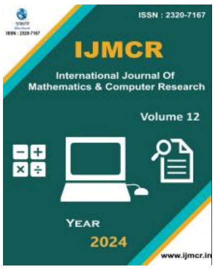
International Journal of Mathematics and Computer Research
ISSN: 2320-7167
Volume 12 Issue 01 January 2024, Page no. 3933-3936
Index Copernicus ICV: 57.55, Impact Factor: 7.362
DOI: 10.47191/ijmcr/v12i1.01



The Comparison some Models (Gamma, Slashed Quasi-Gamma, three Parameters Quasi Gamma, two Parameter Gamma-Exponential, and Modified Log-Logistic Distributions) for Daily Temperature Data

Ernila Sari Gulo[1], Rado Yendra[2], Muhammad Marizal[3], Rahmadeni[4]
[1,2,3,4] Department of Mathematics, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau

ARTICLE INFO
Published Online:
02 January 2024

Corresponding Author:
Muhammad Marizal

ABSTRACT
With the increase in hot temperature events in recent years, there is growing interest in measuring the frequency of recurring temperature events. One of the basis for assessing the frequency of recurrence of temperature events is the probability distribution of temperature events, therefore temperature data is needed to produce statistical modeling, especially in determining the best probability distribution. The study intended to estimate the best-fitted probability model for the daily temperature at the Pekanbaru station in Indonesia from 2010 to 2020 using several statistical analyses. Five continuous probability distributions such as Gamma (GM), Slashed Quasi-Gamma (SQG), Three Parameters Quasi Gamma (TPQG), Two Parameter Gamma-Exponential (TPGE), and Modified Log-Logistic (MLL) distributions were fitted for these tasks using the maximum likelihood technique. To determine the model's fit to the temperature data, several goodness-of-fit tests were applied, including the graphical methos test (density plot) and Numerical criteria method test (AIC, BIC, and - Log Likelihood). The GM and MLL distribution are found to be the best-fitted probability distribution based on goodness-of-fit tests for the daily temperature data at the Pekanbaru station.

KEYWORDS: Gamma Distribution, Slashed Wuasi-Gamma, Three Parameters Quasi Gamma, Two Parameter Gamma-Exponential, Modified Log-Logistic, Temperature.

I. INTRODUCTION

Today's erratic temperature changes can cause great damage to us living through events such as droughts, floods and ecological disturbance because it affects human activities and the economy too. Temperature evaluation shows an important role in reviewing continuous variations in temperature falls or rise. Many different works have been carried out in response to environmental climate change, and conclusions about the critical and difficult fall and rise of ambient temperature have been drawn. Researchers are interested develop appropriate statistical methods for temperature events which provides significant help to this problem. In recent years, there have been several studies on this matter temperature events and analyzed extensively over the past two [1] [2]. Several researchers from around the world and in the country have looked into using probability distributions to analyze temperature data [3] [4] [5] [6] [7]. In this paper, our study focussed on daily temperatures in Pekanbaru city for the years 1990-2020. The purpose of this study is to quantify and describe the behaviour of daily temperature in Pekanbaru city, Indonesia. In this study, five models distribution were tested, they are Gamma (GM), Slashed Quasi-Gamma (SQG), Three Parameters Quasi Gamma (TPQG), Two Parameter Gamma-Exponential (TPGE), and Modified Log-Logistic (MLL) distributions. When fitting a distribution model, assumed that the daily temperature follows the form of a statistical distribution. Identification of a suitable distribution is an important step in selecting best model. Comparison of the proposed distributions with existing distribution functions is done to demonstrate their suitability in describing temperature characteristics. Unknown parameter estimations were obtained with the Maximum Likelihood Method. Graphical methods and Numerical criteria such as AIC, BIC, and - l were used to obtain the distribution which provides the best fit the temperature data. In the following section, the distributions to be used in modelling the temperature data are included.

II. DATA SET AND STUDY AREA

Data used in this work were downloaded from the website of the Federal Office of Meteorology and Climatology of Indonesia (BMKG). The original data consisted of daily temperature records from 2010 to 2020, which were provided by the BMKG of Pekanbaru city, Indonesia. The data and the histogram or characteristic wind speed are presented in Figure 1. Pekanbaru City located in Riau Province, with a tropical climate that has temperature that vary from 74 F^0 to 89 F^0. Initial information on Temperature in the city of Pekanbaru can be seen in the descriptive statistics are presented in Table 1. The variations of temperature data that are not so large (4.58 F^0) indicate that this temperature is quite stable and very good as a source of electrical energy for household purposes

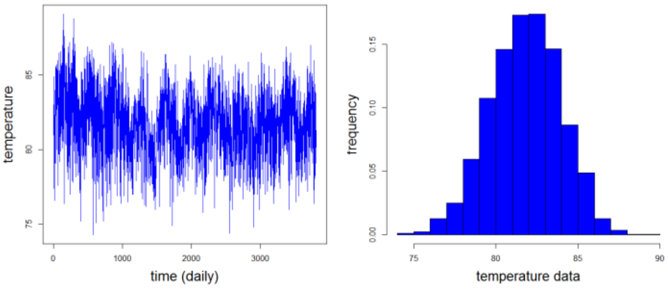


Fig. 1. Plot and histogram daily temperature data on Pekanbaru respectively

Table 1: The descriptive statistics for daily temperature data (F)

Statistics	***Velue***
Mean	81.85
Varians	4.58
Minimum	74.30
Maximum	89.10

III. METHODOLOGY

In this study, five models distribution were tested, they are GM, SQG, TPQG, TPGE, and MLL distributions. When fitting a distribution model, assumed that the daily temperature follows the form of a statistical distribution. Identification of a suitable distribution is an important step in selecting best model. The probability density function and log likelihood function of the distributions are given in table 2 and table 3.

Table 2. Probability Density Function (pdf) of some distributions

The Distributions	**Probability density function (pdf)**
GM	$f(y)=\frac{1}{\beta^{\alpha}\Gamma(\alpha)}y^{\alpha-1}exp\left(-\frac{y}{\beta}\right),$ $y>0,\ \alpha,\beta>0$
SQG [8]	$f(y)=\frac{q\beta^{q}y^{-(q+1)}}{\Gamma\left(\frac{1}{10}\right)}\Gamma\left(\frac{q}{\theta}+\frac{1}{10}\right)F\left(\left(\frac{y}{\beta}\right)^{\theta},\frac{q}{\theta}+\frac{1}{10},1\right),$ $y>0,\quad q,\beta,\theta>0$
TPQG [9]	$f(y)=\frac{2k\theta^{\alpha}y^{2k\alpha-1}exp(-\theta y^{2k})}{\Gamma(\alpha)},$ $y>0,\quad k,\theta,\alpha>0$
TPGE [10]	$f(y)=\frac{k^{2}}{(1+k)}\left(1+\frac{k^{s-2}}{\Gamma(s)}y^{s-1}\right)exp(-ky),$ $y>0,\ \ k,s>0$
MLL [11]	$f(y)=\frac{a(ay)^{b-1}exp(\theta y)(b+\theta y)}{\left(1+(ay)^{b}exp(\theta y)\right)^{2}},$ $y>0,\ \ a,b,\theta>0$

Goodness-of-Fit Tests (GOF)

The goodness-of-fit test is applied to determine the accuracy of the predicted values using the theoretical probability model. It aids in the selection of the best model from among the available distributions. For selecting the best probability distribution, several goodness-of-fit tests were applied, including the graphical methods (density plot) and numerical methods (AIC, BIC). The formula of numerical methods such as AIC and BIC are exhibited in the following Table 4

Table 3. The Log Likelihood of some distributions

The Distributions	**Log Likelihood**
Gamma (GM)	$l(\alpha,\beta)=\alpha-1\sum_{i=1}^{n}log(y_i)-\frac{1}{\beta}\sum_{i=1}^{n}y_i-n\alpha\ log(\beta)-n\ log(\Gamma(\alpha))$
Slashed Quasi-Gamma (SQG) [8]	$l(q,\beta,\theta)=c(q,\beta,\theta)-(q+1)\sum_{i=1}^{n}log(y_i)+\sum_{i=1}^{n}F\left(\left(\frac{y}{\beta}\right)^{\theta},\frac{q}{\theta}+\frac{1}{10},1\right)$ where $c(q,\beta,\theta)=n\ log(q)+nq\ log(\beta)-n\ log\left(\Gamma\left(\frac{1}{10}\right)\right)+n\ log\left(\Gamma\left(\frac{q}{\theta}+\frac{1}{10}\right)\right)$ and $F(y,\alpha,\beta)=\int_{0}^{y}\frac{\beta^{\alpha}}{\Gamma(\alpha)}y^{\alpha-1}e^{-y\beta}\,dy$

Three Parameters Quasi Gamma (TPQG)[9]	$l(k,\theta,\alpha) = n\,log(2) + n\,log(k) + n\alpha\,log(\theta) + (2k\alpha - 1)\sum_{i=1}^{n} log(y_i) - \theta\sum_{i=1}^{n} y_i^{2k} - n\,log(\Gamma(\alpha))$
Two Parameter Gamma-Exponential (TPGE) [10]	$l(k,s) = 2n\,log(k) - n\,log(1+k) + \sum_{i=1}^{n} log(\Gamma(s) + k^{s-2}y^{s-1}) - n\,log(\Gamma(s)) - k\sum_{i=1}^{n} y_i$
Modified Log-Logistic (MLL) [11]	$l(a,b,\theta) = nb\,log(a) + (b-1)\sum_{i=1}^{n} log\,(y_i) + \theta\sum_{i=1}^{n} y_i + \sum_{i=1}^{n} log(b+y_i) - 2\sum_{i=1}^{n} log(1 + (ay)^b exp(\theta y))$

VI. RESULT AND DISCUSSION

In this section, we analyze a daily temperature data set to demonstrate the performance of the GM, SQG, TPQG, TPEG and MLL distributions in practice. The fitting of that distributions was considered using data from the period between 2010 and 2020. Computed parameter values of different probability density function used for the BMKG Pekanbaru stations are presented in Table 5.

Table 4. The formulas of numerical criteria for model evaluation

Numerical Criteria	Formula
AIC	$-2l + 2q$
BIC	$-2l + q\log(n)$

l = log likelihood, q = Number of parameters, n = number of data

Table 5. Computed parameter values of different probability density functions

	GM	SQG	TPQG	TPGE	MLL
α	0.06	-	-	-	-
β	1320.99	-	-	-	-
q	-	55.61	-	-	-
β	-	82.37	-	-	-
θ	-	419.77	-	-	-
k	-	-	1.73	-	-
θ	-	-	1.18×10^{-6}	-	-
α	-	-	4.93	-	-
k	-	-	-	0.90	-
s	-	-	-	74.27	-
a	-	-	-	-	1.05×10^{-8}
b	-	-	-	-	4.42

On the graphical presentation of modelling of the daily temperature data, the density function curve for GM, SQG, TPQG, TPEG and MLL distributions models are seen in Figure 2, 3, 4, 5 and 6 are respectively. When the density functions (pdf) are examined, it was determined that some distributions have yield similar results. From this figure GM and MLL distributions model is able to provide a good result for daily temperature data. However, instead of graphical evaluation, Table 6 provides a more meaningful comparison using AIC, BIC, and $-l$ values.

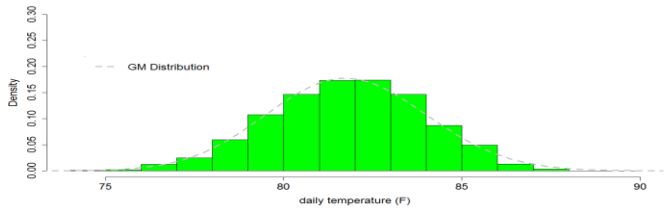


Fig. 2 Fitted pdf plots of GM Distribution.

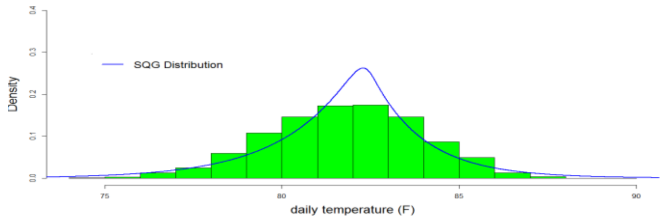


Fig. 3 Fitted pdf plots of SQG distribution

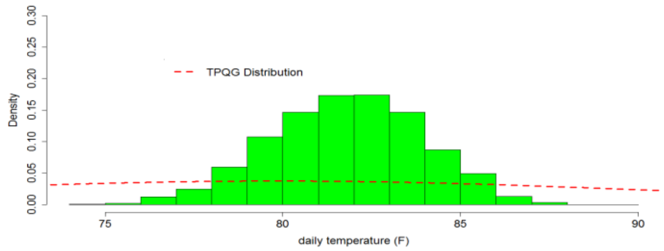


Fig. 4 Fitted pdf plots of SQG distribution

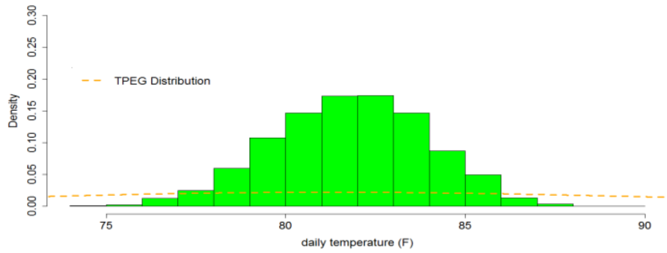


Fig 5. Fitted pdf plots of TPEG distribution

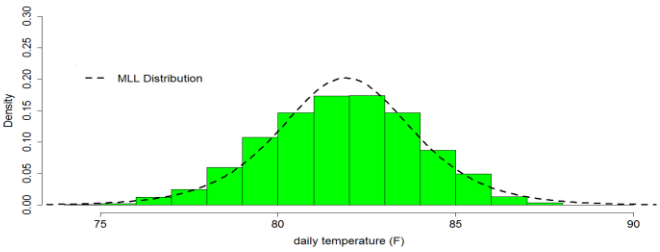


Fig 6. Fitted pdf plots of MLL distribution

Table 6 includes AIC, BIC and - *l* values test statistics values of the goodness of fit test for the fitness of daily temperature data based on Maximum Likelihood Estimators for GM, SQG, TPQG, TPEG, and MML distributions. According to these results, the lowest AIC, BIC and - *l* value are obtained for GM and MML distributions. As a conclusion, it is seen that GM and MML distribution provides better modelling in terms of Numerical criteria

Table 4. AIC, BIC, and – Log Likelihood (*l*) function values

	GM	SQG	TPQG	TPEG	MLL
AIC	**16674.87**	17049.30	25351.19	29314.35	**16754.16**
BIC	**16687.36**	17068.04	25369.94	29335.09	**16772.90**
- *l*	**8335.435**	8521.65	12672.6	14655.18	**8374.079**

V. CONCLUSSION

In the present article, a comparison of distribution models has been undertaken for describing temperature of Pekanbaru stations. Common conventional density GM and extended pdf along with four proposed density function such as SQG, TPQG, TPEG and MLL are used for temperature modeling. It is shown that extended density such as SQG, TPQG and TPEG is inadequate, hence, conventional density functions GM and the extended density MLL are used to model the observed temperature distributions better. Results show clearly that proposed GM and MLL provide viable alternative to other pdf in describing temperature regimes.

REFERENCES

1. K. E. Trenbert and D. J. Shea, Relationships between precipitation and surface temperature, Res. Let., 32, L14703, doi:10.1029/2005GL022760, (2005).
2. V. V. Kharin, F. W. Zwiers, X. Zhang, and M. Wehner, Change in temperature and precipitation extremes in the CMIP5 ensemble, Journal of Climate, 119 (2013), 345-357.
3. Mota, F. S., Rosskoff, J. L. C., & Silva, J. B. (1999). Probabilidade de ocorrência de dias com temperaturas iguais ou superiores a 35 °C no florescimento de arroz no Rio Grande do Sul. Revista Brasileira de Agrometeorologia, 7(2), 147-149.
4. Assis, J. P., Neto, D. D., Manfron, P. A., Martin, T. N., Sparovek, G., & Tim M. T. C. (2004). Ajuste de series históricas de temperatura e radiação solar global diária às funções densidade de probabilidade normal e log-normal, em Piracicaba, SP. Revista Brasileira de Agrometeorologia, 12(1), 113-121.
5. Prela-Pantano, A., Rolin, G. S., & Camargo, M. B. P. (2009). Probabilidade de ocorrência de temperaturas mínimas menores que 5 ºC na região do médio Paranapanema. Bragantia, 68(1), 279-284. https://doi.org/ 10.1590/S0006-87052009000100030
6. Araújo, E. M., Silva, I. N., Oliveira J. B., Cavalcante Junior E. G., & Almeida, B. M. (2010a). Aplicação de seis distribuições de probabilidade a séries de temperatura máxima em Iguatu-Ce. Revista Ciências Agronômica, 41(1), 36-45.
7. Araujo, E. M., Oliveira, J. B. M. G. S., Viana, P. C., & Alves, A. S. (2010b). Análise da aderência de distribuições de probabilidade aos dados de temperatura máxima e mínima do ar em Iguatu-Ce. Revista Caatinga, 23(3), 104-109.
8. Yuri A. Iriarte, Héctor Varela 1 , Héctor J. Gómez and Héctor W. Gómez, A Gamma-Type Distribution with Applications, symmetri, 2020, 12, 1-15, doi:10.3390/sym1205087
9. Qaisar Rashid, Hafiz Muhammad Yaseen, Muhammad Uzair and Muhammad Tariq Jamshaid. Three Parameters Quasi Gamma Distribution and with Properties and Applications, Pak.j.stat.oper.res, 2022, 18(3), 759-773, DOI: http://dx.doi.org/10.18187/pjsor.v18i3.3759
10. Saurav Sarma, Inzamul Ahmed and Anjana Begum, A New Two Parameter Gamma-Exponential Mixture, 2021, Math. Comput. Sci. 11(1), 414-426 https://doi.org/10.28919/jmcs/5146
11. Mohamed Kayid, Applications of Bladder Cancer Data Using a Modified Log-Logistic Model, Applied Bionics and Biomechanics, 2022, 1-12, https://doi.org/10.1155/2022/6600278