International Journal of Mathematics Trends and Technology
ISSN: 2231 – 5373 /doi:10.14445/22315373/IJMTT-V67I5P508
Volume 67 Issue 5, 82-85, May, 2021

The Analisys of Daily Temparature in Pekanbaru City Using Weibull, Gamma, and Nakagami Distribution

Martha Sri Pramadani[1], Arisman Adnan[1], Rado Yendra[2]

[1] *Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Riau, Pekanbaru, 28293, Indonesia*
[2] *Department of Mathematics, Faculty of Science and Technology, UIN Sultan Syarif Kasim Riau, Pekanbaru, 28293, Indonesia*

Abstract — *Changes in temperature are one of the consequences of climate change. Ecosystems and various sectors of human activity are sensitive to high and low temperature, especially if it occurs over a long time. Pekanbaru city is a city in Riau Province, Indonesia, which has tropical climates while daily temperatures varying from 72 F-97 F. This study was focussed on reducing the effects of high temperatures, such as global changes through the behavior of daily temperature data. Main goal of this study is to find the best fitting distribution to the daily temperature measured for the years 1990-2020. The Weibull (W) , Gamma (G) and the Nakagami (N) distributions are fitted to data corresponding to the methods to describe the daily of temperature. Graphical tests of probability density function (pdf), cumulative distribution function (cdf), quantile function (qf) and numerical criteria of Relative Root Mean Square Error (RRMSE) and relative absolute square error (RASE) were used to select the best fit model. In most cases, graphical examine have the same result between G and N distribution but their RRMSE and RASE result different. Finally, we found that the N distribution is the most suitable distribution for modeling the daily temperature of Pekanbaru city.*

Keywords — *climate change, best fitting distribution, Weibull distribution, gamma distribution, Nakagami distribution*

I. INTRODUCTION

One of the major components of climate change is temperature. Temperature have been analyzed extensively over the past two decades [7] [13]. In the presence of hottest days, there is an increase in electricity usage as a result of people switching on all the cooling systems [1]. Therefore, modeling the temperature in the energy field is very important. Increasing temperatures also have an impact on civil construction. The extreme heat puts roads and buildings at risk of collapsing which can danger lives of people [2]. In the agricultural sector, extreme heat has resulted in droughts that in starvation and perish of livestock and health problems [14]. To maintain the risk of a heat wave disaster, accurate and precise statistical modeling of daily temperatures is necessary. Regression analysis, state space, time series, and Kalman filter methods are common used to maintain the risk of a heat wave disaster, accurate and precise statistical modeling of daily temperatures. The techniques are classical statistical techniques were frequently applied in the energy and meteorological fields [5]. These techniques just concentrate only on mean instead the tail of distribution. These results inaccurate estimation and difficulty in estimating appropriate model parameters [6]. The problems due to the use of inappropriate statistical techniques can be overcome by distribution modeling theory. This is because distribution modeling is able to modeling characterization of distribution.

Several researchers have discussed the modeling of probability distribution in climate change, like temperature, precipitation, and wind speed [4], [12], [8], and [10]. Generalized Extreme Value (GEV) distributions are mostly evaluated in the general circulation model (CGM), regional climate models (RCMs), and investigate time series of extreme climate changes. Jeruskova et. al [3] studied the use of the GEV distribution to analyze the annual maximum and minimum temperatures for several European cities by applying the extreme value theory, hypothesis testing, and change point method. Besides the GEV distribution, the Frechet is also known as the inverse Weibull distribution and generalized logistics distribution is most often used in modeling extreme temperature events. 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 particular, we aim to model the daily temperatures by using the Weibull (W), gamma (G), and Nakagami (N) distribution. Then W, G, and N distributions are estimated using Maximum Likelihood.

II. MATERIAL AND METHODS

A. Materials

Pekanbaru is the capital city of Riau and was located among 101°14'- 101°34' East Longitude and 0°25'- 0°45' North Latitude with an height from sea level ranging from 5-50 meters. In general, Pekanbaru has a tropical climate with maximum temperatures ranging from 34.1°C-35.6°C and minimum temperatures between 20.2°C-23.0°C [11]. Original data is in the form of daily temperature records from 1990 to 2020 obtained from NOAA [9]. The number of observations are 10134. The plot and histogram of the data set are presented in Figure 1.

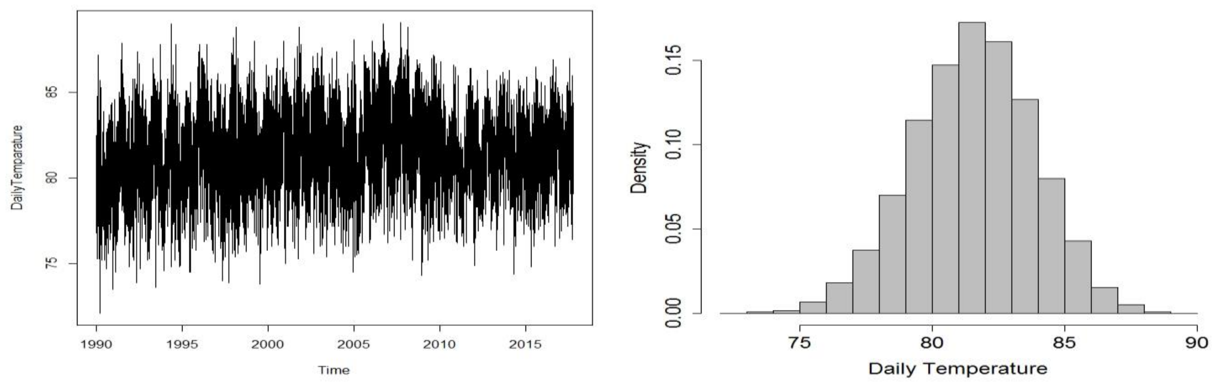

Figure 1. Plot and Histogram Daily Temperature Data on Pekanbaru city

The following table describes the mean, minimum, maximum, variance, and standard deviation of the daily temperature in Pekanbaru city for the years 1999-2020.

Tabel 1. Statistical Description of Daily Temparature

Mean	Minimum	Maximum	Variance	Standard Deviation
81.6	72.1	89.1	5.26	2.29

B. Model Selection for the Daily Temparature

In this study, three models distribution were tested, they are W, G, and N 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, likelihood function and log likelihood function of the distributions are defined as follows:

1. ***Gamma Distribution:***

$$f(x;\gamma,\lambda)=\frac{\lambda(\lambda x)^{\gamma-1}e^{-\lambda x}}{\Gamma(\gamma)},\ \ x\geq 0, \gamma>0 \text{ and } \lambda>0$$

$$F(x;\gamma,\lambda)=\frac{1}{\Gamma(\gamma)}\Upsilon(\gamma,\lambda x)$$

$$L(\gamma,\lambda|x_i)=\frac{\lambda^{n\gamma}}{\Gamma(\gamma)^n}\prod_{i=1}^{n}{x_i}^{\gamma-1}\exp\left(-\sum_{i=1}^{n}\lambda x_i\right)$$

$$\ell(\gamma,\lambda|x_i)=n\gamma\ln(\lambda)-n\ln\Gamma(\gamma)+(\gamma-1)\sum_{i=1}^{n}\ln x_i-\sum_{i=1}^{n}\lambda x_i$$

2. ***Weibull Distribution:***

$$f(x;\alpha,\beta)=\frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1}e^{-\left(\frac{x}{\alpha}\right)^{\beta}},\ \ x\geq 0, \alpha>0 \text{ and } \beta>0$$

$$F(x;\alpha,\beta)=1-e^{-\left(\frac{x}{\alpha}\right)^{\beta}}$$

$$L(\alpha,\beta|x_i)=\frac{\beta^n}{\alpha^{n\beta}}\prod_{i=1}^{n}{x_i}^{\beta-1}\exp\left(-\sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta}\right)$$

$$\ell(\alpha,\beta|x_i)=n\ln(\beta)-n\beta\ln(\alpha)+(\beta-1)\sum_{i=1}^{n}\ln x_i-\sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta}$$

3. ***Nakagami Distribution:***

$$f(x;m,\Omega)=\frac{2m^m}{\Gamma(m)\Omega^m}x^{2m-1}e^{\left(-\frac{m}{\Omega}x^2\right)},\ \ \ x\geq 0, m\geq 0.5 \text{ and } \Omega>0$$

$$F(x;m,\Omega) = 1 - \frac{P\left(m, \frac{mx^2}{\Omega}\right)}{\Gamma(m)}$$

$$L(m,\Omega|x_i) = \left(\frac{2}{\Gamma(m)}\right)^n \left(\frac{m}{\Omega}\right)^{nm} \prod_{i=1}^{n} {x_i}^{2m-1} exp\left(-\frac{m}{\Omega}\sum_{i=1}^{n} {x_i}^2\right)$$

$$\ell(m,\Omega|x_i) = n\ln(2) - n\ln(m) + nm\ln(m) - nm\ln(\Omega) + (2m-1)\sum_{i=1}^{n} \ln x_i - \frac{m}{\Omega}\sum_{i=1}^{n} {x_i}^2$$

The selection of the best distribution for modeling daily temperature data in Pekanbaru city can be determined using the goodness of fit test, that is RRMSE and RASE. This test involves on the difference between the expected value and the value observed in the distribution. The tests are given as follows:

$$RRMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_{i:n} - \hat{Q}(F_i)}{x_{i:n}}\right)^2}$$

$$RASE = \sum_{i=1}^{n}\left|\frac{x_{i:n} - \hat{Q}(F_i)}{x_{i:n}}\right|$$

where $x_{i:n}$ is observed values for ith order statistics of random sample of size n, $\bar{Q}(F_i) = \frac{1}{n}\sum_{i=1}^{n} \hat{Q}\,(F_i)$ is the estimated quantile values associated with Gringorton plotting position F_i. The distribution with the smallest value of RRMSE and RASE will chosen as the best fitting model for the data.

III. ANALYSIS AND RESULTS

Estimation of distribution parameters is the main thing that must be done to determine the characteristics of the distribution. MLE was used to estimate parameter the W, G, and N distribution. Parameter estimate of daily temperature distribution (W, G, N) and goodness of fit test (RRMSE and RASE) are given in the Table 2. Based on the smallest RRMSE and RASE values, it is shown that N is the best distribution in modeling daily temperature data in Pekanbaru city.

Tabel 2. Parameters and Goodness of Fit

Distributions	Parameters	RRMSE	RASE
Weibull	γ (shape) = 82.71 λ (scale) = 37.66	1.04×10^{-14}	5.38×10^{-14}
Gamma	α (shape) = 0.27 β (scale) = 299.71	4.61×10^{-17}	1.24×10^{-13}
Nakagami	m (shape) = 315.93 Ω (scale) = 6664.45	3.11×10^{-17}	2.09×10^{-12}

Figure 2 was presented the histogram for the data with the fitted densities and we can observe that there are a typical observations and the right tail of the Nakagami model is slightly heavier and the same figure presents the empirical and fitted cumulative distribution functions and we can see that the fitted *cdf* of the Nakagami model is closer to the empirical distribution.

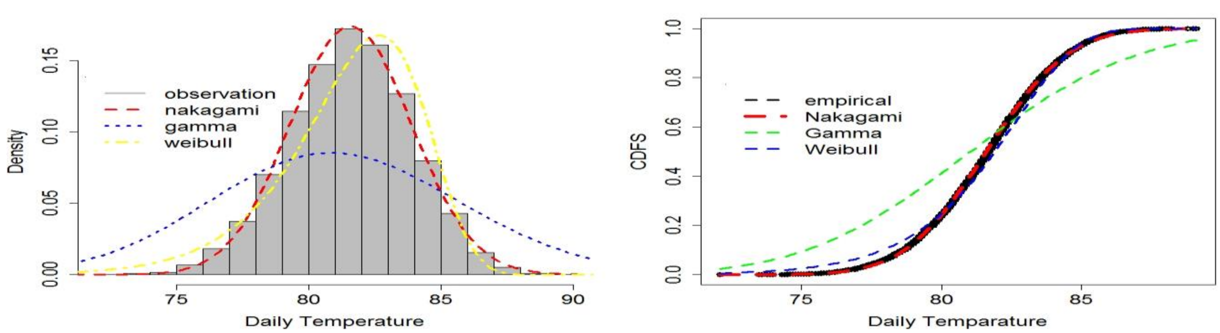

Figure 2. The Comparison Empirical and fitted using Density and CDFs of W, G, and N Respectively

To examine the degree of suitability of the model visually, a quantile-quantile plot were used. From visual point, there was differences in choosing a distribution to describe the data used in study. For example, the daily temperature in Pekanbaru city displayed in Figure 3. The Q-Q plot has shown that the Nakagami distribution is fitted better because most of the data are on a straight line compared to the gamma and Weibull distribution. In many cases, finding the best model from the visual display of the Q-Q plot is very difficult. Therefore, a statistical test is needed to determine the best distribution for modeling data. One of the statistical tests is the RRMSE and RASE tests.

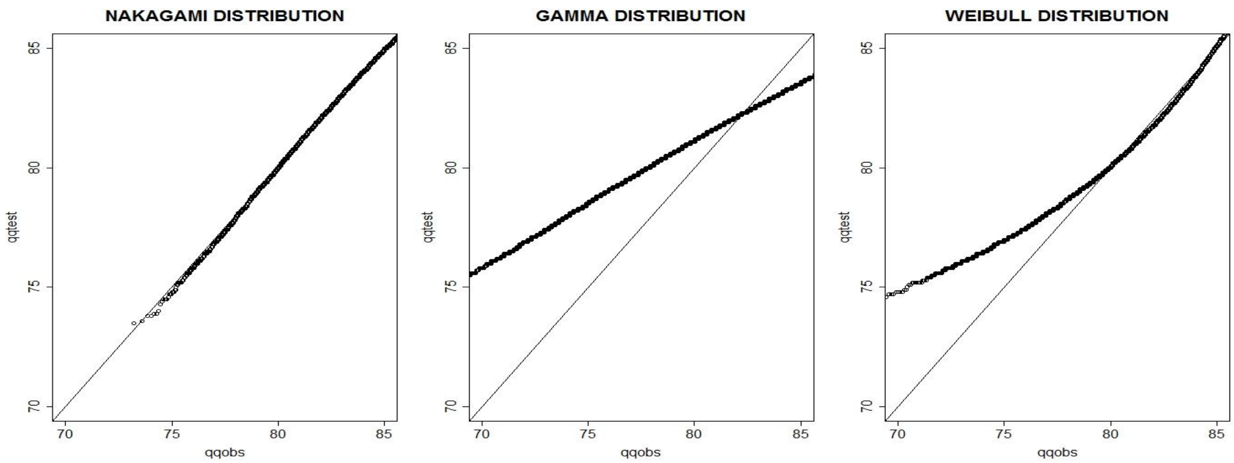

Figure 3: Q-Q plots for the fitted daily temperature distributions

IV. CONCLUSIONS

In this paper, the occurrence probability of the daily temperature events was analyzed at the station, in Pekanbaru City. Three of the two parameter probability distributions, the W, G, and N distributions are used to modeling daily temperature data in Pekanbaru city. The parameter estimation of the W, G, and N distributions was carried out using Maksimum Likelihood Method (MLE). We can conclude that, based on the graphical methods (PDF, CDF, QF) and numerical methods (RRMSE and RASE), the Nakagami distribution better than the gamma and Weibull distribution for daily temperature data in Pekanbaru city for the years 1990-2020. Hence, we recommend to implement the Nakagami distribution to model daily temperature.

REFERENCES

[1] A. Munoz, E. F. Snchez-beda, and J. Marn, Short-term forecasting in power system: A guide tour, in Hanbook of power system II, Springer-Verlag, Berlin Hedelberg, (2010), 129-160.

[2] B. Lyon, Southern Africa summer drought and heat waves: Observations and coupled model behavior, Journal of Climate, 22(22) (2009), 6033-6046.

[3] D. Jaruskova dan M. Rencova, Analysis of annual maximal and minimal temperatures for some Eropean cities by change point methods, Envirometrics, 19 (2008), 221-233.

[4] F. W. zwiers dan V. V. Kharin, Change in extremes of the climate simulated by CCC under CO2 dobling, Journal of Climate, 11 (1997), 2200-2222. High resolution fiber distributed measurements with coherent OFDR, in Proc. ECOC'00, paper 11(3)(2000) 109-125.

[5] H. Hahn, S. Meyer-Nieberg, and S. Pickl, Electrical load forecasting methods: Tools for decision making, European Journal of Operational Research, 3 (2009), 902-907.

[6] H. N. Bystrom, Extreme value theory and extremely large electricity price changes, International Review of Economics and Finance, 1 (2005), 41-55.

[7] K. E. Trenbert and D. J. Shea, Relationships between precipitation and surface temperature, Res. Let., 32, L14703, doi:10.1029/2005GL022760, (2005).

[8] K. Pangaluru, I. Velicogna, T. C. Sutterley, Y. Mohajerani, E. Ciraci, J. Sumpolli, and S. V. B. Rao, Estimating change of temperatures and precipitation extremes in India using the Generalized Extreme Value (GEV) distribution, Hydrology and Earth System Sciences, (2018), 1-33.

[9] NOAA, National Centers for Environmental Information, https://www.ncei.noaa.gov accessed 22 Februari 2019.

[10] P. Guttorp and J. Xu, Climate change, trends in extremes, and model assessment for a long temperature time series from Sweden, Environmetrics, (2009), 456-463.

[11] Pekanbarau. Go. Id, Wilayah Geografis, 13 Februari 2021, https://www.pekanbaru.go.id/index.php/p/menu/profil-kota/wilayah-geografis accessed 18 Maret 2021.

[12] V. V. Kharin dan F. W. Zwiers, Estimating extremes in transient climate change simulations, Journal of Climate, 18 (2004), 1156-1173.

[13] 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.

[14] W. Steffen, L. Hughes, and S. Parkins, Heat waves: Hotter, longer, more often, Climate Council of Australia Limited, Second Major Technical Report of the Climate Council, Sydney, (2009).