DATA ANALYTICS AND APPLIED MATHEMATICS
E-ISSN: 2773-4854
VOLUME 6, ISSUE 1, 2025, 1 –8
DOI: https://doi.org/10.15282/daam.v6i1.11140

RESEARCH ARTICLE

Conway-Maxwell-Poisson model fitting to HFMD data in Malaysia

Noraishah Mohammad Sham[1] and Khoo Wooi Chen[2]

[1]Environmental Health Research Centre, Institute for Medical Research, National Institutes of Health, Ministry of Health Malaysia, 40170 Shah Alam, Selangor, Malaysia
[2]Institute of Actuarial Science and Data Analytics, UCSI University, Cheras 56000 Kuala Lumpur, Malaysia

ABSTRACT - Since the first Hand, Foot and Mouth Disease (HFMD) outbreak occurred in Sarawak, Malaysia in 1997, the number of reported cases has remained in a cyclical pattern. Numerous HFMD research involve clinical and laboratory findings. However, there is limited research that models the distribution of the HFMD dataset in Malaysia using a count data model fitting. This study aims to demonstrate the flexibility of the Conway-Maxwell-Poisson (COM-Poisson) regression model. All daily reported cases of HFMD from 2009 to 2019 were analysed and presented by each state in Malaysia. A normal Poisson was extensively used, but it has limitations in terms of the equi-dispersion assumption. Thus, the performance of the COM-Poisson model was investigated and compared to the Poisson model by looking at the goodness of fit (log-likelihood, AIC, and BIC) test values. The results showed that COM-Poisson models fit the HFMD dataset well with lower AIC and BIC values compared to the Poisson model. The parameter coefficients' estimated values also indicated smaller values than in the Poisson model. Given the versatility of the COM-Poisson distribution, it is effective for statistical applications and procedures, such as modelling count data. In addition, the potential for related research was also being examined to improve the accuracy of the predictive model.

ARTICLE HISTORY
Received : 13th Aug 2024
Revised : 11th Feb 2025
Accepted : 5th Mac 2025
Published : 31st Mac 2025

KEYWORDS
HFMD
Over-dispersed
Time series
Malaysia

1. INTRODUCTION

Hand, Foot, and Mouth Disease (HFMD) is a common infection in infants and children. The highly contagious cohort includes a disproportionate number of children under the age of six years old. From April to June 1997, 29 infants under the age of six died of cardiorespiratory failure in Sarawak, Malaysia [1]. It is believed that the condition is caused by human enteroviruses, specifically the pathogens Coxsackie A16 and Enterovirus 71 (EV71), which are the leading causes of death in children. Pathogens, such as saliva, blister fluid, and nasal and throat secretions, are easily transmitted through direct touch. Since the outbreak, Malaysia has included HFMD as one of the notified diseases in its enforcement [2]. The surveillance aims to detect an impending outbreak, monitor the circulating infectious agent, and estimate the magnitude of HFMD and HFMD complications in the population at risk.

Health Facts Malaysia reported that between 2002 and 2021, the national incidence rate of HFMD showed an upward and downward trend. A significant increment was depicted from 2002 with 10.58 cases to 2018 with 232.62 per 100,000 population [3]. Following that, the incidence rate decreased due to the COVID-19 pandemic until 2021. However, following the government's announcement that all types of businesses could resume their operations, HFMD cases in Malaysia increased in 2022. These figures have climbed above the warning level with 889 reported outbreaks. The three states with the highest number of outbreaks were Wilayah Persekutuan (WP) Kuala Lumpur and Putrajaya, which had 309 outbreaks (35%), followed by Selangor and Perak with 111 outbreaks (12%) and 105 outbreaks (12%) respectively.

Since the first outbreak, many statistical and mathematical models have shown significant findings towards HFMD data all over the world. Each of the models has its limitations. For example, ARIMA models are good for forecasting over a short period [4]. Most of them used time series models, namely Seasonal Autoregressive Integrated Moving Average (SARIMA), Autoregressive Integrated Moving Average (ARIMA), Neural Network Nonlinear Autoregressive (NNAR), and Long Short-Term Memory Neural Network (LSTM) to predict and forecast the number of HFMD cases at the selected areas [5–8].

Despite the inclusion of clinical and laboratory findings in many HFMD studies, there is a scarcity of research that examines the method of handling count data that encounter the problem of over- or under-dispersion for HFMD using statistical tools for interpretation. Therefore, the goal of this research was to determine HFMD transmission using the Conway-Maxwell Poisson (COM-Poisson) distribution with the assumption of unequal mean and variance of the dataset. It is hoped to provide significant output from the research, which would yield substantial results from the inquiry.

***CORRESPONDING AUTHOR | Noraishah Mohammad Sham** | ✉ noraishah.ms@moh.gov.my

2. METHODOLOGY

2.1 COM-Poisson distribution

Poisson distribution is widely known as a discrete probability distribution that indicates the likelihood of a particular number of events occurring in a given time or space period. This distribution is frequently applied to various statistics to demonstrate how many events are expected to occur during a certain duration. The probability mass function with parameter λ>0 is presented below (equation 1).

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!} \tag{1}$$

where x is a random variable (a non-negative integer), and λ is a rate parameter.

On the other hand, COM-Poisson is a two-parameter generalisation or an extension to the Poisson distribution that allows for variable amounts of dispersion. It is frequently used to represent both under- and over-dispersed data. Conway and Maxwell in 1962 established this model to address the problem of equi-dispersion in count datasets, which is represented by the Poisson model, as well as over-dispersed count data in the negative binomial model [9]. The distribution used to represent both under- and over-dispersed data. The following probability mass function (equation 2) was used to visualise this flexible distribution, which had been reparameterization by Guikema and Coffelt in 2018, where two parameter augmentations were included, which allows for variable amounts of dispersion [10].

$$p(y|\lambda, v) = \left(\frac{\lambda^y}{y!}\right)^v \frac{1}{Z(\lambda, v)}, \lambda > 0, v > 0, y = 0,1,2, \ldots \tag{2}$$

where $Z(\lambda, v) = \sum_{y=0}^{\infty}(\lambda^y/y!)^v$, y is random variable (a non-negative integer), v is the dispersion parameter that controls over-dispersion, and $Z(\lambda, v)$ is the normalising constant, ensuring that the probabilities equal to 1.

2.1.1 Maximum Likelihood Estimation (MLE) for COM-Poisson

This study used R software, namely as *COMPoissonReg* package to estimate MLE for the COM-Poisson distribution. This package is the simplest way to fit the COM-Poisson model because it automatically handles the complex normalising constant $Z(\lambda, v)$. The package provides the *compois( )* function, which simplifies the process of fitting the model and estimating the parameters using MLE. The summary for both parameters used in the COM-Poisson distribution is

(i) MLE for λ is the sample mean, $\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} X_i$

(ii) MLE for v is solved numerically where the expression for the likelihood with respect to v is complex due to the normalising constant

2.2 Model selection criterion

Model selection issues are frequently addressed using information-theoretic criteria such as Akaike's Information Criteria (AIC) [11] and Bayesian Information Criteria (BIC) [12]. These two approaches were commonly used to describe data evidence in support of a model. Information criteria approaches, in particular, have been established to focus on limiting the amount of information necessary to convey the data and model [13].

2.2.1 Akaike Information Criterion (AIC)

AIC seeks to evaluate and compare the fit of different models for a given dataset. AIC is not roughly related to the size of the sample. As a result, AIC lacks crucial features of asymptotic consistency [14]. AIC is specifically designed to discover the best approximation model to the unknown real data producing process, which has been implemented by others [11,14,15]. The equation is shown in equation 3.

$$\mathrm{AIC} = 2k - 2\ln(\hat{L}) \tag{3}$$

where k is the number of parameters in the model, and $\hat{L}$ is the maximised value of the likelihood function for the estimated model given the data.

Penalty term, $2k$ penalises models that use too many parameters. Adding more parameters can increase the likelihood, but might lead to overfitting. As the number of parameters k increases, the penalty term, $2k$ increases. This discourages adding unnecessary parameters, thus preventing overfitting, where the model fits the training data well but generalises poorly to new data.

2.2.2 Bayesian Information Criterion (BIC)

BIC is developed inside a Bayesian framework, which reflects sample size, and has asymptotic consistency qualities. BIC aims to discover the genuine model, despite its similar pattern to AIC. BIC applies a bigger penalty for adequate sample sizes than AIC, therefore, when all other criteria are equal, it tends to pick simpler models than AIC. This justifies the use of BIC from a Bayesian perspective. BIC is generally defined as in equation 4.

$$\text{BIC} = -2ln(\hat{L}) + kln(n) \tag{4}$$

where n is the number of observations.

The penalty term, $ln(\hat{L})$ increases with both the number of parameters k and the number of data points n, meaning BIC becomes more conservative as the dataset grows. BIC places a stronger penalty on additional parameters than AIC, especially as the sample size n increases. This results in BIC generally favouring simpler models compared to AIC.

AIC and BIC are both used to select the best model by balancing model fit and complexity. AIC tends to favour models that predict future data well and is more lenient with adding parameters, making it suitable when readers want a good fit, even if the model is slightly more complex. BIC, on the other hand, penalises complexity more heavily, especially with large datasets, and is more likely to favour simpler models to avoid overfitting.

AIC is preferred when the goal is to predict future data accurately, while BIC is better for large datasets where readers want to avoid overfitting and choose a simpler model. Both criteria help by identifying the model with the best trade-off between fit and complexity. The model with the lowest AIC or BIC is typically the best. These criteria also help avoid overfitting, as they penalise overly complex models, guiding readers to choose simpler, more generalizable models.

2.3 Real data analysis

2.3.1 Data collection

The number of HFMD reported cases from 2009 to 2019 was obtained from the Disease Control Division, Ministry of Health Malaysia, surveillance system in daily units. The 11-year dataset was presented by 14 states in Malaysia. All acquired data goes through a cleaning process, in which any missing data or duplicates are reviewed. The missing data were double-checked with the data provider in order to be filled in, while some of them were replaced with substituted values, such as the mean case of the day.

2.3.2 Descriptive analysis

In order to describe a dataset, it should be examined using basic features. This study used a sequence plot, histogram and boxplot to give a summary of the daily data of HFMD for the study period. The sequence plot gives a visual indication of observed data convergence over time. This step is essential because it enables to scan when there is any patterns that exist in the dataset.

A histogram is a graphical representation of data in which the bars portray the frequency of the data. The count data of HFMD patients was depicted in a histogram in this case to show the frequency in each state from 2009 to 2019. A boxplot is another descriptive analysis that is used to depict the characteristics of any data, as a complement to the histogram. This plot shows how the values of any data are spread out and enables rapid graphical interpretation. Although being quite comprehensive and more direct in terms of numbers displayed than the histogram, this plot takes up significantly less space than the histogram. In this study, boxplots represent the distribution of HFMD cases in each state throughout the study period. The outliers and values of the data may be determined using this approach. The symmetrical, tightness and skewness of the dataset have also been calculated by the plot.

All routines were implemented in the R statistical computing environment [16] using a few packages that are available, such as *readxl*, *tidyverse*, *ggplot2*, *dplyr*, *lubridate*, *haven*, *mgcv* and *COMPoissonReg*. A graphical display of data was plotted using the information gathered from descriptive statistics for each state. In this specific case, a histogram and a box plot are utilised to illustrate the distribution of HFMD cases more accurately.

3. RESULTS AND DISCUSSION

Malaysia had recorded 358,188 cases of HFMD throughout the 11 years of the study period. Figure 1 shows a line graph presenting the time series of HFMD cases in Malaysia from approximately 2009 to 2019. The x-axis represents the timeline in years, while the y-axis shows the number of reported HFMD cases. The data reveals clear seasonal patterns and periodic surges in HFMD activity over the years, with outbreaks typically occurring annually and varying in intensity.

There is a noticeable increase in both the frequency and magnitude of peaks starting around 2012, with the most significant outbreak with 1,094 cases occurring on 30th July 2018, where the number of cases spiked dramatically to exceed 1,000 cases at its peak. Prior to 2012, case numbers remained relatively low and stable, rarely exceeding 200. However, from 2012 onward, multiple peaks each year became apparent, suggesting a trend of increasing HFMD activity over time. These peaks appear to follow a roughly cyclical pattern, possibly influenced by seasonal factors or environmental conditions. After the 2018 peak, case numbers gradually declined but still showed periodic surges.

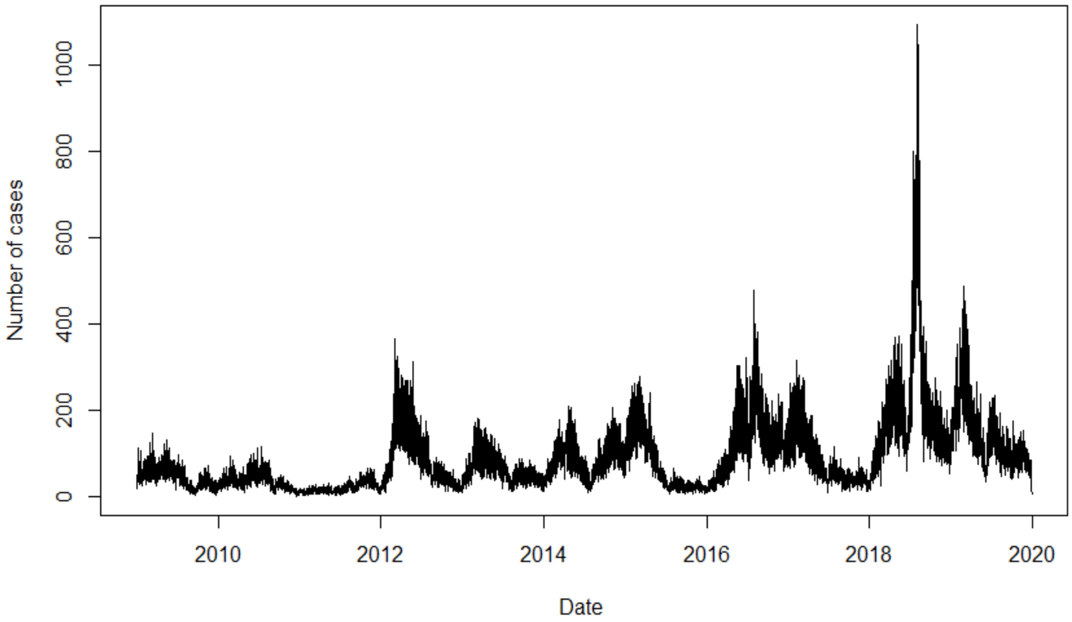

Figure 1. Number of HFMD cases in Malaysia from 2009 to 2019

Table 1 presents descriptive measures for the number of HFMD cases in each state. Sarawak has the highest number of occurrences (93,980) and represents 26.24% of the total dataset. It also has the highest mean value (23.400) and a variance of 552.290, indicating a large spread of data, but its skewness is relatively low at 1.982, suggesting a moderate asymmetry. Selangor, with 85,542 occurrences (about 23.88% of the total), has a high mean (21.300) and also high variance (642.106), indicating significant variability, though its skewness is 3.295, indicating a positively skewed distribution. Penang and Perak have moderate means (5.256 and 5.173, respectively), with Penang having a higher variance (47.430), suggesting more variability. Johor has a moderate mean (6.150) and a variance of 58.460, similar to Penang, and a higher skewness of 3.761, suggesting a more right-skewed distribution.

Smaller states like Perlis and Terengganu have lower means (around 0.718 for Perlis and 1.050 for Terengganu), with Perlis showing a higher skewness of 3.384, suggesting a strong right skew. Pahang also has a low mean (1.451) with a variance of 6.665, indicating less variability in the data. States like Negeri Sembilan and Melaka show relatively low means and moderate variances (18.245 and 19.091, respectively), with skewness values indicating moderately right-skewed distributions.

Finally, Wilayah Persekutuan and Sabah have mean values of 7.934 and 5.517, respectively, with Wilayah Persekutuan showing a slightly higher variance (108.508) and skewness (3.541), indicating higher variability and right-skewed distributions.

Table 1. Descriptive statistics of HFMD cases and rates by state from 2009 to 2019

State	Number	Percentage (%)	Mean	Median	Mode	Variance	Skewness	95% percentile
Perlis	2,883	0.80	0.718	0	0	1.704	3.384	3
Kedah	13,209	3.69	3.288	2	0	21.252	4.678	10
Penang	21,112	5.89	5.256	3	1	47.430	4.738	16
Perak	20,778	5.80	5.173	3	0	55.757	3.973	18
Selangor	85,542	23.88	21.300	13	2	642.106	3.295	65
Negeri Sembilan	10,630	2.97	2.646	1	0	18.245	4.895	10
Melaka	10,809	3.02	2.691	1	0	19.091	4.204	10
Johor	24,703	6.90	6.150	4	0	58.460	3.761	20
Pahang	5,829	1.63	1.451	1	0	6.665	3.880	6
Terengganu	4,218	1.18	1.050	0	0	3.762	3.963	4
Kelantan	10,488	2.93	2.611	1	0	22.486	3.775	11
Sarawak	93,980	26.24	23.400	16	3	552.290	1.982	70
Sabah	22,161	6.19	5.517	2	0	82.109	3.539	21
Wilayah Persekutuan	31,846	8.89	7.934	5	0	108.508	3.541	26

This study plotted a boxplot as in Figure 2, providing insights into the median, spread, and variability of cases in each state. The red line within each box represents the median number of cases, which is notably higher in Selangor and Sarawak compared to other states, indicating a consistently larger number of HFMD cases in those areas. In contrast, states such as Perlis, Terengganu, Kelantan, and Negeri Sembilan have much lower medians, typically under five cases, reflecting relatively mild or infrequent outbreaks.

The height of each box indicates the interquartile range (IQR), which represents the middle 50% of data points. Selangor and Sarawak show significantly taller boxes, highlighting a high variability in their case numbers, while many other states, including Terengganu and Kelantan, exhibit narrow boxes that suggest their case counts are tightly clustered around the lower end. The whiskers, which extend from the boxes, represent the range of values within 1.5 times the IQR from the quartiles. These are especially long for Selangor and Sarawak, reaching up to nearly 70 cases, which suggests a substantial number of high-value data points and possible outliers in these states. Although individual outliers are not explicitly plotted, the length of the whiskers in some regions implies occasional spikes in HFMD cases.

States such as Johor, Sabah, and Wilayah Persekutuan display moderate variability, with wider boxes and longer whiskers than smaller or less populous states. On the other hand, states like Perlis, Melaka, and Terengganu consistently show low and narrow boxplots, suggesting a limited and steady number of cases.

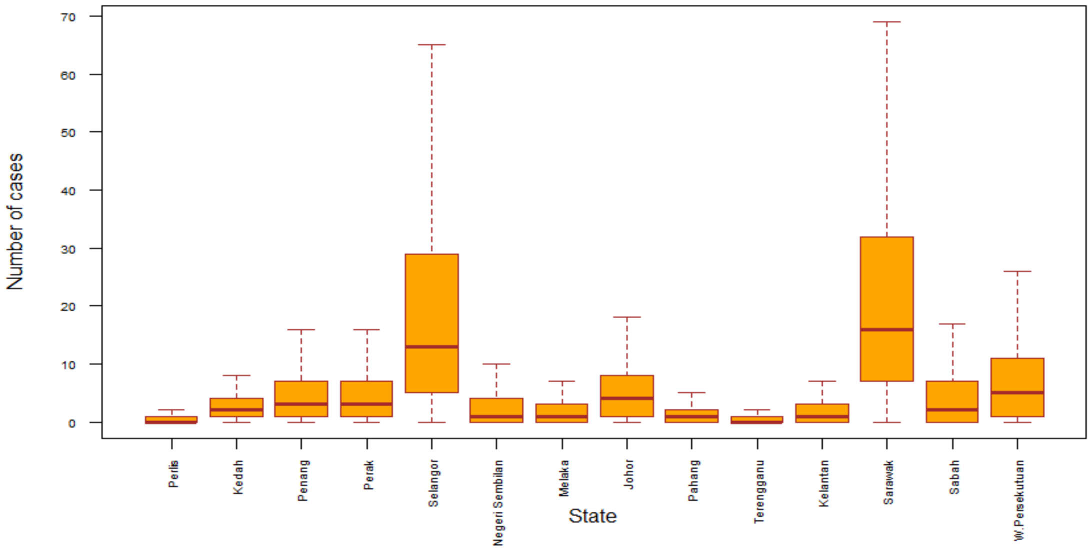

Figure 2. Boxplot of HFMD cases by states from 2009 to 2019

There were two sets of results that were analysed to compare the Poisson and COM-Poisson models. The results from the COM-Poisson regression show notable variations across the states in terms of expected counts, over-dispersion, and model fit. Perlis has the lowest expected count (λ=0.418) and a very low over-dispersion parameter (ν=0.088), indicating significant over-dispersion in the data, which suggests that the variance in the data is much higher than the mean. Despite the over-dispersion, the model for Perlis fits well with a relatively low AIC and BIC. States like Kedah (λ=0.767), Penang (λ=0.840), and Perak (λ=0.838) show moderate expected counts and values of ν close to 1, indicating that these states follow the typical assumptions of a Poisson distribution without much over-dispersion or under-dispersion. The models for these states are statistically significant with very low p-values and have reasonable AIC/BIC values, reflecting good model fit. Selangor has a relatively high λ=0.840, but its over-dispersion parameter (ν=0.119) indicates that the model may slightly underestimate the variance in the data. Sarawak also has a high expected count (λ=0.968), but with a ν=0.278, it suggests slight under-dispersion, meaning the data's variability is lower than expected under the Poisson assumption. Larger states like Selangor and Sarawak have higher AIC and BIC values, indicating more complex models, while smaller states like Perlis have lower AIC/BIC, suggesting simpler models with better fit. All models show extremely low p-values, indicating strong statistical significance and that the models are meaningful in predicting the expected counts for each state.

The Poisson regression analysis across various states reveals significant variation in the baseline expected counts, as indicated by the intercept values. Sarawak and Selangor exhibit the highest expected counts, with Sarawak showing approximately 23.4 and Selangor around 21.3. These states are predicted to have notably high counts when all predictors are zero. On the other hand, Perlis has the lowest expected count, with a value of approximately 0.72, reflecting a relatively lower baseline. States like Kedah and Penang fall in between, with expected counts of around 3.3 and 5.3, respectively. The variation in intercepts suggests that different states have distinct baseline levels, which could be influenced by local factors captured in the model's predictors.

Table 2. COM-Poisson results on HFMD cases by state from 2009 to 2019

State	Intercept	λ	v	AIC	BIC	*p-value
Perlis	-0.872	0.418	0.088	9,378.963	9,391.558	2.446e-251
Kedah	-0.265	0.767	0.125	18,706.380	18,718.975	0.000
Penang	-0.174	0.840	0.075	22,077.844	22,090.439	0.000
Perak	-0.177	0.838	0.089	21,960.719	21,973.314	0.000
Selangor	-0.174	0.840	0.119	32,776.881	32,789.476	0.000
Negeri Sembilan	-0.320	0.726	0.072	17,204.828	17,217.423	0.000
Melaka	-0.316	0.729	0.064	17,318.696	17,331.291	0.000
Johor	-0.151	0.860	0.254	23,239.492	23,252.088	0.000
Pahang	-0.524	0.592	0.051	13,312.694	13,325.289	0.000
Terengganu	-0.669	0.512	0.048	11,410.917	11,423.512	0.000
Kelantan	-0.324	0.723	0.132	17,113.321	17,125.917	0.000
Sarawak	-0.032	0.968	0.278	33,511.163	33,523.758	0.000
Sabah	-0.167	0.847	0.168	22,434.946	22,447.541	0.000
Wilayah Persekutuan	-0.119	0.888	0.102	25,144.906	25,157.501	0.000

*p-values from Chi-squared test for equi-dispersion, reject H_o if p-value < 0.05

Table 3. Poisson results on HFMD cases by state from 2009 to 2019

State	AIC	BIC	Intercept
Perlis	10,520	10529.870	-0.331
Kedah	26,590	26601.120	1.191
Penang	35,280	35286.890	1.660
Perak	39,790	39792.270	1.644
Selangor	105,900	105890.900	3.059
Negeri Sembilan	25,530	25533.100	0.974
Melaka	26,470	26474.580	0.991
Johor	40,120	40129.660	1.817
Pahang	17,660	17664.970	0.373
Terengganu	14,040	14042.560	0.050
Kelantan	28,480	28484.180	0.960
Sarawak	97,690	97692.250	3.153
Sabah	49,060	49069.470	1.709
Wilayah Persekutuan	51,870	51875.900	2.071

When examining the top three states with the highest incidence rate, namely Sarawak, Wilayah Persekutuan and Selangor, it is believed that these values were high due to the states' population size, as follows: 2,453,677, 1,982,112 and 6,994,423 [17]. In terms of computing the estimates, the COM-Poisson distribution outperformed the Poisson distribution in better fitting. COM-Poisson are able to determine the statistically significant parameters where Poisson is not able to do it due to the natural limitation of the distribution itself. The aforementioned characteristics are heavily contributed to by the presence of the dispersion parameter, v. This feature allows for the calculation of the magnitude of the decay in the distribution of the data. Thus, more accurate estimates can be calculated from these fits.

Additionally, the v parameter also makes it feasible to handle of inconsistent and irregular dataset. It is known that COM-Poisson handled these particular types of datasets, namely as under- and over-dispersed data, which are not handled by other existing distributions [18]. A study on the number of deaths in 633 cities of São Paulo State, Brazil, between 1 March and 30 April 2020, from cardiopathy, conduction disorders, cardiac arrhythmias, COVID-19 and diabetes mellitus revealed that the data are zero-inflated. Since the sample variance values vary, we can say that the data are seemingly slightly over-dispersed. In order to analyse correlated and over-dispersed multivariate count data, the multivariate COM-Poisson model is a helpful tool [19].

Another study on COVID-19 in the mainland United States presented the importance of understanding the spatiotemporal pattern of the disease using a spatio-temporal Conway Maxwell Poisson model, which performs better than a commonly used spatio-temporal Poisson model when applied to data on COVID-19 deaths [20]. However, this study has a limitation in that it did not incorporate any influential factor which may be associated with HFMD, such as meteorological parameters. This study simply demonstrates that examining different types of distributions has an effect on the plan as a preparation step for HFMD.

4. CONCLUSIONS

This study highlights a growing public health concern, with a trend toward more frequent and severe HFMD outbreaks over the decade, especially notable from 2012 onward. This pattern may indicate changes in virus transmission dynamics, reporting practices, or population susceptibility. The data shows substantial variation across states, with Sarawak and Selangor having the highest mean values and variance, while Perlis and Terengganu have the lowest. This could be due to larger populations or urbanisation. The skewness values across most states indicate a right-skewed distribution, with more pronounced skewness in some states like Perlis.

When the results of both fits were examined, it can be concluded that the COM-Poisson distribution works significantly better than the Poisson distribution in fitting, particularly when comparing the AIC values, which yielded substantially lower values. COM-Poisson regression was also provided, with most of the estimated parameters being significant. This model works incredibly as it can be adapted to both over- and under-dispersion count data. Apart from that, a projection on the average of HFMD cases for each state was also suggested to be made using the COM-Poisson regression fitted model. The projection will be working as an early warning system, which estimates the number of HFMD cases, leading the health authorities to prepare for any kind of expected scenario. Additionally, the AIC and BIC values, which are used to assess model fit, support a more robust understanding of the model's predictive power across states, with states like Wilayah Persekutuan and Johor showing relatively higher AIC and BIC values, indicating a potentially more complex or less fit model for those areas. For future work, this study recommends expanding the analysis to include other Poisson-based models, such as Generalised Additive Models (GAM), Generalised Linear Models (GLM), and other alternative approaches, to further explore their applicability and performance in modelling count data.

ACKNOWLEDGEMENTS

Institution(s)
All the authors would like to express their gratitude to General of Health Malaysia for the support and Institute for Medical Research for helping us with data cleaning and endless support.

Fund
The authors would like to extend their appreciation to the Director General of Health Malaysia for his permission to publish this article and for providing financial support under project ID NMRR-19-2710-50873.

Individual Assistant
NA

AUTHOR CONTRIBUTIONS

Noraishah Mohd Sham (Conception of the work; Data collection; Data analysis and interpretation; Drafting and critical revision of the article; Final approval for publication; Multiple rounds of manuscript revision to ensure clarity, accuracy, and scholarly rigour). Khoo Wooi Chen (Guidance and intellectual input throughout the research process-particularly in refining the study design; Validating analytical approaches; Offering critical feedback during manuscript development).

DECLARATION OF ORIGINALITY

The authors declare no conflict of interest to report regarding this study.

REFERENCES

[1] Chan LG, Parashar UD, Lye MS, Ong FG, Zaki SR, Alexander JP, Ho KK, Han LL, Pallansch MA, Suleiman AB, Jegathesan M. Deaths of children during an outbreak of hand, foot, and mouth disease in Sarawak, Malaysia: clinical and pathological characteristics of the disease. Clinical Infectious Diseases. 2000 1;31(3):678-83.

[2] Ministry of Health (Malaysia). Prevention and Control of Infectious Diseases Act 1988: Incorporating latest amendment – P.U.(A) 374/2006. 2006.

[3] Ministry of Health (Malaysia). Health Facts. 2002. Available from: https://www.moh.gov.my/index.php/-pages/view/58?mid=19

[4] Sham NM, Krishnarajah I, Shitan M, Lye MS. Time series modelon hand, foot and mouth disease in Sarawak, Malaysia. Asian Pacific Journal of Tropical Disease. 2014;4(6):469-72.

[5] Gu J, Liang L, Song H, Kong Y, Ma R, Hou Y, Zhao J, Liu J, He N, Zhang Y. A method for hand-foot-mouth disease prediction using GeoDetector and LSTM model in Guangxi, China. Scientific Reports. 2019;9(1):17928.

[6] Liu L, Luan RS, Yin F, Zhu XP, Lü Q. Predicting the incidence of hand, foot and mouth disease in Sichuan province, China using the ARIMA model. Epidemiology & Infection. 2016;144(1):144-51.

[7] Tian CW, Wang H, Luo XM. Time-series modelling and forecasting of hand, foot and mouth disease cases in China from 2008 to 2018. Epidemiology & Infection. 2019;147:e82.

[8] Yu G, Feng H, Feng S, Zhao J, Xu J. Forecasting hand-foot-and-mouth disease cases using wavelet-based SARIMA–NNAR hybrid model. PLoS one. 2021;16(2):e0246673.

[9] Conway RI. A queueing model with state dependent service rate. Journal of Industrial Engineering. 1961;12:132.

[10] Guikema SD, Goffelt JP. A flexible count data regression model for risk analysis. Risk Analysis: An International Journal. 2008;28(1):213-23.

[11] Akaike H. Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, editors. Proceedings of the 2nd International Symposium on Information Theory. Budapest: Akademiai Kiado; 1973.

[12] Schwarz G. Estimating the dimension of a model. The Annals of Statistics. 1978:461-4.

[13] Acquah H. On the comparison of Akaike information criterion and consistent Akaike information criterion in selection of an asymmetric price relationship: Bootstrap simulation results. AGRIS on-line Papers in Economics and Informatics. 2013;5(1):3-9.

[14] Bozdogan H. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika. 1987;52(3):345-70.

[15] Zucchini W. An introduction to model selection. Journal of mathematical psychology. 2000;44(1):41-61.

[16] R Core Team. R: A language and environment for statistical computing. 2020.

[17] Department of Statistics Malaysia. MyCensus 2020. 2023.

[18] Hayati M, Sadik K, Kurnia A. Conwey-Maxwell Poisson distribution: approach for over-and-under-dispersed count data modelling. In IOP Conference Series: Earth and Environmental Science 2018 (vol. 187, no. 1, p. 012039). IOP Publishing.

[19] Santana RA, Conceição KS, Diniz CA, Andrade MG. Type I multivariate zero-inflated COM–Poisson regression model. Biometrical Journal. 2022;64(3):481-505.

[20] Li X, Dey DK. Estimation of COVID-19 mortality in the United States using Spatio-temporal Conway Maxwell Poisson model. Spatial Statistics. 2022;49:100542.