INTERNATIONAL JOURNAL OF DEVELOPMENT MATHEMATICS

ISSN: 3026-8656 (Print) | 3026-8699 (Online)

journal homepage: https://ijdm.org.ng/index.php/Journals

A Study of Asthma Cases among Under-Five Children using Mixture Poisson Ar(Q) Model

Yaska Mutah[a*], Saidu S. Abdulkadir [b], Adesupo A. Akinrefon[b], Emmanuel Torsen[b]

[a]Department of Mathematics and Statistics, Federal Polytechnic, Mubi, Adamawa State, Nigeria

[b]Department of Statistics, Faculty of Physical Sciences, Modibbo Adama University, Yola, Adamawa State, Nigeria

ARTICLE INFO

Article history:
Received 02 March 2025
Received in revised form 20 June 2025
Accepted 25 June 2025

Keywords:
Asthma; Under-five children; Mixture Poisson model; Autoregressive process; Temporal dependence; Multimodality; Time series analysis

MSC 2020 Subject classification:
60E05

ABSTRACT

This study investigates asthma cases among children under five in Mubi North LGA of Adamawa State. Secondary data was obtained from National Health Management Information System as recorded in the Mubi Zonal center of Adamawa State for a period of 3 years. Through the use of Mixture Poison Autoregressive AR(q) process, a time series model was estimated. The results revealed through descriptive statistics analysis a multimodal distribution in the cases of asthma in the study area, indicating several peaks that may be attributed to distinct risk factors or subpopulations affecting the incidence of asthma. In the case of the Mixture Poisson Autoregressive model, the results also show that a three-component Mixture Poisson Autoregressive model offers the best fit to the observed asthma data. The estimated parameters of the model show considerable temporal correlations among the components and fluctuating rates of asthma cases. The significant ($p<5\%$) mixing proportions suggests that the third component plays an important role in the model, indicating a higher average rate of asthma cases. The results for the out-of-sample forecasts demonstrated considerable variability in predicted asthma cases for the next 6 days, reflecting the influence of external factors.

1. Introduction

According to Asher and Pearce (2014), asthma is one of the most common respiratory illnesses which affect children worldwide, with a very high prevalence among children under the age of five years (<5 years). Frequent hospitalization depicts the significant health burden the disease has on the healthcare system and the family. Children in this age group display high susceptibility as a result of the undeveloped nature of their immune systems, leading to high sensitivity to environmental factors, alongside the difficulty in management of asthma (Diaconu *et al*., 2024). Of interest from a public health standpoint is the temporal and spatial variation of asthma severity among this cohort (Wang *et al*., 2024). In the statistical analysis of asthma case data, the special statistical issues caused by overdispersion, time dependence (which are characteristics of count data) and the presence of underlying subpopulations with varying risk profiles has been a serious limitation (Jamaludin *et al*. 2020).

In analyzing case of asthma, Time-series methods have been frequently used by many researchers to examine the trends and patterns of the disease. Magalhães *et al*. (2021) applied a time-series study design in conducting an analysis of trends in asthma hospitalization among children's populations in Brazil. Chang *et al*. (2019) in their study analyzed asthma hospital admissions and fine particulate matter air pollution (PM2.5) in Jackson, MS, based on a time-series analytical approach. However, traditional time-series models may not accurately capture the complex and heterogeneous nature of being count data (e.g. asthma cases). Although conventional time-series approaches have been commonly employed to analyze asthma trends (e.g., Chang *et al*., 2019; Magalhães *et al*., 2021), these approaches might not properly reflect the complicated and heterogeneous nature.

* Corresponding author. Tel.: +234802 729 9629
E-mail address: yaskamutah@gmail.com (Yaska Mutah)
https://doi.org/10.62054/ijdm/0202.15

The recent development in distribution theory and statistical modeling has provided novel techniques for the analysis of data (e.g. count) that enable researchers to identify patterns and trends that might otherwise not be apparent (David et al., 2024, Mathew et al., 2024). One such method is the Mixture Poisson Autoregressive model (Mixture Poisson AR(p)), proposed by Yaska *et al*. (2025), which possesses a systematic approach to the modeling of count data that may possess overdispersion and time dependencies. The model combines the flexibility of mixture models, which identify and detect latent subgroups in the population, with the structure of autoregressive processes, which model temporal dependencies in the data. This article is a study of asthma in children less than five years old based on this Mixture Poisson AR(q) model for analyzing the prevalence of asthma in children less than five years old in Mubi North Local Government Area of Adamawa State.

2 Literature Review

Yang *et al*. (2024) analyzed global regional trends in childhood asthma prevalence, mortality, and disability in the 2019 Global Burden of childhood disease study. Their study found that prevalence and incidence rates are increasing, also, death rates and disability-adjusted life years (DALYs) have decreased, especially in the Western Pacific countries such as Japan have made great progress in reducing the asthma burden. The study emphasized the need for continued collaborative efforts to address persistent disease burdens worldwide. Magalhaes, *et al*. (2021) utilized time series analysis to examine hospital admissions for asthma in children and adolescents in Brazil from 2008 to 2017. Their findings revealed that the proportion of admissions in Higher hospital rates for children and boys aged 5-9 years were highest in the Northeast region. Trend analysis indicates that hospital admission rates nationwide are decreasing which indicated an improvement in asthma management during the study period.

Jamaludin *et al*. (2020) conducted a detailed investigation into asthma cases in Johor Bahru, a region experiencing rapid development and facing significant pollution challenges, which have contributed to the prevalence of asthma in the area. The study aimed to analyze the behavior of asthma cases using a count-based statistical approach, specifically the Poisson Integer Generalized Autoregressive Conditional Heteroscedasticity (Poisson-INGARCH) and Negative Binomial INGARCH (NB-INGARCH) models, incorporating both identity and log link functions. René *et al*. (2019) examined the impact of biomass combustion on asthma risk in children under five in Yopougon municipality. They found that indoor air pollution from biomass fuel significantly increased asthma risk, with butane gas emerging as a safer alternative to reduce pollutants. The findings underscored the role of environmental interventions in mitigating asthma triggers.

Chang *et al*. (2019) investigated the relationship between ambient PM2.5 pollution and asthma hospitalizations in Jackson, Mississippi, from 2003 to 2011. They observed a 7.2% increase in daily asthma admissions for every 10 $\mu g/m^3$ rise in PM2.5 levels, particularly among males and Black individuals. These findings highlight the adverse health effects of fine particulate matter, even in areas with relatively low pollution levels. Adelou *et al*. (2016) studied the prevalence and pattern of acute asthma exacerbations in children presenting to the emergency room of the University of Nigeria Teaching Hospital, Enugu. Their study showed that in cases Most are mild to moderate. The prevalence is higher during the rainy season and also reported that children taking controlled drugs were less likely to develop severe disorders.

3 Methods

This section presents the methods utilized in the analysis as well as the source of the data. The data on Asthma was collected from National Health Management Information System as recorded in the Mubi Zonal center of Adamawa State.

Autoregressive Processes AR (p)

The autoregressive process implies a model which explains the present value of the series, C_t, as a function of the lagged values of the response variable $(C_{t-1}, C_{t-2}, \ldots, C_{t-q})$. In mathematical form, an autoregressive process of order pp is written as:

$$C_t = \phi_1 C_{t-1} + \phi_2 C_{t-2} + \cdots + \phi_p C_{t-p} + e_t \equiv \sum_{i=1}^{q} \phi_i C_{t-i} + e_t \tag{1}$$

where e_t is white noise, ϕ_i are the autoregressive coefficients, and pp is the order of the autoregressive process.

For a stationary mean μ,

$$E\left[E\left(C_t \mid C_{t-1}\right)\right] = E\left[\sum_{i=1}^{p} \rho_i C_{t-i} + \lambda\right] \text{and} \quad E[C_t] = \sum_{i=1}^{q} \rho_i E[C_{t-i}] + \lambda \tag{2}$$

It can be shown that equation (2) converges to a geometric series if $\sum_{i=1}^{p} \rho_i < 1$, yielding: $\lambda = \left(1 - \sum_{i=1}^{q} \rho_i\right)\mu$.

Since $E(C_0) = \mu$, then equation (2) can be expressed as: $E[C_t \mid C_{t-1}] = \sum_{i=1}^{q} \rho_i C_{t-i} + \left(1 - \sum_{i=1}^{q} \rho_i\right)\mu$, which defines a stationary linear AR(p) process. This model has two major components: $\Pr(C_t \mid m_t) = \frac{e^{-m_t} m_t^{C_t}}{C_t!}$, a Poisson distribution describing the conditional mean of a homogeneous Markov process, and the transition equation: $m_t = \sum_{i=1}^{q} \rho_i C_{t-i} + \left(1 - \sum_{i=1}^{q} \rho_i\right)\mu$.

Mixture of Poisson Autoregressive (MPAR) model

Yaska *et al*. (2025) showed that the Mixture Poisson Autoregressive (q) model can be expressed as:

$$g(c) = \sum_{j=1}^{k} \pi_j \left\{ \sum_{i=1}^{q} \rho_{ij} c_{t-1} + \left(1 - \sum_{i=1}^{q} \rho_{ij}\right) \mu_j \right\} \tag{2}$$

where; π_j= mixing weight, $\sum_{j=1}^{k} \pi_j = 1$ and π_j is nonnegative; k= number of components in the mixture distribution; ρ_{ij}= coefficient of autoregressive parameters for the k^{th} components; $q =$ total number of lags; c_t= current observation; c_{t-1}= immediate past observation; μ_j= mean of a give k^{th} component; c_{t-i}= previous c_{t-1} observations

Expectation–Maximization (EM) Algorithm

Steps Involves in Fitting the Model

The likelihood for a mixture model can be determined as:

$$L(\omega) = f\left(y_1, y_2, \ldots, y_{t-i}; \rho_1, \rho_2, \ldots, \rho_k\right) \doteqdot \prod_{i=1}^{n} f\left(y_i; \rho_1, \rho_2, \ldots, \rho_k\right)$$

$$L(\omega)=\prod_{i=1}^{n}\sum_{j=1}^{k}\pi_k f\left(y_i;\rho_k\right)L\left(\rho_1,\rho_2,\dots,\rho_k\right)=\sum_{i=1}^{n}\log\sum_{j=1}^{k}\pi_i f\left(y_i;\rho_k\right) \tag{3}$$

where ≝ is used for the assumption that $y_1,\dots,y_n$ are independent and identically distributed.

Obtaining the log likelihood is to determine which component does the variable or sample data belongs to. Let $Z_{i,k}$ being the unobserved variable assuming two components

$$Z_{i,k}=\begin{cases}1 & \text{if } y_i \text{ belongs to } k,\\ 0 & \text{otherwise.}\end{cases}$$ and its corresponding probabilities stands as: $$\begin{cases}P(Z_{i,k}=1)=\pi_k\\ P(Z_{i,k}=0)=1-\pi_k\end{cases}$$

Therefore, the log-likelihood can be written as:

$$\ell\left(\rho_1,\dots,\rho_k\right)=\sum_{i=1}^{n}\left[\sum_{j=1}^{q}Z_{ij}\log\left(\pi_i f\left(c_i;\rho_p\right)\right)\right] \tag{4}$$

The Z_{ij} here is the incomplete (missing) datum because we do not know whether it is $Z_{ij}=0$ or $Z_{ij}=1$ for c_i and a specific k. Therefore, using the EM algorithm, we try to estimate it by its expectation.

The E-step in EM-algorithm

$$Q(\omega)=\sum_{i=1}^{k}\left[\sum_{j=1}^{q}\left(E[Z_{i,k}\mid y_{t-1},\lambda,\rho]\cdot\log\left(\pi_k f(y_{t-1},\lambda,\rho)\right)\right)\right] \tag{5}$$

The $Z_{i,k}$ is either 0 or 1; therefore:

$$E[Z_{i,k}\mid y_{t-1},\lambda,\rho]=0\cdot P(Z_{i,k}=0\mid y_{t-1},\lambda,\rho)+1\cdot P(Z_{i,k}=1\mid y_{t-1},\lambda,\rho)+1\cdot P(Z_{i,k}=1\mid y_{t-1},\lambda,\rho)$$

Using Bayes rule,

$$P(Z_i=1\mid y_{t-1},\lambda,\rho)=\frac{P(y_{t-1},\lambda,\rho,Z_i=1)}{P(y_{t-1},\lambda,\rho)}P(Z_i=1\mid y_{t-1},\lambda,\rho)=\frac{P(y_{t-1},\lambda,\rho\mid Z_i=1)P(Z_i=1)}{\sum_{j=0}^{k}P(y_{t-1},\lambda,\rho\mid Z_i=j)P(Z_i=j)} \tag{6}$$

Assuming for k component the marginal probability in the denominator is:

$$P(y_{t-1},\lambda,\rho)=\pi_1 f_1(y_{t-1};\rho_1)+\pi_2 f_2(y_{t-1};\rho_2)+\cdots+\pi_k f_k(y_{t-p};\rho_k)$$

Thus assume

$$\hat{\upsilon}_{ik}^{v}=\frac{\hat{\pi}^{v-1}f_1(y_{t-1},\lambda,\rho)}{\hat{\pi}_1^{v-1}f_1(y_{t-1},\lambda,\rho)+\hat{\pi}_2 f_2(y_{t-1},\lambda,\rho)+\cdots+\hat{\pi}_k f_k(y_{t-1},\lambda,\rho)} \tag{7}$$

where $\hat{\upsilon}_i=E[Z_i\mid y_{t-1},\lambda,\rho]$ is called responsibility of y_i (Friedman *et al,* 2009).

$$Q(\omega)=\sum_{i=1}^{n}\left[\hat{\upsilon}_1\log\left[\pi_1 f_1(y_{t-1},\lambda,\rho)\right]+\hat{\upsilon}_2\log\left[\pi_2 f_2(y_{t-1},\lambda,\rho)\right]+\cdots+\hat{\upsilon}_k\log\left[\pi_k f_k(y_{t-p},\lambda,\rho)\right]\right] \quad (8)$$

The M-Step

The M-Step in a mixture model, the algorithm computes new parameter values that maximize the expected log-likelihood,

$$\hat{\varnothing}_k,\hat{\pi}_k=\arg\max_{\varnothing_1,\pi_k}\left[Q(\rho_1,\ldots,\rho_p,\pi_1,\ldots,\pi_k)\right] \quad (9)$$

$$iif \sum_{k=1}^{K}\pi_k=1$$

The equation (9) is a problem that can be solve through optimization as follows:

Using Lagrange multiplier by introducing a new variable ∝ known as Lagrange multiplier

$$L(\rho_1,\ldots,\rho_k,\pi_1,\pi_2,\ldots,\pi_K,\alpha)=Q(\rho_1,\ldots,\rho_k,\pi_1,\pi_2,\ldots,\pi_K)-\alpha\left(\sum_{k=1}^{K}\pi_k-1\right)$$

$$=\sum_{i=1}^{n}\sum_{k=1}^{K}\left[\hat{\upsilon}_{i,k}\log\pi_k+\hat{\upsilon}_{i,k}\log f_k(x;\rho_k)\right]-\alpha\left(\sum_{k=1}^{K}\pi_k-1\right) \quad (10)$$

$$\frac{\partial L}{\partial\rho_k}=\sum_{i=1}^{n}\left[\frac{\hat{\upsilon}_{i,k}}{f_k(c_i;\rho_k)}\frac{\partial f_k(c_i;\rho_k)}{\partial\rho_k}\right]=0 \quad (11)$$

$$\frac{\partial L}{\partial\rho_k}=\sum_{i=1}^{n}\frac{\hat{\upsilon}_{i,k}}{\pi_k}-\alpha=0 \quad\Rightarrow\quad \pi_k=\frac{1}{\alpha}\sum_{i=1}^{n}\hat{\upsilon}_{i,k}\ ;\ \frac{\partial L}{\partial\alpha}=\sum_{k=1}^{K}\left(\pi_k-1\right)=0 \quad\Rightarrow\quad \sum_{k=1}^{K}\pi_k=1$$

$$\therefore\sum_{k=1}^{K}\frac{1}{\alpha}\sum_{i=1}^{n}\hat{\upsilon}_{i,k}=1 \quad\Rightarrow\quad \alpha=\sum_{i=1}^{n}\sum_{k=1}^{K}\hat{\upsilon}_{i,k}$$

$$\therefore\hat{\pi}_k=\frac{\sum_{i=1}^{n}\hat{\upsilon}_{i,k}}{\sum_{i=1}^{n}\sum_{k=1}^{K}\hat{\upsilon}_{i,k}} \quad (12)$$

Chose the value of ω, say $\omega^{(v)}$, that maximizes $Q(\omega,\omega^{(v)})=L_c(\omega)$ from E-step with Z_{ij} replaced with υ_{ij}. The E-step and M-step are alternate repeatedly, wherein their subsequent executions, where the initial fit parameters ω replaced by the current $\omega^{(v)}$, say $\omega^{(v-1)}$ at the *k-th* stage. The process stops when it converges, say when the difference between the current parameters and the previous or initial parameters is 0.0001 or $L(\omega^{(v+1)})\geq\omega^{(v)}$. Using equations (11) and (12) we obtain the new values of the estimations $\hat{\rho}_k,\hat{\lambda}_k$ *and* $\hat{\pi}_k$ in every iteration.

Estimating the number of components

To estimate the number of required components in a given set of data, the Akaike Information Criterion (AIC) was adopted. The expected number of components is $k > 1$ since. The number of components with least BIC is said to be the best fit for the data.

$$AIC = n \times In(m_t) + k \times In(n) \tag{13}$$

where m_t is conditional mean with parameters as $\rho, \mu\ and\ \lambda$

Hartigan's Dip Test Statistic for Unimodality

The Dip test statistic determine whether a dataset is unimodal or multimodal. The test statistic, known as the Dip statistic, measures the maximum difference between the empirical distribution function and the unimodal distribution function at any point in the sample (Hartigan & Hartigan 1985).

Test Statistic: The test statistic is called the "dip" statistic, denoted as;

$$D = \max\left|F\ \ x\ -U\ \ x\right| \tag{14}$$

where: F(x) is the empirical distribution function *U(x)* is the unimodal distribution function that minimizes the maximum difference.

4 Results

The data on Asthma was collected from National Health Management Information System as recorded in the Mubi Zonal center of Adamawa State. The data consisted of daily and weekly records of counts on Asthma among children of under five years was for a period of 240 days a total of 1334 reported cases.

Table 1: Descriptive Statistic

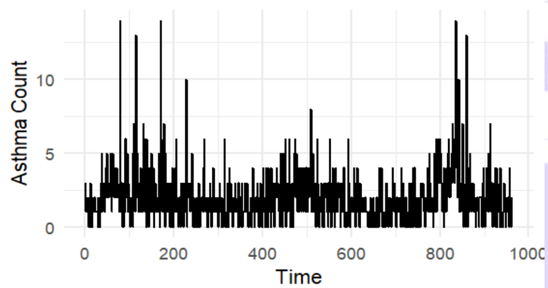
Cases	Minimum	Maximum	Mean	Standard Deviation	Variance	Skewness	Kurtosis
Asthma	0	14	2.000	1.7553	3.081	2.237	10.061

Table 2: Hartigans' Dip Test of Unimodality

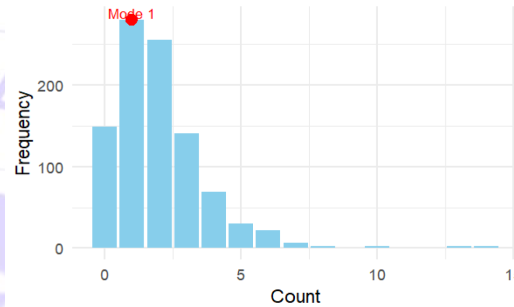
	Test Statistic	P-value
Asthma	0.13292	2.2e-16

Table 1, the descriptive statistics provide a summary of the central tendency, dispersion, and shape of the asthma case data among under-five children. The mean number of cases is 2.000, with a minimum of 0.00 and a maximum of 14.00. The variance is 3.081, indicating moderate variability in asthma incidence. The skewness is 2.237, which indicates a highly positive skew, suggesting that while most children experienced few or no asthma symptoms, a small proportion of the population experienced very high frequencies of asthma. Furthermore, the kurtosis value of 10.061 implies a leptokurtic distribution, indicating a higher peak and heavier tails than a normal distribution. This suggests that extreme values i.e., children with recurrent asthma are more frequent than expected under normality assumptions. These characteristics highlight the heterogeneous nature of the data and suggest the presence of subpopulations with varying asthma risks.

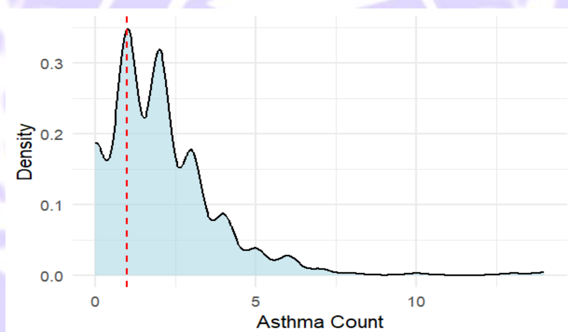
Table 2 presents the Hartigan's Dip Test of Unimodality for asthma cases, which assessed whether the distribution follows a unimodal or multimodal pattern. The D-statistic, with a value of 0.13292, is relatively large, and the p-value is extremely small (2.2e-16). Therefore, this finding leads to the rejection of the null hypothesis of unimodality, implying that the distribution of asthma cases among children aged less than five years is multimodal. The existence of more than one peak in the distribution, as revealed in Figure I, may be due to the presence of various underlying subpopulations, varying etiologies, or different risk factors influencing asthma-related admissions. The results also validate the suitability of applying the Mixture Poisson AR(p) model to the study, given its capability in adequately addressing both the temporal dependence and the intrinsic heterogeneity of the data.



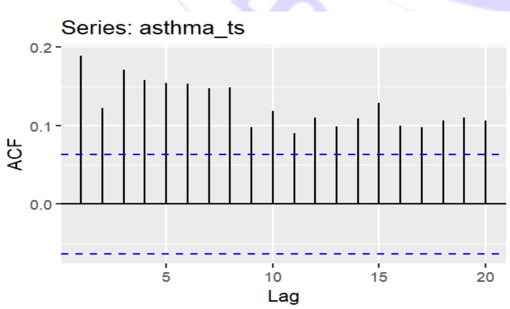
***a*:Original Time series Plot of Asthma data**



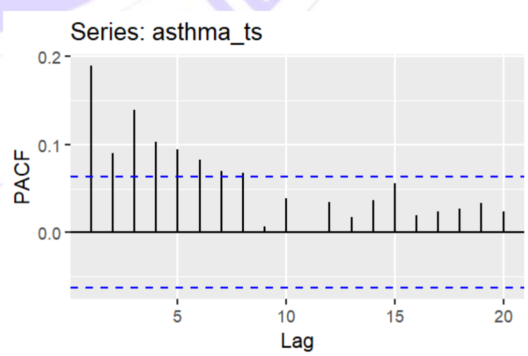
***b:* Histogram of Asthma data**



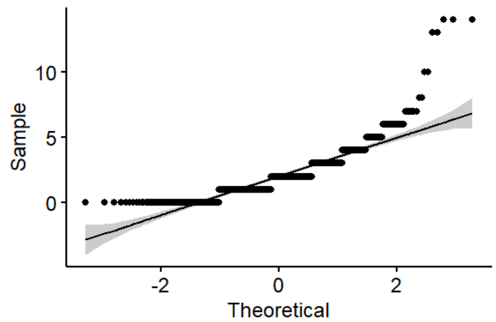
***c:* Density Plot**



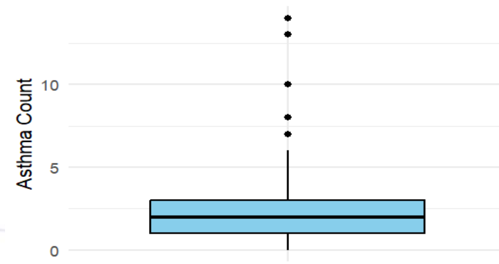
***d:* Autocorrelation plot for Asthma data**



***e:* Partial Autocorrelation for Asthma Data**



***f:* Probability Plot for Asthma data**



***g:* Boxplot for Asthma data**

Figure 1: Visualizations Summarizing the Distribution and Temporal Behaviour of Asthma Cases

Figure 1 present various visualizations of summarizing the distribution and temporal behavior of asthma data. *a:* The original time series plot of asthma cases with temporal variation. *b:* Histogram of asthma data showing frequency distribution. *c:* Density plot displaying the shape and skewness of the distribution. *d:* Autocorrelation plot showing the existence of serial correlation in the data. *e:* Partial autocorrelation plot indicating the strength of correlation at varying lags. *f:* A probability plot verifies if the dataset is normally distributed and *g:* A boxplot shows range, mean, and outliers in asthma data.

Figure 1 indicates the characteristics associated with the Asthma data, the top left graph is the original time series plot of the data this shows that the data is time series as the graph line indicates sine waves up and down movement of the points depending on the number of counts at that particular time, and the histogram at the top right hand is showing multiple components or subpopulations as the bars shows grouping due factors influencing the data, both the Autocorrelation and the partial autocorrelation located at the center implies that the data is non-stationary as some of the spikes are outside the control line of the data, the probability plots shows non-normality of the Asthma data this is because some of the data falls below the normality line of the plot and the boxplot indicates overdispersion of the Asthma data this is because some data falls outside the whisker line, and non-symmetric of Asthma data as one tail of the boxplot is longer than the other, several data points appeared to be an outlier as they are outside the whisker lines.

Model Selection

The number of components with the least AIC shall be selected as the best fit for the set of data.

Table 3: Model Selection for Asthma Cases

Datasets	Components	AIC
Asthma	2	4624.735
	3	4234.709*
	4	5051.558
	5	4614.215

** Best fit model*

Table 3 presents the results of model selection for a Mixture of Poisson Autoregressive models applied to Asthma dataset. The models are compared based on the number of components (2, 3, 4, and 5) and Akaike Information Criterion (AIC) was adopted as the measurement value. The model with the lowest AIC value is considered the best fit for the data, as indicated by an asterisk (*). The results revealed that the model with 3 components is also the best fit, with AIC value of 4234.709. This indicates that a three-component model is most appropriate for modelling the asthma data.

Application of MPAR Model to Asthma Data Set

Table 4: Parameter Estimates for Asthma data

Components	π	Parameters	Estimates	Std Errors	Z-score	p-value
1	0.2497	λ	0.244	0.0623	3.9183	0.0001
		ρ	0.8071	0.0089	90.8756	0.0000
2	0.1811	λ	0.177	0.0623	2.8426	0.0045
		ρ	0.4129	0.0089	46.4966	0.0000
3	0.5692	λ	0.5559	0.0623	8.9283	0.0000
		ρ	0.3715	0.0089	41.8324	0.0000

Table 4 presents the parameter estimates of the Mixture Poisson Autoregressive (Mixture Poisson AR) model that has been applied to the asthma dataset. The model comprises three distinct components, each described by different set of parameters: λ (mean), ρ (parameter at lag 1), and π (mixing proportion). In regard to the first component, an estimated value of λ is 0.2440 with a standard error of 0.0623. A corresponding Z-score of 3.9183 is statistically significant (p-value = 0.0001), and therefore this component is significantly contributing towards the model. The autoregressive parameter ρ is estimated to be 0.8071 with a corresponding standard error of 0.0089 and a highly significant Z-score of 90.8756 (p-value = 0.0000), and hence there is strong temporal dependence. A mixing proportion for this component is estimated to be 0.2497.

Regarding the second factor, the estimate of λ is 0.1770 with a standard error of 0.0623. The Z-score of the latter is 2.8426, which conveys statistical significance (p-value = 0.0045). Additionally, the estimate of ρ is 0.4130 with a standard error of 0.0089 and a Z-score of 46.4966, denoting high statistical significance (p-value = 0.0000). Further, the mixing proportion is 0.1811. The third component has an even higher estimated λ value of 0.5559 with a standard error of 0.0623 and a Z-score of 8.9283, which is statistically significant (p-value = 0.0000). The autoregressive parameter ρ is estimated to be 0.3715 with a standard error of 0.0089 and a Z-score of 41.8324, which is also statistically significant (p-value = 0.0000). The mixing proportion for this component is estimated to be 0.5692, meaning it contributes the most to the mixture overall.

The mean parameter λ is the number of asthma instances predicted for a specified time. A λ of 0.2440 in the first component indicates a relatively low mean rate of asthma incidence in children who are less than five years old. The second component with a λ of 0.1770 indicates an even lower rate, maybe a subpopulation with limited exposure or diminished risk factors. Contrarily, the third factor offers a λ value of 0.5559, suggesting an elevated average rate that may be associated with children showing greater vulnerability or those experiencing frequent exposure to asthma triggers. The autoregressive coefficient ρ quantifies the temporal correlation observed in asthma events. The high ρ value of 0.8071 for the first component indicates strong persistence of the symptoms of asthma over time, which means that following an initial attack of asthma in a child, subsequent attacks are probable. Additionally, the ρ values of 0.4130 and 0.3715 for the second and third components, respectively, also give evidence for the presence of moderate but strong temporal correlation, indicating that asthma events are not independent over time.

The mixing proportions, denoted by π, are the relative sizes of the subpopulations. The third component (π = 0.5692) is the dominant component in the distribution, implying a high percentage of children belong to this high-risk category. In contrast, the first and second components with proportions 0.2497 and 0.1811, respectively, imply smaller populations with lower prevalence of asthma. The importance of λ and ρ parameters in all the components shows the strength of the Mixture Poisson AR model in its ability to satisfactorily explain both the frequency of asthma events and the respective temporal trends. The results corroborate the assumption that asthma events in children under the age of five years in the Mubi North Local Government Area are influenced by underlying subpopulations with varying asthma risk profiles and temporal trends, thereby justifying the application of a mixture-based time series modeling framework.

Forecasting

The out-of-sample forecast values for asthma cases (presented in Table 5) over the next 6 days show a more variable pattern, with forecast values ranging from 0 to 9. The results suggests that asthma cases may not follow a clear seasonal trend but rather fluctuate with peaks and troughs throughout the period. The highest number of cases is predicted on day 3 (9 cases), and the lowest on day 4 having no reported cases with several days showing no predicted cases. This variability may reflect the influence of multiple factors such as air quality, weather changes, or individual triggers that can vary on a daily basis. The forecasts for the health conditions in Mubi North Local Government Area provide valuable insights into potential future health service needs and can inform public health planning and resource allocation.

Table 5: Forecasting Values for Asthma Cases in the next 6 days in year 2024

Days	1	2	3	4	5	6
Forecast Value	2	3	9	0	2	8

5. Discussion

Results of the Mixture Poisson Autoregressive (Mixture Poisson AR) model applied to asthma incidence data in Mubi North show the presence of three subpopulations with varying risk levels of asthma and time dependencies. The estimated mean rates (λ) and mixing proportions (π) show that a large proportion of children belong to a high-risk group with a very high average incidence of asthma (λ = 0.5559; π = 0.5692). These observations are consistent with the reports of Yang *et al*. (2024) and René *et al*. (2019), who pointed out regional variations in asthma prevalence and the roles of environmental exposures, such as indoor air pollution, in increasing asthma risk in children. Additionally, identification of subgroups with lower incidence is in line with findings from Magalhães *et al*. (2021) of improved asthma control and the need for healthcare intervention in susceptible groups. The autoregressive coefficients (ρ) in all model components being significant suggest the presence of strong temporal dependence of asthma events, i.e., past events strongly influence future episodes. This is consistent with the modelling approach of Jamaludin *et al*. (2020), who modelled similar dynamics for asthma data using autoregressive count models. That asthma attacks are persistent, particularly for the highest-risk group, underscores the chronic nature of the disease and the need for constant monitoring and intervention. These results complement Chang *et al*. (2019) and Adelou *et al*. (2016), who recognized increased asthma events associated with environmental determinants and seasonal trends.

6. Conclusion

The study examined childhood asthma incidence in Mubi North Local Government Area using a Mixture Poisson Autoregressive model. Results on multimodal distribution of cases of asthma presented many peaks reflecting various risk factors, with the best fit being by the three-component model with strong temporal correlations. Out-of-sample prediction demonstrated significant variability owing to exogenous factors like meteorological and environmental conditions, and predictive significance was evident in the high-risk subgroup. The findings validate population heterogeneity for asthma incidence and establish the applicability of Mixture modelling for capturing complex inherent in time series count data. Identification of distinct risk subpopulations validates targeted intervention strategies for particular exposure types and evidence-based policy. Compared to previous drawing on aggregated data, this research used individual-level observations within a probabilistic context and provided greater precision in characterizing patterns of incidence and forming a methodological foundation for epidemiological surveillance activity in limited-resource environments.

References

Asher, I., & Pearce, N. (2014). Global burden of asthma among children. *The international journal of tuberculosis and lung disease, 18*(11), 1269-1278.

Chang, H. H., Pan, A., Lary, D. J., Waller, L. A., Zhang, L., Brackin, B. T., ... & Faruque, F. S. (2019). Time-series analysis of satellite-derived fine particulate matter pollution and asthma morbidity in Jackson, MS. Environmental monitoring and assessment, 191, 1-10.

David I. J., & Mathew, S. (2024). Statistical Properties and Applications of Transmuted Skew Student t Distribution. *Reliability: Theory & Applications, 19*(3), 144-156.

Diaconu, I. D., Gheorman, V., Grigorie, G. A., Gheonea, C., Tenea-Cojan, T. S., Mahler, B., ... & Văruț, R. (2024). A Comprehensive Look at the Development of Asthma in Children. Children, 11(5), 581.

Adelu, B. O., Eze, J. N., Ayuk, A. C., & Oguonu, T. (2016). Prevalence and pattern of asthma exacerbation in children seen at the University of Nigeria Teaching Hospital, Enugu. *Nigerian Journal of Paediatrics, 43*(2), 78-82.

Hartigan, J. A., & Hartigan, P. M. (1985). The dip test of unimodality. *The annals of Statistics*, 70-84.

Jamaludin, A. R., Yusof, F., & Suhartono, S. (2020). Modelling asthma cases by count analysis approach: Poisson INGARCH and negative binomial INGARCH. Matematika, 15-30.

Magalhães, L. S., Policena, G. M., Carneiro, V. S. M., Costa, L. D. C., Costa, M. S. N. D., & Vieira, M. A. D. S. (2021). Analysis of the trend of hospitalizations for asthma in children and adolescents in Brazil. *Jornal de Pediatria, 97*(3), 309-314.

Mathew, S., David I. J. & Mutah Y. (2024). A Study on One-Parameter Entropy-Transformed Exponential Distribution and Its Application. *Journal of Reliability and Statistical Studies. 17*(1); 17–44.

René, K. A. K., Kouadio, K., Siele, S., Harvey, A. T., M'begnan, C., Leandre, K., & Veronique, Y. (2019). Prevalence of asthma in children under 5 years old exposed to air pollution in Abidjan,(Côte D'ivoire). *International Journal of Recent Scientific Research 10*(7) 33353-33358.

Wang, J., Cortes-Ramirez, J., Gan, T., Davies, J. M., & Hu, W. (2024). Effects of climate and environmental factors on childhood and adolescent asthma: A systematic review based on spatial and temporal analysis evidence. Science of the Total Environment, 175863.

Yang, C. H., Li, X. Y., Lv, J. J., Hou, M. J., Zhang, R. H., Guo, H., & Feng, C. (2024). Temporal Trends of Asthma Among Children in the Western Pacific Region From 1990 to 2045: Longitudinal Observational Study. *JMIR Public Health and Surveillance, 10*(1), e55327.

Yaska, M., Abdulkadir, S. S., Akinrefon, A. A., & Torsen, E. (2025). A study on mixture Poisson autoregressive (P) model. *Mikailalsys Journal of Mathematics and Statistics, 3*(2), 382-394.