeangliae*	8	1	Died
28/08/19	2019	East Java	PLTU Paiton	-7.711318, 113.583328	*Rhincodon typus*	5	1	Survived
18/08/19	2019	West Java	nn	-6.081699, 107.425979	*Delphinidae*	nn	1	Died
16/08/19	2019	Banten	Binuangen	-6.839042, 105.900045	*Delphinidae*	nn	1	Died
04/08/19	2019	Bali	nn	-8.433401, 114.826593	*Stenella longirostris*	nn	1	Died
29/07/19	2019	Maluku	nn	-3.603043, 128.709608	*Megaptera novaeangliae*	nn	1	Died
23/07/19	2019	West Kalimantan	nn	0.421578, 108.942288	*Delphinidae*	nn	1	Died
22/07/19	2019	North Sumatera	nn	3.915655, 98.641087	*Delphinidae*	nn	1	Died
11/07/19	2019	East Java	Pantai Bambang	-8.290448, 113.105731	*Megaptera novaeangliae*	11	1	Died
11/07/19	2019	West Sulawesi	nn	-3.457637, 119.422586	*Delphinidae*	nn	1	Survived
09/07/19	2019	Bengkulu	Teluk Sepang	-3.904638, 102.280918	*Delphinidae*	nn	1	Died
26/06/19	2019	West Java	Pantai Cikadai	-7.190663, 106.443775	*Not Identified*	nn	1	Died
13/06/19	2019	East Nusa Tenggara	Pantai Doreng	-8.740595, 122.408641	*Delphinidae*	nn	1	Died
10/06/19	2019	Bali	Pantai Melaya	-8.289554, 114.498472	*Carcharodon carcharias*	1	1	Died
28/05/19	2019	West Papua	nn	-0.340917, 130.945255	*Delphinidae*	nn	1	Died
09/05/19	2019	Bali	Pantai Mengiat	-8.808089, 115.231902	Delphinidae	nn	1	Died
09/04/19	2019	Papua	nn	-4.777610, 136.542187	*Physeter macrocephalus*	nn	1	Died
09/03/19	2019	Aceh	Panga	4.549890, 95.694539	*Delphinidae*	nn	1	Died
09/03/19	2019	Aceh	nn	4.550072, 95.691889	*Stenella longirostris*	nn	1	Died
01/03/19	2019	South East Sulawesi	Pulau Bokori	-3.939953, 122.659604	*Not Identified*	nn	1	Died
05/02/19	2019	Bali	Kuta	-8.720257, 115.168991	Delphinidae	nn	1	Survived
04/02/19	2019	Papua	nn	-4.777610, 136.542187	*Balaenoptera brydei*	nn	1	Died

Continued Table 1: Event counts stranded marine animals in Indonesia

Event	Year	Province	Beach	Latitude Longitude	*Type*	Size	N	Status
27/01/19	2019	North Sumatera	Sungai Kualuh	2.611761, 100.095708	*Delphinidae*	nn	2	Survived
24/01/19	2019	East Java	Bancamara	-7.003523, 114.171978	*Rhincodon typus*	3	1	Died
14/01/19	2019	Maluku	Pulau Buru	-3.110435, 126.862872	*Balaenoptera musculus*	18	1	Died
09/01/19	2019	Aceh	nn	4.550906, 95.690956	*Delphinidae*	nn	1	Died
08/01/19	2019	East Java	nn	-8.329642, 112.221151	*Delphinidae*	nn	1	Died
19/11/18	2018	South East Sulawesi	nn	-5.326570, 123.465253	*Physeter macrocephalus*	9.5	1	Died
06/10/18	2018	East Java	Pantai Kajaran	-8.290518, 113.125963	*Megaptera novaeangliae*	15	1	Died
26/09/18	2018	Riau	nn	1.445870, 102.154649	*Orcaella brevirostris*	nn	1	Died
04/09/18	2018	Riau	nn	2.048109, 101.564324	*Dugong dugon*	nn	1	Died
15/07/18	2018	Aceh	Pantai Seragihan	2.210150, 98.067943	*Delphinidae*	nn	1	Died
17/06/18	2018	Aceh	nn	5.059513, 97.665810	*Megaptera novaeangliae*	nn	1	Died
06/06/18	2018	East Java	Randutatah	-7.700458, 113.482848	*Megaptera novaeangliae*	nn	1	Died
04/04/18	2018	East Java	Pantai Praureme k	-8.323072, 111.626680	*Rhincodon typus*	4	1	Died
29/03/18	2018	West Nusa Tenggara	Tabuan	-8.899194, 116.448733	*Physeter macrocephalus*	10	1	Died
21/03/18	2018	Bengkulu	nn	-4.702177, 103.266769	*Physeter macrocephalus*	12.5	1	Died
20/03/18	2018	Bali	nn	-8.074885, 115.138295	*Physeter macrocephalus*	15	1	Died
02/03/18	2018	East Java	nn	-7.712337, 114.182269	*Physeter macrocephalus*	17	1	Survived
01/03/18	2018	East Java	nn	-7.714664, 114.181058	*Physeter macrocephalus*	20	1	Died
01/02/18	2018	South East Sulawesi	nn	-4.801613, 121.637438	*Not Identified*	nn	1	Died
13/11/17	2017	Aceh	Ujung Kareung	5.652282, 95.423781	*Physeter macrocephalus*	nn	4	Died
13/11/17	2017	Aceh	Ujung Kareung	5.652282, 95.423781	*Physeter macrocephalus*	nn	6	Survived
30/10/17	2017	East Java	nn	-7.722134, 113.110528	*Megaptera novaeangliae*	nn	1	Died
27/09/17	2017	North Sulawesi	nn	1.117452, 124.339524	*Physeter macrocephalus*	11	1	Died
30/08/17	2017	Bali	Pantai Yeh Kuning	-8.400932, 114.659160	*Megaptera novaeangliae*	8	1	Died
15/08/17	2017	East Java	Pantai NgGenjor	-8.294141, 111.768709	*Not Identifiedi*	10	1	Died
26/12/16	2016	East Java	Pantai Sidem	-8.256870, 111.798004	*Rhincodon typus*	nn	1	Died
12/08/16	2016	Central Java	Pantai Kemiren	-7.697235, 109.057250	*Delphinidae*	nn	1	Died
15/06/16	2016	East Java	nn	-7.732271, 113.177689	*Globicephala macrorhynchus*	nn	15	Died
15/06/16	2016	East Java	nn	-7.732271, 113.177689	*Globicephala macrorhynchus*	nn	17	Survived
15/05/16	2016	East Java	Pantai Randupitu	-7.774278, 113.320638	*Physeter macrocephalus*	4	7	Died
26/12/15	2015	East Java	Pantai SIdem	-8.256870, 111.798004	*Rhincodon typus*	7.5	1	Died

Table 2: Accuracy based on the type of approximation and distribution

Approximation	Distribution	log-likelihood	Df	AIC	AICc	BIC
LA	Tweedie	-42.66616	12	109.3323	161.3323	120.6656
LA	Negative Binomial	-44.68026	12	113.3605	165.3605	124.6938
LA	Poisson	-45.06127	8	106.1225	120.5225	113.678
LA	ZIP	-45.06127	12	114.1225	166.1225	125.4558
LA	Gaussian Link Identity	-31.07117	12	86.14233	138.1423	97.4756
VA	Gaussian Link Identity	-40.57117	12	105.1423	157.1423	116.4756
VA	Negative Binomial	-44.68026	12	136.9484	188.9484	148.2817
VA	Poisson	-55.19088	8	126.3818	140.7818	133.9373

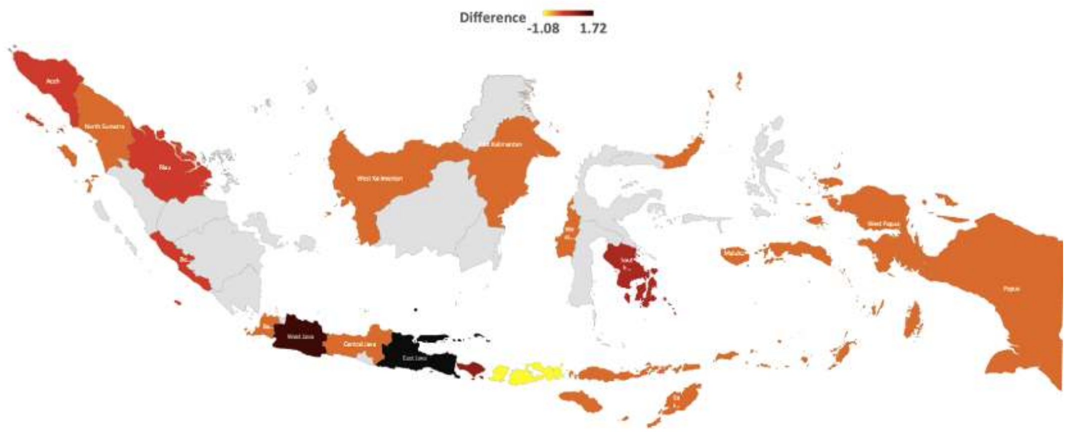

Fig. 3: Intercept models based on stranded marine animal's event counts

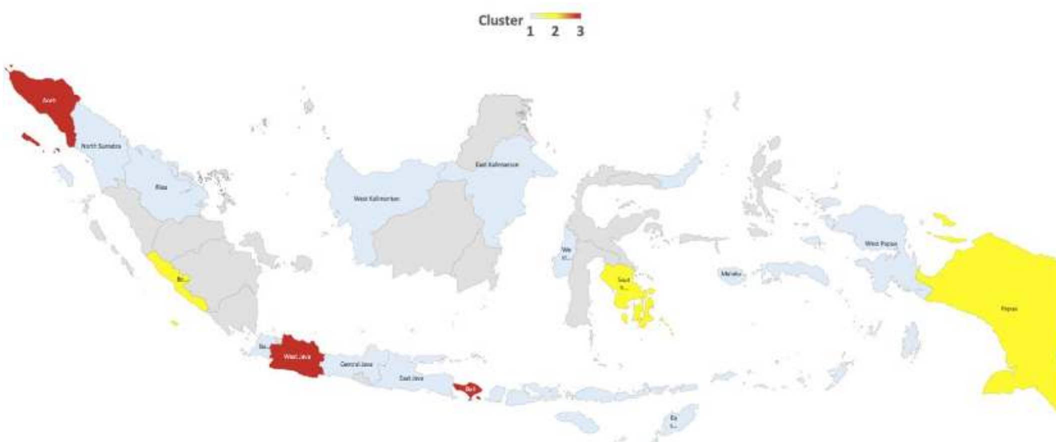

Fig. 4: Clustering based on intercept value stranded marine animal's event counts

be conducted on wind speed and direction at the location. This is important considering the fact that the changes in these factors have the ability to change the migration path of marine animals, thereby, providing opportunities for them to be stranded. Wind has been discovered to have an influence on ocean currents and waves with the wind-generated currents having varying speeds depending on their depth and, apart from its horizontal movement, wind also causes vertical water currents known as upwelling and down welling in certain areas. Upwelling is an oceanographic phenomenon which involves a solid, cold, and typically wind-driven motion which brings a nutrient-rich mass of water to the surface of the sea. Moreover, wind conditions also affect high and low waves occurring continuously at the sea level with the changes in its speed considered to be increasing sea waves and this consequently has the ability to moving water to the surface. The upwelling zone expands in the tidal sea region and this is usually followed by high wind speeds towards the land. This means an increase in the strength of ocean waves encourages mammals to search for food nutrients or fish flocks and they usually end up being drawn into coastal areas.

Bengkulu, 21 March 2018

The dominant wind recorded in Bengkulu region in March 2018 blew from South to Northwest and later in the night to Northeast with an average speed of 5 knots and a maximum speed of 24 knots and the weather was observed at this period to be dominated by rain from night to morning as well as fog and haze which were majorly experienced in the morning.

Madden–Julian oscillation (MJO) is observed to be getting stronger when it is outside the circle and in Phases 3, 4, and 5 which are its position in the Indonesian territory. The MJO in the first and second week of March 2018 is presented in Fig. 5 including phases 2 and 3 while the third week moves towards the Neutral phase and at the end of the month, it shifts from phase 6 to 7. This movement shows the MJO is quite significant and influences the growth of rain clouds in Western Indonesia especially Bengkulu in the first and second weeks after which it did not affect the increase in rainfall as indicated in its shift to neutral during the third to fourth week.

Bali, 10 May 2019

The wind in the Bali region is always not far from the western and eastern winds and this is observed from the very much influence of the winds from these areas on the region has indicated on the regional meteorological scale. It periodically blows every 6 months alternately from mainland Asia to Australia and vice versa through the concept of monsoon or western wind when it blows from the west or Asia and the east wind when blowing from the east or Australia. May is the final period of the first transition period from the rainy to the dry season which usually occurs in June, July, and August and the east wind has been observed to be very dominant at this period. Therefore, May is usually known as the end of the transition period into the dry season, especially when there are no disturbances such as El-Niño and La-Niña. MJO was, however, discovered to be non-active in this month because Bali is not in quadrants 4 and 5, therefore, a little rain is usually experienced, even for just two times, with the intensity not exceeding 1 mm per day.

East Java 15 June 2016

The MJO diagram is divided into 8 phases with phase 1 in Africa, phase 2 in the western Indian Ocean, phase 3 in the eastern Indian Ocean, phases 4 and 5 in Indonesia, phase 6 in the western Pacific region, phase 7 in the central Pacific, and phase 8 in

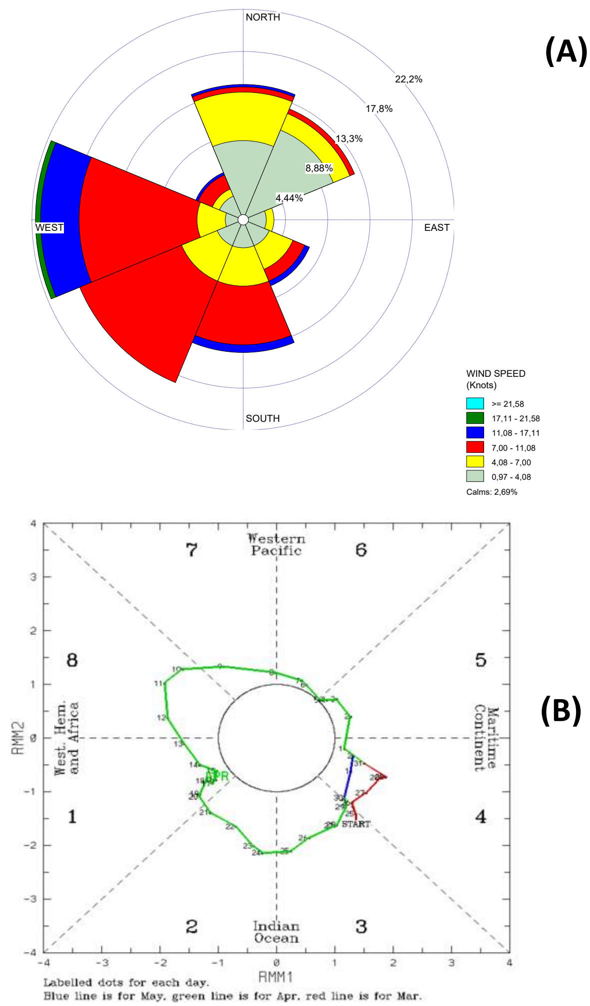

Fig. 5: Wind Speed (A) and MJO (B) in Bengkulu, Indonesia

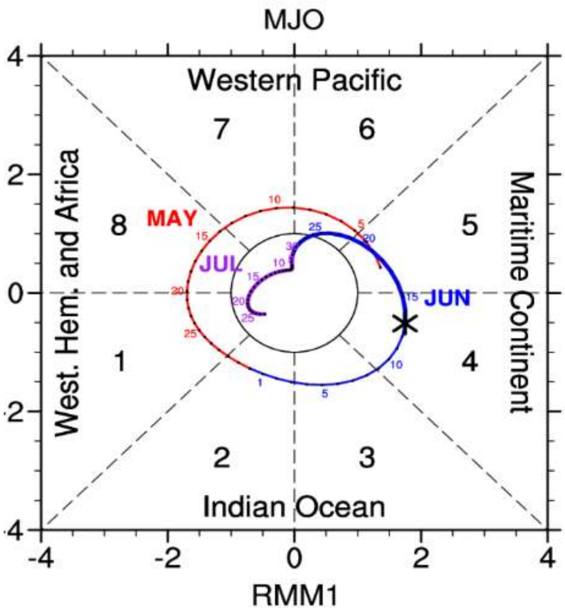

Fig. 6: Madden–Julian oscillation Bali, Indonesia

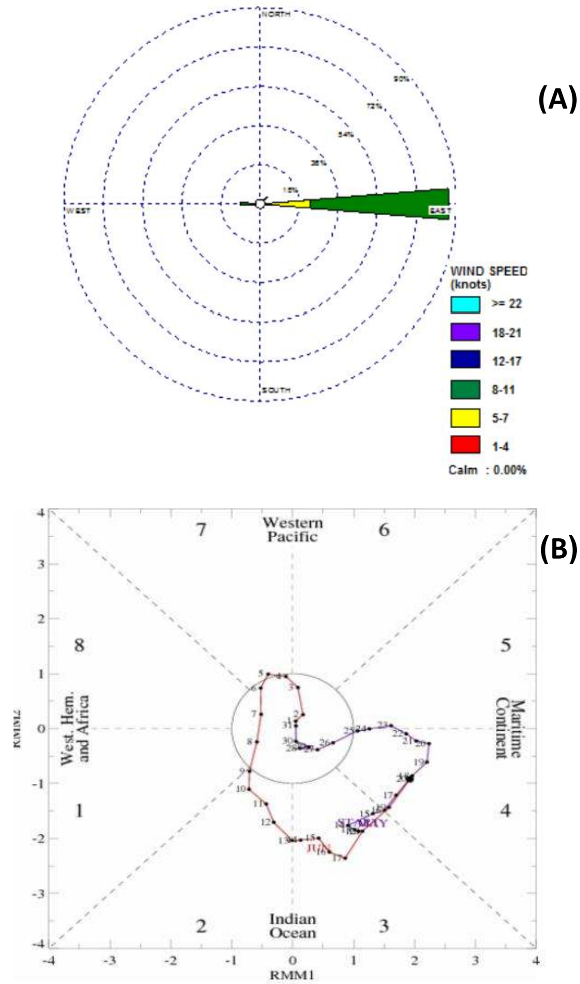

Fig. 7: Wind Speed (A) and MJO (B) in East Java, Indonesia

the convection regions of the Western Hemisphere. Fig. 6 represents the placement of the track in the small circle at the middle shows the MJO is in a weak condition while an outside placement shows it is strong or active. Fig. 7 shows MJO started to be strengthened in the West Indian Ocean of Sumatra on June 14, 2016, and observed to be entering the Sumatra and Java on June 19-20, 2016, thereby, being one of the causes of high rainfall. Another cause is the high sea surface temperature anomaly in the Madura Strait compared to other waters. In June 2016, most of the wind conditions, 87%, were from the East with an average speed of 8-11 knots while the Northeast direction has 4% with an average wind speed of 5-7 knots as shown in the Wind rose diagram.

CONCLUSION

The most frequently stranded whale mammals are those living in the deep sea and the location for the strandedness is usually very shallow areas. This is not surprising considering the fact that these animals are accustomed to swimming in the deep sea and find it difficult to return due to the less effectiveness of the echolocation capabilities they use in navigating when they are in such environments. This means it is possible the majority of whales are stranded due to navigation errors, for example, when they hunt prey to remote and dangerous areas. It was also found that there are cases the marine animals are dragged by the tide and this mostly leaves them stranded on the coast, thereby, becoming dehydrated and dead. Moreover, one or two whales were also discovered to have gotten lost and become stranded in the middle of the road while some are attached to the changes in the Earth's magnetic field caused by sunspots and high-level radio waves emitted by solar storms. There is also a large part of the radiofrequency (RF) waves range reaching the Earth and its noise has been found to be interfering with the magnetic orientation of several species. Some stranded marine mammals were also discovered after dissection to have a large lump of marine debris in their stomach containing several ropes, plastic cups, and plastic bags and this indicates the relationship between the type of food and eating behavior of whale sharks and their appearance in certain locations. This is also observed in their appearance in groups due to the abundance of planktons floating freely in waters and they are also found to be centered when they feed in the same area. Those discovered to have stranded and entangled in the South Coast and Banggai Islands were reported to be allegedly due to the existence of food sources in both waters and this was supported by the lack of differences in the appearances of both male and female whale sharks found at each of the study locations during the study period of whale shark behavior which tended to be the same between males and females. The only variation observed was the number of occurrences with the male having a higher appearance compared to the female due to the significant differences in the number of individuals. Furthermore, the wounds found on the fins, body, and mouth area are possibly due to the friction with the net of the chart or entangled fishing line of fishermen. Mass tourism activities, fishing, as well as ship collisions in Indonesian waters also have the ability to cause injuries to these marine mammals as observed in the injuries on the bodies of several whale sharks in the waters. Meanwhile, those reported to be stranded in West Sumatra have no known exact cause while the entangled fishing nets used in the Banggai Islands are observed to pose a serious threat to the population of the fishes. Therefore, it is important to apply Marine Ranching in different regions, reorganize the marine ecosystem to restore its original state, and develop the newly formed ecosystem to benefit fishermen. This is achievable by the creation of floating fish shelters and by placing blocks to protect the biota and the sea dam.

AUTHOR CONTRIBUTIONS

R.E. Caraka leads this study and has reviewed related kinds of literature, designed and developed the concept of all analysis prepared, writing, and edited the manuscript text. R.C. Chen performed the supervision and provide the study grant. Y. Lee performed the supervised the project and helped to provide the study grant. T. Toharudin performed the supervision and provides the research grant. C. Rahmadi provided and curated the dataset stranded marine animals and edited the manuscript. M. Tahmid provided and curated the climate dataset and edited the manuscript. A.S. Achmadi provided the dataset stranded marine animals and edited the manuscript.

ACKNOWLEDGMENTS

This study is fully supported by the National Research Foundation of Korea grants [NRF-2019R1A2C1002408] and the Ministry of Science and Technology (MOST) under Grant [107-2221-E-324-018-MY2] and [106-2218-E-324-002]. The research is a part of the Chaoyang University of Technology and the Higher Education Sprout Project, Ministry of Education (MOE), Taiwan, under the project name: "The R&D and the cultivation of talent for health-enhancement products". Therefore, it was conducted in Lab Hierarchical Generalized Linear Model (H-GLM), Department of Statistics, College of Natural Sciences Seoul National University with data and information retrieved from the preliminary research conducted by Research Center for Biology, Indonesian Institute of Sciences (LIPI) on stranded marine animals in East Java and Indonesia in 2019. The CR and ASA appreciate Anang Setyo Budi and Apandi (RC Biology LIPI) for assisting during the fieldwork in Lumajang (East Java) and for visiting the stranded *Rhincodon typus* in 2019.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

ABBREVIATIONS

AIC	Akaike information criterion
AICc	Akaike Information Criterion corrected
BIC	Bayesian Information Criterion
BMKG	Meteorology, Climatology, and Geophysical Agency is an Indonesian non-departmental government agency for meteorology, climatology, and geophysics.
cm	centimeter
df	Degree of Freedom
Eq.	Equation
Fig.	Figure

GLM	generalized linear model
Laplace approximation (LA)	Approximating Bayesian parameter estimation and Bayesian model comparison
m	meter
MGLLVM	Multivariate generalized linear latent variable models
MJO	Madden–Julian oscillation
mm	millimeter
NB	Negative Binomial
nn	Nomen nescio
Variational approximation (VA)	Techniques for making approximate inference for parameters in complex statistical models
$y_1,\ldots y_n$	sequence of independent random variables
ZIP	Zero Inflated Poisson
Var (Y)	Variance of *Y*
E (Y)	Expectancy of *Y*
ϕ	dispersion parameter

REFERENCES

Abreo, N.A.S.; Blatchley, D.; Superio, M. D., (2019). Stranded whale shark (Rhincodon typus) reveals vulnerability of filter-feeding elasmobranchs to marine litter in the Philippines. Mar. Pollut. Bull., 141: 79–83 **(5 pages).**

Anggraini, Y.; Notodiputro,K.A.; Saefuddin,A.; Toharudin,T., (2020). Latent factor linear mixed model (LFLMM) for modelling Flanders. J. Comm. Math. Biol. Neuro., 2020: 1–14 (**14 pages**).

Bartholomew, D.; Knott, M.; Moustaki, I., (2011). Latent variable models and factor analysis: A Unified Approach: 3rd Edition, Latent Variable Models and Factor Analysis: A Unified Approach: 3rd Edition **(294 pages).**

Bianconcini, S.; Cagnone, S., (2012). Estimation of generalized linear latent variable models via fully exponential Laplace approximation. J. Multivariate Anal., 112: 183–193 **(11 pages).**

Caraka, R.E.; Shohaimi S.; Kurniawan, I.D.; Herliansyah, R.; Budiarto, A.; Sari, S.P.; Pardamean, B., (2018). Ecological Show Cave and Wild Cave: Negative Binomial Gllvm's Arthropod Community Modelling. Procedia Comput. Sci., 135: 377–384 **(8 pages).**

Caraka, R. E.; Chen, R.C.; Lee, Y.; Toharudin, T.; Kurniawan, I.D.; Asmawati.S.; Ropawandi, D.; Akbar, M.F.R.; Indra; Rustan, M.F., (2020a). Latent GLM Tweedie Distribution in Butterflies Species Counts. Ad. Int. Syst. Res., 172: 306-314 (**9 pages**).

Caraka, R. E.; Chen, R.C.; Lee, Y.; Noh, M.; Toharudin, T.; Pardamean,B.; Saputra,A., (2020b). Variational approximation multivariate generalized linear latent variable model in diversity termites. Sylwan., 164(1): 161–177 (**16 pages**).

Caraka, R.E.; Nugroho, N.T.; Tai, S.K.; Chen, R.C.; Toharudin, T.; Pardamean, B., (2020c). Feature importance of the aortic anatomy on endovascular aneurysm repair (EVAR) using Boruta and Bayesian MCMC. J. Comm. Math. Biol. Neuro., 2020:1–23 (**23 pages**).

Chan, D.K.P.; Tsui, H.C.L.; Kot, B.C.W., (2017). Database documentation of marine mammal stranding and mortality: Current status review and future prospects. Disease of AQ Organisms (**10 pages**).

Colman, J. G., (1997). A review of the biology and ecology of the whale shark. J. Fish Biol., **(15 pages**).

Crawley, M.J., (2012). Regression chapter 10. The R Book., 449–497 (**1051 pages**).

Demos, T. C.; Achmadi, A.S.; Giarla, T.C.; Handika,H.; Mahardatunkamsi .; Rowe, K.C .; Esselsstyn J.A., (2016). Local endemism and within-island diversification of shrews illustrate the importance of speciation in building Sundaland mammal diversity. Molecular Ecol., 25(20): 5158–5173 (**16 pages**).

Duignan, P.J., (2003). Disease investigations in stranded marine mammals 1999–2002., DOC Science Internal Series (**32 pages**).

Fasano, G.: Franceschini, A., (1987). A multidimensional version of the Kolmogorov-Smirnov. Month Not of the Roy. Astr. Soc., 225(1): 155-170 (**16 pages**).

Ha, I.D.; Lee, Y.; Song, J.K., (2002). Hierarchical-likelihood approach for mixed linear models with censored data. Lifetime Data Anal., 2002(8): 163–176 (**13 pages**).

Herliansyah, R.; Fitia, I., (2018). Latent variable models for multi-species counts modeling in ecology, Biodi J. Biol. Div., 19(5): 1871–1876 (**6 pages**).

Huber, P.; Ronchetti, E.; Victoria-Feser, M.P., (2004). Estimation of generalized linear latent variable models. J. Royal Stat Soc. Series B: Statistic. Method. 66(4): 893-908 (**16 pages**).

Hui, F.K.C.; Wartodm, D.I.; Ormerod, J.T.; Haapaniemi, V.; Taskinen,S., (2017). Variational approximations for generalized linear latent variable model. J. Comput. Graph. Stat., 26(1): 35-43 **(9 pages).**

Jamilatuzzahro; Caraka, R.E.; Herliansyah, R., (2018).Aplikasi generalized linear model dengan R. 1st ed., Yogyakarta: INNOSAIN GI Yogyakarta (**104 pages**).

Jamilatuzzahro; Caraka, R.E.; Aprinaldy, D.; Mahadi, A., (2019). Generalized linear model multivariate poisson with artificial marginal (GLM-MPAM): Application of vehicle insurance. AIP Conference Proceedings., 2111 (020013):1-7 (**7 pages**).

Kristensen, K.; Nielsen, A.; Berg, C.W.; Skaug, H.; Bell, B.M., (2016). TMB: Automatic differentiation and laplace approximation. J. Stat. Soft., 70: 1–21 (**21 pages**).

Kuha, J., (2004). AIC and BIC: Comparisons of assumptions and performance. Soc. Meth. Res., 33(2): 188-229 **(31 pages)**

Kurniawan, I.D.; Rahmadi, C.; Caraka, R.E.; Ardi, T.E., (2018). Short communication: Cave-dwelling arthropod community of semedi show cave in gunungsewu karst area, Pacitan, East Java, Indonesia. Biodiversitas., 19(3): 857–866 (**10 pages**).

Kurniawan, I.D.; Soesilhadi, H. R.; Rahmadi,C.; Caraka, R.E.; Pardamean,B., (2018). The difference on arthropod communities structure within show caves and wild caves in Gunungsewu karst area, Indonesia. Int. J. Ecol. Environ. Conserv., 24(1): 81–90 **(10 pages).**

Lee, Y.; Nelder, J., (2001). Modelling and analyzing correlated non-normal data. J.Statistical Modeling An Int. J., 1(1): 3–16 (**14 pages**).

Lee, Y.; Noh, M. (2012). Modelling random effect variance with double hierarchical generalized linear models. J. Statistic.

Modeling. An Int. J., 12(6): 487–502 (**16 pages**).

Lomba, A.; Pellissier, L.; Randin, C.; Vicente, J.; Moreira, F.; Honrado, J.; Guisan, A., (2010). Overcoming the rare species modelling paradox: A novel hierarchical framework applied to an Iberian endemic plant. J. Biol. Conserv., 143(11): 2647–2657 (**11 pages**).

Lusher, A. L.; Milian, G.H.; Berrow, S.; Rogan, E.; O'Connor, I., (2018). Incidence of marine debris in cetaceans stranded and bycaught in Ireland: Recent findings and a review of historical knowledge. J. Environ. Pollut., 232: 467–476 (**10 pages**).

Maryanto, I.; Achmadi, A.S.; Kartono, A.P., (2008). Mamalia dilindungi perundang-undangan Indonesia. LIPI Press (**254 pages**).

Mouillot, D.; Bellwood, D.R.; Baraloto, C.; Chave,J.; Galzin, R.; Harmelin-Vivien, M.; Kulbicki, M.; Lavergne, S.; Lavorel, S.; Mouquet, N.; Paine, N.M.C.; Renaud, J.; Thuiller, W., (2013). Rare Species Support Vulnerable Functions in High-Diversity Ecosystems. PLoS One. 11(5): 1–11 (**11 pages**).

Niku, J.; Warton, D.I.; Hui, F.K.C.; Taskinen, S., (2017). Generalized Linear Latent Variable Models for Multivariate Count and Biomass Data in Ecology. J. Agric. Biol. Environ. Stat., 22(4): 498–522 (**25 pages).**

Niku, J.; Brooks, W.; Herliansyah, R.; Hui, F.K.C.; Taskinen, S.; Warton, D.I., (2019a). Efficient estimation of generalized linear latent variable models. PLoS ONE., 14(5): 1–20 (**20 pages**).

Niku, J.; Hui, F.K.C., (2019b). Gllvm: Fast analysis of multivariate abundance data with generalized linear latent variable models. J. Meth. Ecol. Evol., 1–10 (**10 pages**).

Noh, M.; Lee, Y.; Oud, J.H.L.; Toharudin, T., (2019). Hierarchical likelihood approach to non-Gaussian factor analysis. J. Stat. Comput. Simulation. 89(3): 1555–1573 **(19 pages).**

Norman, B., (2002). CITES Identification Manual Whale Shark (Rhincodon typus Smith 1829) (**18 pages**).

Rahman, D. A.; Herliansyah, R.; Rianti, P.; Rahmat, U.M.; Firdaus, A.Y.; Syamsudin, M., (2019). Ecology and conservation of the endangered banteng (Bos javanicus) in Indonesia Tropical Lowland Forest. Hayati J. Biol.., 26(2): 68–80 (**13 pages**).

Rudolph, P.; Smeenk, C., (2009). Indo-West Pacific marine mammals. In Encyclopedia of Marine Mammals. Academic Press. 608–616 (**9 pages**).

Stevens, J.D., (2007). Whale shark (Rhincodon typus) biology and ecology: A review of the primary literature. Fish. Res., (**6 pages**).

Verfuss, U.K., (2018). Comparing methods suitable for monitoring marine mammals in low visibility conditions during seismic surveys. J. Marine Pollut. Bull., 26: 1–18 (**18 pages**).

Warton, D.I., (2005). Many zeros does not mean zero inflation: Comparing the goodness-of-fit of parametric models to multivariate abundance data. Environmetrics. 16(3): 275–289 **(15 pages).**

Wibowo, F.; Dharmadi, K., (2014). Stranding of Marine Megafauna in Indonesian Waters: Causes of the Phenomenon and Handling Methods (Seastar2000). In Design Symposium on Conservation of Ecosystem. The 13th SEASTAR2000 workshop. 19–25 **(7 pages).**

AUTHOR (S) BIOSKETCHES

Caraka, R.E., Ph.D. Candidate, Department of Information Management, College of Informatics, Chaoyang University of Technology, Taiwan and Postdoctoral Researcher, Department of Statistics, College of Natural Sciences, Seoul National University, Shin Lim-Dong, Kwan Ak Ku, South Korea. Email: *rezzyekocaraka@gmail.com*

Chen, R.C., Ph.D., Distinguished Professor, Department of Information Management, College of Informatics, Chaoyang University of Technology, Taiwan. Email: *crching@cyut.edu.tw*

Lee, Y., Ph.D., Professor, Department of Statistics, College of Natural Sciences, Seoul National University, Shin Lim-Dong, Kwan Ak Ku, South Korea. Email: youngjo@snu.ac.kr

Toharudin, T., Ph.D., Associate Professor, Department of Statistics, Padjadjaran University, West Java, Indonesia. Email: *toni.toharudin@unpad.ac.id*

Rahmadi, C., Ph.D., Researcher, Museum Zoologicum Bogoriense, Research Center for Biology, Indonesian Institute of Sciences, Indonesia. Email: *cahyo.rahmadi@lipi.go.id*

Tahmid, M., S.T., Researcher, Indonesian Agency for Meteorological, Climatological and Geophysics, BMKG, Manado, Indonesia. Email: *muhammad.tahmid@bmkg.go.id*

Achmadi, A.S., Ph.D., Researcher, Museum Zoologicum Bogoriense, Research Center for Biology, Indonesian Institute of Sciences, Indonesia. Email: *anang.setiawan.achmadi@lipi.go.id*

HOW TO CITE THIS ARTICLE

Caraka, R.E.; Chen, R.C.; Lee, Y.; Toharudin, T.; Rahmadi, C.; Tahmid, M.; Achmadi, A.S., (2021). The use of multivariate generalized linear latent variable models to measure the difference in event count for stranded marine animals. Global J. Environ. Sci. Manage., 7(1): 117-130.

DOI: 10.22034/gjesm.2021.01.09

url: https://www.gjesm.net/article_44366.html