The Probability Distribution of Maximum Temperature to Assess the Suitable Statistical Models: Take the North-East and Southern Regions of Pakistan

Tasir khan (tasirkhan.state@gmail.com)
Lanzhou University https://orcid.org/0000-0003-0651-0418

Yejuan Wang
Lanzhou University

Research Article

Keywords: Probability distribution function, Kumaraswamy distribution, Statistical analysis, temperature data, Maximum likelihood estimator

Posted Date: July 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-568429/v1

The probability distribution of maximum temperature to assess the suitable statistical models: Take the north-east and southern regions of Pakistan

Tasir Khan[a], , Yejuan Wang[a *]

[a] School of Mathematics and Statistics, Gansu Key Laboratory of Applied Mathematics and Complex Systems, Lanzhou University, Lanzhou 730000, China

*Corresponding author Email:

Dr. Tasir khan

School of Mathematics and Statistics

Lanzhou University, Lanzhou 730000, China

tasirkhan.state@gmail.com

The probability distribution of maximum temperature to assess the suitable statistical models: Take the north-east and southern regions of Pakistan

Abstract

Precise maximum temperature probability distribution information is indeed of accurately significance for numerous temperature uses. The purpose of this research to assess the appropriateness of these functions likelihood for evaluating the temperature models at different sites in southern part of Pakistan. The Kumaraswamy distribution function is used initially to approximation the models of maximum temperature. Compare the presentation of the Kumaraswamy distribution with twelve commonly used the probability functions. The consequences obtained show that the more effective functions are not similar across all sites. The maximum temperature features, quality and quantity of the noted temperature observation can be regarded as a factors that affect the presentation of the function. Similarly, the skewness of the noted maximum temperature observations may affect the precision of Kumaraswamy distribution. For the Hyderabad, Lahore and Sialkot sites, the Kumaraswamy distribution obtainable the topmost presentation, however for the Karachi, Multan stations, the generalized extreme value (GEV) distributions provided the best fit, respectively. According to the calculations, the Kumaraswamy distribution usually be regarded as a valid distribution because it runs 3 best fit sites and ranks 2 to 3 among the remaining sites. Though, the tight presentation of the Kumaraswamy and GEV and the flexibility of the Weibull distribution which has been usually verified, more evaluations of the presentation of the Kumaraswamy distribution are needed.

Key words; Probability distribution function; Kumaraswamy distribution; Statistical analysis; temperature data; Maximum likelihood estimator

Introduction

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 (Karpouzos et al., 2010; Cui et al., 2017) (Huang et al., 2019). In case of weather variation or global temperature, small changes in the usual temperature value can cause major

changes in the scale and frequency of extreme measures, with high temperatures (Yan et al., 2016).According to the fifth valuation report of the Intergovernmental Panel on Weather Change, from 1880 to 2012, the global average surface temperature increased by 0.85°C. There is no doubt that global warming has had a serious influence on human society(Chen et al., 2019, Zhou et al., 2017). As the temperature rises, the ability to retain water in the atmosphere increases, and the temperature of extreme value has changed significantly, leading to forest fires and frequent droughts. At the same time, with temperature changes, the global usual ocean level rise by 19 cm between 1900 and 2011. In other words, changes in extreme temperatures increase the intensity and frequency of extreme weather events (such as extreme temperatures, heavy rains, droughts, and floods. When the temperature rises, the desert will increase, and the infiltration and soil moisture intensity will change. The change in water rotation is caused by an increase in temperature. Rising temperature will also change the restructuring of river excess and the characteristics of water resources in the basin. Extreme temperature events is the main measure of extreme climate events. Therefore, the study of extreme temperature changes under global warming is great significance.

A current analysis of the detection of extreme temperature trends in Europe (Irannezhad et al., 2019) confirmed that 20th century over time, the warmth of day and night. Also concluded that the frequency and intensity of high temperature (low temperature) (Shaby and Reich, 2012, Parey et al., 2013, Naveau et al., 2014, Huang et al., 2016).Since the temperature probability, it is a feasible method to predict extreme precipitation based on the relationship between temperature and extreme precipitation under weather change. Therefore, in all over the world is studying the basic process of detecting extreme precipitation changes with temperature (Donat et al., 2016, Gao et al., 2018).(Wang et al., 2017) found that, compared with the historical temperature dependency of extreme precipitation, the peak temperature of extreme precipitation will rise with climate warming, which means that the peak arrangement does not mean the probable higher limit of extreme precipitation in the future. As mentioned above, the relationship between extreme precipitations and has been fully studied, but the association among highest extreme rainfall and highest temperature (Teshome and Zhang, 2019).

All selection periods fit the GEV distribution and estimate the parameter. The likelihood ratio test shows that the best model in which the position parameter increases linearly, and the parameters shape and scale are constant. Model diagnosis including quantile plots, probability

plots and density plot showed a good degree of fit. The GOF test Anderson Darling and Kolmogorov Simonov show that modeling can almost gave same fitting result (Hasan et al., 2012, Hughes et al., 2007). A univariate extreme value (EV) model based on the limit temperature data of the block maximum method. The block size selection is important because when size is large, valuable information may be wasted. Using a block length of one year will generate too few maximum sequences (20 data) and result in a higher estimate variance. However, a block length that is too short will not meet the limit approximation of extreme value temperature (EVT). Therefore, the length is half-yearly, quarterly and monthly blocks are still feasible. From the analysis of quantile graph, monthly and annual blocks are more suitable than quarterly and semi-annual blocks. Generally, LM is better than MLE in estimating parameters. Numerical results indicate that the maximum temperature will gradually increase (Amin et al., 2018).

The absolute burden of heat waves that is not conducive to health. This is in the entire community, but workers who work in various hot places are particularly vulnerable. Therefore, the impact is also economical. Since this growing hazard, the health authorities of the Republic of Djibouti would play an important role conducting research on the actual impact of high temperatures on morbidity and mortality, and promoting, leading and evaluating a series (Ozer and Mahamoud, 2013). During the confirmation period statistical downscaling model (SDSM) obtained better results from monthly seasonal and daily time sequence. In other words, the results based on the seasonal sequence are slightly better than the monthly sequence. The performance of this model is suitable for SDSM when trying to simulate future extreme temperature indexes. In the verification process, the intensity limit index is better than the frequency index (Mahmood and Babel, 2014). This asymmetry is easily clarified by the irregularity of seasonal temperature distribution. Under simulated warmer weather situations, the temperature change (standard deviation of the distribution) in summer is relatively unchanged, but in winter (especially in higher latitudes) it is significantly reduced (Holmes et al., 2016). The extreme temperature increase occurred in the last ten years. The influence of temperature is the cause of changing Pakistan's climate. For example, heat waves are increasing across the country (Afzaal et al., 2009). Accurately estimating the long-term trends of global and regional climate change are essential for the impact and prediction attributable to climate change (Li et al., 2020, Li et al.).

The objective of this study due to the effective application of the Kumaraswamy distribution in different fields, it may be interesting to evaluate its proficiency in the designated

case study to estimate the maximum temperature distribution. Thus, the presentation of the Kumaraswamy distribution was tested for certain before used distribution functions (including exponential, normal, invers-Gaussian, logistic, log-logistic, log-normal 3, Gumbel Generalize Extreme Value, Weibull 3, Pearson type 3, and Generalize-Gamma distribution. Statistical evaluation of the efficiency of all twelve distribution functions based on broadly used statistical parameters. To determine the most suitable theoretical function by using the In addition, KS, AD, AIC and BIC are used as goodness of fit indicators.

Methodology

Probability distribution modeling of maximum temperature

Generally, understanding the probability distribution of maximum temperature is essential for characterizing temperature behavior, evaluating maximum temperature performance. Therefore, it is important to determine the most suitable function for temperature data. In this study, twelve PDFs were used to describe the frequency distribution of temperature. These distribution functions are exponential, N, IG,L, LL3, LN3, GUM, KUM, GEV, WEI3, PE3, GG Weibull, gamma, lognormal, Gaussian inverse, logarithmic, generalized extreme and upper middle. In this study, the Kum distribution is used for the first time to describe the maximum temperature distribution in the selected case study. MLE is used to estimate the parameters that define each PDF. MLE is generally regarded as a very reliable technique, which can determine the parameter value that maximizes the likelihood of the data used [41]. For more sample sizes, MLE is most efficient than estimate approaches such as the Linear moments (LMOM), method of moments (MOM) and produces a lesser mean square error.

Table 1

Appraising the fitting performance

The special probability distribution is affected by a number of aspects, therefore the PD parameter evaluation methods, comparison methods and the availability of rainfall data. In this research, chi square and KS test was used to assess the fitness of certain PDs. The KS and chi square test calculated test statistics that define the theoretical value and actual value estimate from distributions. In addition, in order to check the visual estimation of the goodness of fit. The advantage of the above test plot and fit test applied to the maximum rainfall data (Morgan et al.,

2009, Lollchund et al., 2014, Cassalho et al., 2018, Fawad et al., 2019). Weather test statistics and fitted test described above can consistently use to select the best fit distribution.

In order to evaluate the effectiveness checked PDF for modeling the probability distribution of temperature, statistical displays of AD, KS, Bayes Information Criterion (BIC) and Akaike Information Criterion (AIC) are used . Statistically used to examine the deviation among the forecast data and the observed data using a probability function. These indicators are briefly introduced here [43]

The KS statistics is usually based on the empirical distribution function (EDF), and the sample comes from continuous distribution. A random sample assumes that there are $x_1, x_2, x_3, \dots, x_n$ in a certain distribution. Then you can define EDF in the following ways

$$F_n(x) = \frac{1}{n}[number\ of\ observation\ \leq x] \quad (1)$$

The KS test the theoretical probability distribution as

$$D = max_{1\leq i\leq n,}\left[F(x_i) - \frac{i-1}{n}, \frac{i}{n} - F(x_i)\right] \quad (2)$$

In equation (2), $F(x_i)$ is the cumulative distribution function, x_i is the ith order statistic and n denote the sample size.

The Anderson Darling (A-D) test associates the observed fitting of the distribution. The Anderson darling test assigns a higher weighted distribution to the tail, (Fawad et al., 2019) . The A-D statistics A^2 test as:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\ln F(x_i) + ln\left(1 - F(X_{n-i+1})\right)\right] \quad (3)$$

Where A^2 relates to the test result, X is the variable, $F(X_i)$ is the distribution function, and n is the sample size and uses the statistical model with the smallest AD score for the data of wind speed as the most-fitting distribution model.

2. Study Area

Pakistan's climate has many unique cyclical changes. The most serious variables affecting the climate are humidity, temperature, wind speed and rainfall. The deserts in some areas are still very hot and dry.

Karachi is the most populous city in Pakistan and the capital of Sindh Province. Karachi is located in southern Pakistan in Sindh Province. Therefore, the summer is not hot and humid from December to February. Compared with the hot season that started in March and lasted until the June monsoon, it was dry and pleasant. Hyderabad is also located in Sindh province, the desert

climate in Hyderabad is hot and warm throughout the year. This city is famous for its tempering the originally hot climate. As a result, houses in Hyderabad have traditionally been equipped with "induction wind" towers that blow breeze in to residential areas to reduce heat. From mid-April to late June is the hottest period of the year, with the highest peak in May at 41.4^{o} C. The maximum temperature recorded on May was 50°C, and the lowermost temperature recorded on February was 1°C (34°F).

Lahore, the main city and cultural and historical midpoint of Punjab. The weather in Lahore is semi-arid. In June, the rainy season begins. The temperature rises in July.

Sialkot is located in Panjab in the northwest. It has four sub seasonal humid and subtropical characteristics. The weather in Sialkot is still hot during the day, but cool at night and low humidity. In winter, the climate is a bit warm and there is a lot of precipitation. Multan is located in the southern part of Punjab and has witnessed the most extreme temperatures in Pakistan. Multan has an arid climate, with hot summers and cold winters. Summer begins in May and lasts until September. In Multan, summer is the longest season, while in the monsoon season there is heavy rain.

This article will use all of the above five stations to characterize the maximum temperature features. The maximum temperature sequence of weather sites Hyderabad, Karachi, Lahore, Multan and Sialkot are composed from the Pakistan Meteorological Department (PMD). Thirty-six years of data will be used to find the possible trends in the highest temperature. The AMT data from 1981 to 2016 will be analyzed in several years. The locations of the eight stations are given in Figure 1.

Fig. 1. Map of five selected locations in Pakistan

Table 2. Descriptive analysis of five stations

Results and discussion

In this research, different distribution the goodness of fit test is evaluated to describe the distribution of temperature at five areas in the **north-east and southern** of Pakistan. In this affection, the Kum distribution is 1st time and previously related with some distribution.

In Table 2 these stations are measured descriptive of the entire database. Gives a detailed description of the selected sites, which contains information about the statistical characteristics of

the recording time, geographic location and temperature data. Some descriptive statistics, including the maximum temperature, standard deviation, mean, skewness, and kurtosis used by the selected sites. Table 3 displays all statistical distributions functions of Hyderabad, Karachi, Lahore, Multan and Sialkot stations. In table 3 last column shows the rank of separately distribution function. It should be noted that based on all numerical parameters, Kum is determined to be the most appropriate distribution for Hyderabad, Sialkot and Lahore sites. However, for Karachi and Multan stations the GEV distribution functions run the best fit to measure temperature data. For Hyderabad, Karachi, Lahore, Multan and Sialkot stations, the parameters to obtain the best distribution function are Kum (α_1=1.9422 α_2=1.4754 a=21.301 b=43.15), GEV(k=-0.61938 σ=3.5027 μ=31.764) Kum (α_1=1.8367 α_2=1.6086 a=14.834 b=43.175), GEV (k=-0.47692 σ=7.985 μ=30.561) and Kum (α_1=1.8116 α_2=1.5544 a=13.266 b=42.257). The consequences also express that, exclude for Hyderabad site, the Wei3 function can operate normally in other stations, so it can be used as the third or fourth best distribution. In addition, the poorest performance showed that Exponential distribution represents its lowest priority.

The main results from table 3 is more effective approaches which are not same between sites. About parameters such as temperature features, the quality and quantity of noted the temperature data may affect the validity of the distribution function to denote the temperature distribution to check the location. Further, the results indicate that skewness may be the main parameter affecting the accuracy of Kum model. It can be seen that the Kum distribution ranks first in the Lahore, Sialkot and Multan sites with lower skewness, whereas it indicate that lower efficiency for other sites with greater skewness (see table 2 and 3 association). However, in future studies, the effects of all the above parameters must be properly studied to draw conclusions.

Table 3

Additional outcome displays that for all five sites, the Gam model indications comparative flexibility for all temperature data and ranks in the top four in expressions of proficiency. In addition, the Kum, Wei3 and GEV distribution models are most flexible because they can provide good performance for all sites with different temperature features. In general, it can be decided that the Kum distribution is usually measured an effective function because it offers the most

suitable in 3 sites, and it ranks 3rd to 5th among the remaining 3 sites. However, due to the tightness of the Kum and Wei distribution functions and the flexibility of the Wei function as its broadly verified characteristics in previous studies, the efficiency of the Kum distribution should be examined more in future studies. For this reason, more situation studies with different temperature features must be estimated.

In order to illustrate that the four most suitable distribution function describe the temperature in different ranges, figure 2(a-e) shows that CDF and PDF and curves fitted by all stations. For pdf and cdf graphs, the horizontal axis is the range of temperature data. For pdf plots, the shows the probability density, which varies between the highest and lowest probable value. For the cdf graph, the perpendicular axis shows the cumulative density, as we move from left to right on the parallel axis, the value increase from 0 to 1.

Fig. 2. Cumulative distribution (CDF) and Histograms with fitted distributions for: (a) Hyderabad (b) Karachi (c) Lahore (d) Multan (e) Sialkot

The precise return level of extreme temperature must be estimate. The GEV distribution is used to fit extreme temperature sequence of each site. The change position series is installed in the station with the variation point (Yan et al., 2016). Dort and David (2016) a consistent method was used to examine the modulation of the extreme probability of temperature and rainfall, and it was found that the extreme maximum temperature has a statistically significant long-term increase, but has obvious seasonal and regional changes. They used the Wilcoxon test and Boxplots summarized the results of four AEPs (50%, 10%, 5%, and 1%) and fixed and non-fixed generalized extreme value (GEV) distributions. The occurrence of temperature waves between 1985 and 2005 in north Pakistan and the fast melting of glaciers proved the increasing trend of weather warming in Pakistan (Rasul et al., 2008). A bivariate stochastic model for the space time field of maximum and minimum temperature. The bivariate field splits in to two parts “weather” and “local climate”. The climate factor is spatially related to bivariate simulation. The statistical model adds the blocking effect of spatial variation to allow small scale variability of local variation model and successfully adapts to the stationarity of cross covariance and direct covariance functions over time.(Kleiber et al., 2013).

The estimator AIC is the prediction error of sample, and therefore the comparative superiority of the statistical model for data set. (McElreath, 2020, Taddy, 2019) given the set of models used for

the data, AIC is the quality of estimations of individually model relative to every model. Therefore, AIC provides a method of model selection. AIC is based on particular theory, when using a statistical model denote the process of generating data, the representation is almost certainly not accurate. Therefore, a model to represent the procedure will lose some information. AIC estimations the comparative amount of data missing a specified model, when assessing the total data lost, AIC will weigh the models goodness-of-fit and model simplicity. For identification of mode the application of BIC widely used in linear regression and time series. However, it can be widely applied to models based on maximum likelihood. Because the interested model is equal to number of parameters.

The other method is based on the relative measure of information loss when fitting the model to describe the data. This approach includes Akaike information criteria AIC and BIC. However, these two technique is the most popular measure. In the sense of hypothesis testing, AIC is not a model test. Rather, the process and scoring provide a method for comparing data models and a tool for model selection. The general formula for AIC and BIC is

$$AIC = -2\log(L) + 2k \quad (4)$$

$$BIC = -2\log(L) + k\log(n) \quad (5)$$

Where K is the number of parameter and L is likelihood of the fitted model [74-75]. These criteria take in to account the simplicity of the model because they contain penalties that increase with the number of parameters. AIC and BIC penalized the logarithmic probability criterion, thereby maintaining a balance between good fit and complexity. The model selection of the best fit distribution on the basis of AIC and BIC lowest value for maximum temperature.

Conclusion

In this research, the efficiency of different function was estimate to denote the function of maximum temperature at 5 sites in the Pakistan.in this study, a new function called kumarsawamy (Kum) was estimated first time. MLE is an actual parameter estimation method used to analyze the related parameters. The results shows that is impossible to present one appropriate distribution for all check sites. For Hyderabad, Lahore and Sialkot sites, the Kumaraswamy distribution was found to be most suitable for maximum temperature data, while for Karachi and Multan stations GEV functions is most appropriate. It was originate that certain parameters such as maximum temperature features, the quality and quantity of noted temperature data can be measured the

distribution as an effect on the performance. In addition, skewness is the main parameter that affects the precision of the Kumaraswamy distribution, so it ranks first in the Hyderabad, Lahore and Sialkot stations with lower skewness, and shows lower efficiency for other sites with higher skewness. However, in the future research, the impact of the above revealed important parameters of the function should be properly studied to draw conclusions. For all sites the results show that Kum, Wei3, and GEV are more flexible in distribution because they can display better performance.

Generally, this research shows that the Kum distribution function is an actual distribution because it runs the most fit in 2 stations, and it ranks 2nd among the remaining 2 stations. However, due to the tightness of the Kum, GEV Wei3 distribution functions and the flexibility of the GEV and wei3 function as its widely proven characteristics in previous studies, the effectiveness of the Kum distribution should be evaluated more in future study. For this reason, more situation study with different temperature features would also be estimated.

Acknowledgments

We thank our respected reviewers and especially Yejuan Wang for their valuable comments and suggestions that helped us to improve this paper.

Conflict of interest

The authors declare that they have no conflict of interest

Funding. No

Author's contribution

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by [Tasir khan] and [yejuan wang]. The first draft of the manuscript was written by [Tasir khan] and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.**Tasir Khan.** Data curation, Writing- Original draft preparation. **Yejuang wang;** Supervision**:**

Data availability statement

All the authors of this manuscript confirmed that the data supporting the findings of this study are available in the article. All the required data is available and easily accessible.

https://www.rstudio.com/products/rstudio/download/

https://easyfit.en.softonic.com/

Pakages
L-mom
lmomco-package
L-Moments

```
data(groundbeef)
serving <- groundbeef$serving
(fitg <- fitdist(serving, "gamma"))
gofstat(fitg)
(fitln <- fitdist(serving, "lnorm"))
gofstat(fitln)

gofstat(list(fitg, fitln))

data(toxocara)
number <- toxocara$number

fitp <- fitdist(number,"pois")
summary(fitp)
plot(fitp)

fitnb <- fitdist(number,"nbinom")
summary(fitnb)
plot(fitnb)

set.seed(1234)
x4 <- rweibull(n=1000,shape=2,scale=1)
# fit of the good distribution
f4 <- fitdist(x4,"weibull")

# fit of a bad distribution
f4b <- fitdist(x4,"cauchy")

gofstat(list(f4,f4b),fitnames=c("Weibull", "Cauchy"))

lmoments<-Lmoments(x);
lmomcov<-Lmomcov(x);
estim_params<-lmom2normpoly4(lmoments);
hist(x,30,freq=FALSE)
plotpoints<-seq(min(x)-1,max(x)+1,by=0.01);
```

```
lines(plotpoints,dnormpoly(plotpoints,estim_params),col='red');
lines(plotpoints,dnormpoly(plotpoints,true_params),col='blue');
```

Consent for publication

All the authors agree to publish this paper

Ethical statement

All experimental procedure were approved by the animal welfare and ethics committee of Lanzhou University (LZU-201805-224)

Consent to participate

This study involve no living organisms or their products so don't need any consent of participate

References

AMIN, N. A. M., ISMAIL, M. S. & HAMID, H. A. Modelling extreme temperature in Perlis using block maxima method. AIP Conference Proceedings, 2018. AIP Publishing LLC, 020010.

CASSALHO, F., BESKOW, S., DE MELLO, C. R., DE MOURA, M. M., KERSTNER, L. & ÁVILA, L. F. 2018. At-site flood frequency analysis coupled with multiparameter probability distributions. *Water resources management,* 32**,** 285-300.

CHEN, T., AO, T., ZHANG, X., LI, X. & YANG, K. 2019. Climate change characteristics of extreme temperature in the Minjiang river basin. *Advances in Meteorology,* 2019.

DONAT, M. G., LOWRY, A. L., ALEXANDER, L. V., O'GORMAN, P. A. & MAHER, N. 2016. More extreme precipitation in the world's dry and wet regions. *Nature Climate Change,* 6**,** 508-513.

FAWAD, M., YAN, T., CHEN, L., HUANG, K. & SINGH, V. P. 2019. Multiparameter probability distributions for at-site frequency analysis of annual maximum wind speed with L-Moments for parameter estimation. *Energy,* 181**,** 724-737.

GAO, X., ZHU, Q., YANG, Z., LIU, J., WANG, H., SHAO, W. & HUANG, G. 2018. Temperature dependence of hourly, daily, and event-based precipitation extremes over China. *Scientific reports,* 8**,** 1-10.

HASAN, H., RADI, N. A. & KASSIM, S. Modeling of extreme temperature using generalized extreme value (GEV) distribution: A case study of Penang. World Congress on Engineering, 2012. 181-186.

HOLMES, C. R., WOOLLINGS, T., HAWKINS, E. & DE VRIES, H. 2016. Robust future changes in temperature variability under greenhouse gas forcing and the relationship with thermal advection. *Journal of Climate,* 29**,** 2221-2236.

HUANG, W. K., STEIN, M. L., MCINERNEY, D. J., SUN, S. & MOYER, E. J. 2016. Estimating changes in temperature extremes from millennial-scale climate simulations using generalized extreme value (GEV) distributions. *Advances in Statistical Climatology, Meteorology and Oceanography,* 2**,** 79-103.

HUGHES, G. L., SUBBA RAO, S. & SUBBA RAO, T. 2007. Statistical analysis and time-series models for minimum/maximum temperatures in the Antarctic Peninsula. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 463**,** 241-259.

IRANNEZHAD, M., MORADKHANI, H. & KLØVE, B. 2019. Corrigendum to "Spatio-temporal Variability and Trends in Extreme Temperature Events in Finland over the Recent Decades: Influence of Northern Hemisphere Teleconnection Patterns". *Advances in Meteorology,* 2019.

KLEIBER, W., KATZ, R. W. & RAJAGOPALAN, B. 2013. Daily minimum and maximum temperature simulation over complex terrain. *The Annals of Applied Statistics*, 588-612.

LOLLCHUND, R. M., BOOJHAWON, R. & RUGHOOPUTH, S. D. 2014. Statistical modelling of wind speed data for Mauritius. *International Journal of Renewable Energy Research (IJRER),* 4, 1056-1064.

MAHMOOD, R. & BABEL, M. S. 2014. Future changes in extreme temperature events using the statistical downscaling model (SDSM) in the trans-boundary region of the Jhelum river basin. *Weather and Climate Extremes,* 5, 56-66.

MCELREATH, R. 2020. *Statistical rethinking: A Bayesian course with examples in R and Stan*, CRC press.

MORGAN, E., LACKNER, M., VOGEL, R. & BAISE, L. 2009. Probability distributions for offshore wind speeds. *AGUFM,* 2009, A31F-0179.

NAVEAU, P., GUILLOU, A. & RIETSCH, T. 2014. A non-parametric entropy-based approach to detect changes in climate extremes. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, 861-884.

OZER, P. & MAHAMOUD, A. 2013. Recent extreme precipitation and temperature changes in Djibouti City (1966–2011). *Journal of Climatology,* 2013.

PAREY, S., HOANG, T. T. H. & DACUNHA-CASTELLE, D. 2013. The importance of mean and variance in predicting changes in temperature extremes. *Journal of Geophysical Research: Atmospheres,* 118, 8285-8296.

SHABY, B. A. & REICH, B. J. 2012. Bayesian spatial extreme value analysis to assess the changing risk of concurrent high temperatures across large portions of European cropland. *Environmetrics,* 23, 638-648.

TADDY, M. 2019. *Business data science: Combining machine learning and economics to optimize, automate, and accelerate business decisions*, McGraw Hill Professional.

TESHOME, A. & ZHANG, J. 2019. Increase of extreme drought over ethiopia under climate warming. *Advances in Meteorology,* 2019.

WANG, G., WANG, D., TRENBERTH, K. E., ERFANIAN, A., YU, M., BOSILOVICH, M. G. & PARR, D. T. 2017. The peak structure and future changes of the relationships between extreme precipitation and temperature. *Nature Climate Change,* 7, 268-274.

YAN, B., XIA, Z., HUANG, F., GUO, L. & ZHANG, X. 2016. Climate change detection and annual extreme temperature analysis of the amur river basin. *Advances in Meteorology,* 2016.

ZHOU, B., XUE, H. & SHANG, G. 2017. Characteristics of extreme climate in downstream catchment of Yangtze River during 1960–2012. *Water Power,* 43, 26-30.

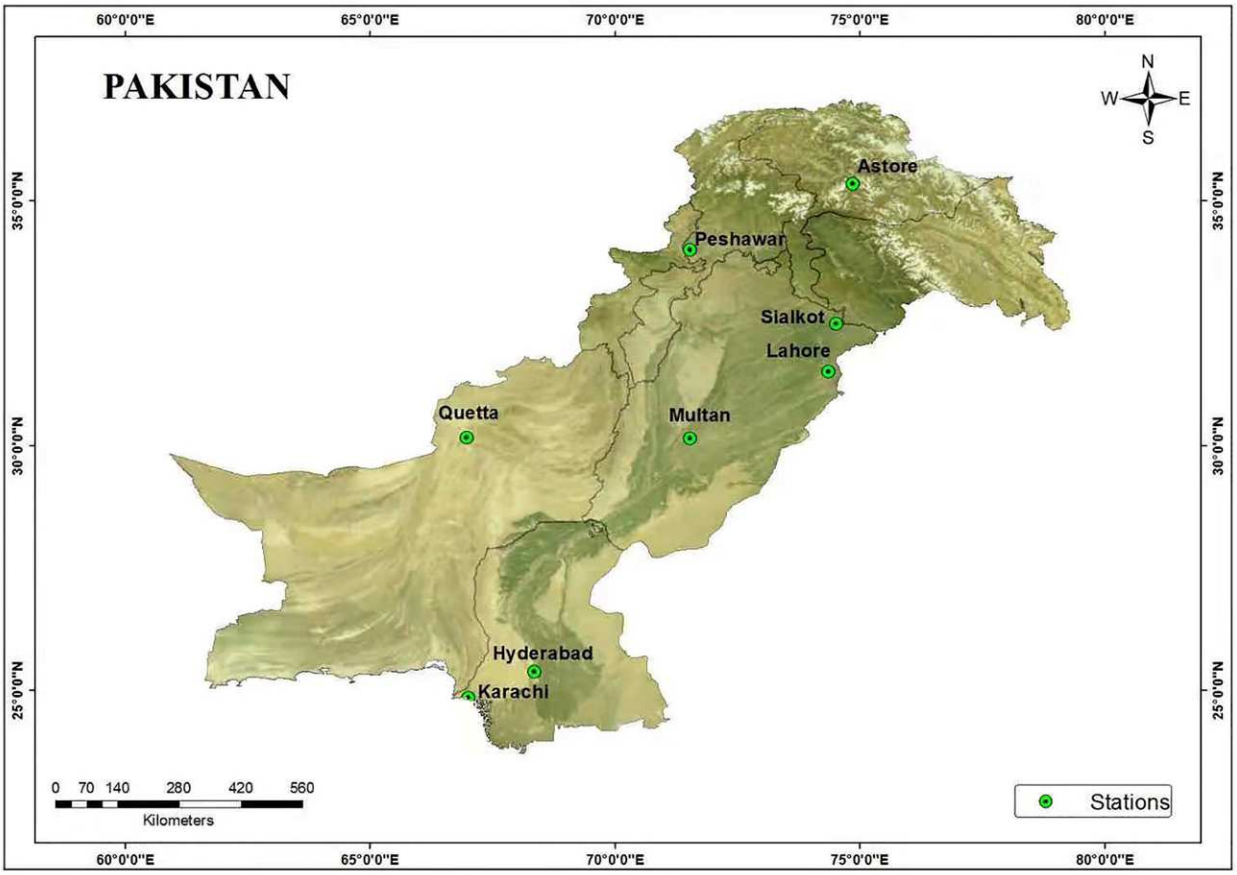
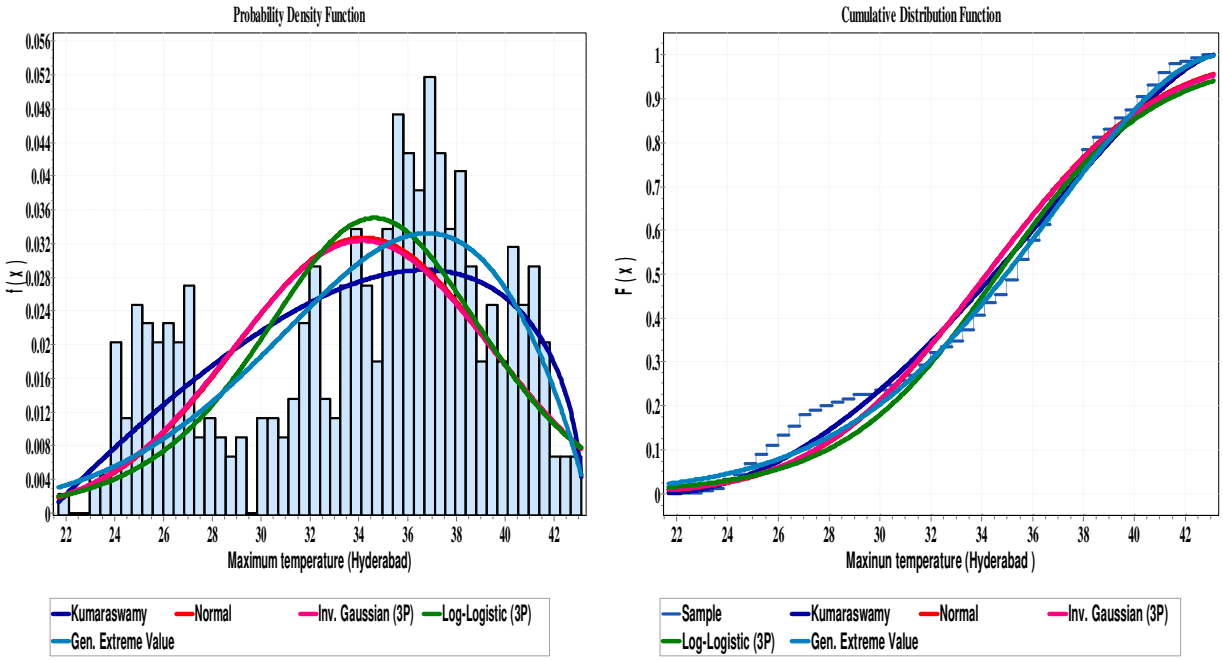
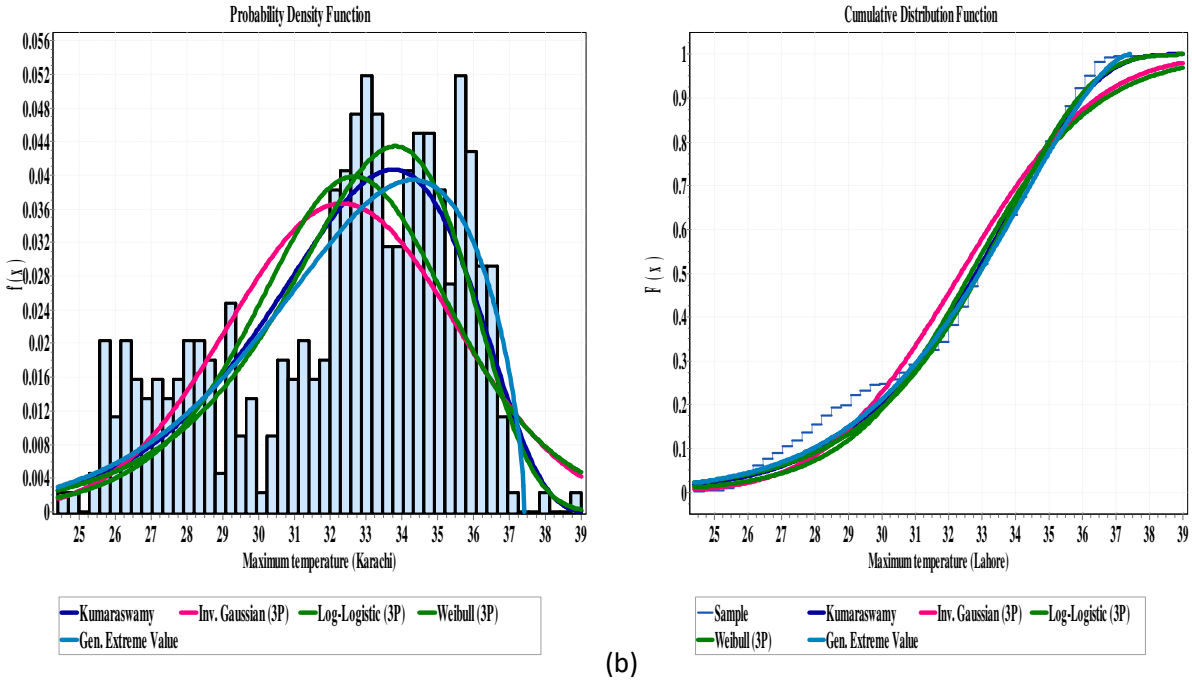


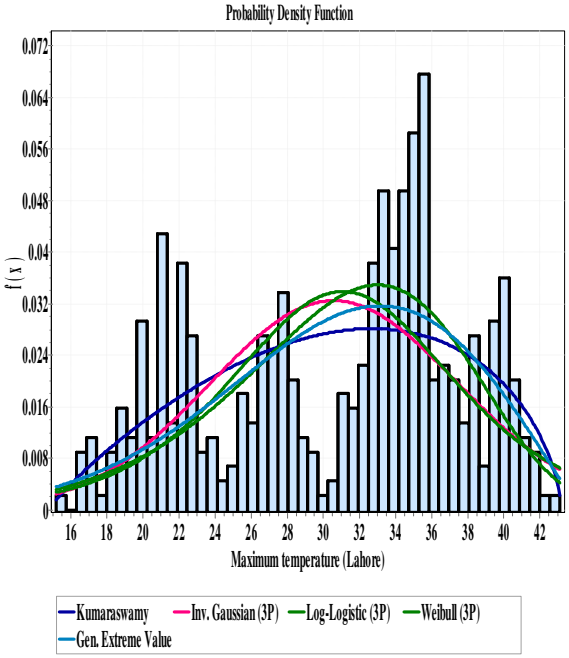
Fig. 1. Map of five selected locations in Pakistan

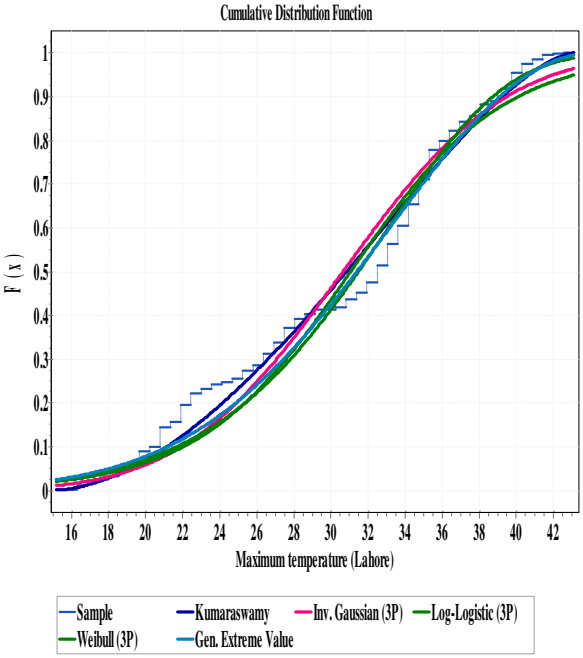

Probability Density Function
f (x)
Maximum temperature (Hyderabad)
Kumaraswamy
Normal
Inv. Gaussian (3P)
Log-Logistic (3P)
Gen. Extreme Value
Cumulative Distribution Function
F (x)
Maxinun temperature (Hyderabad)
Sample
Kumaraswamy
Normal
Inv. Gaussian (3P)
Log-Logistic (3P)
Gen. Extreme Value

(a)

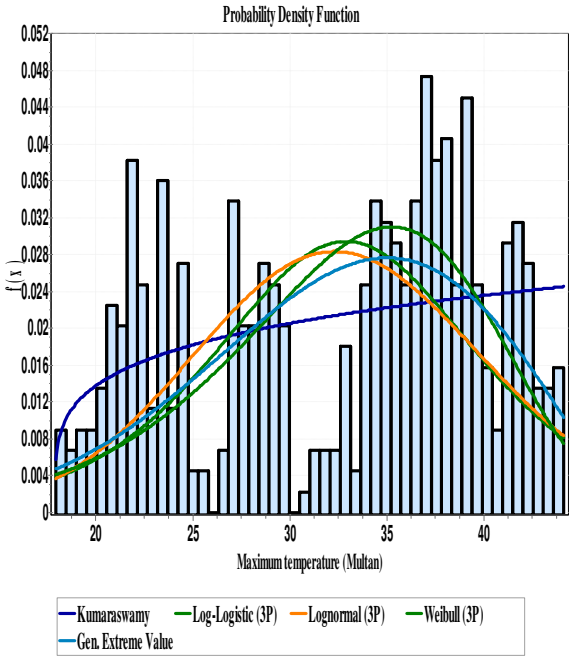

Probability Density Function
f (x)
Maximum temperature (Karachi)
Kumaraswamy
Inv. Gaussian (3P)
Log-Logistic (3P)
Weibull (3P)
Gen. Extreme Value
Cumulative Distribution Function
F (x)
Maximum temperature (Lahore)
Sample
Kumaraswamy
Inv. Gaussian (3P)
Log-Logistic (3P)
Weibull (3P)
Gen. Extreme Value

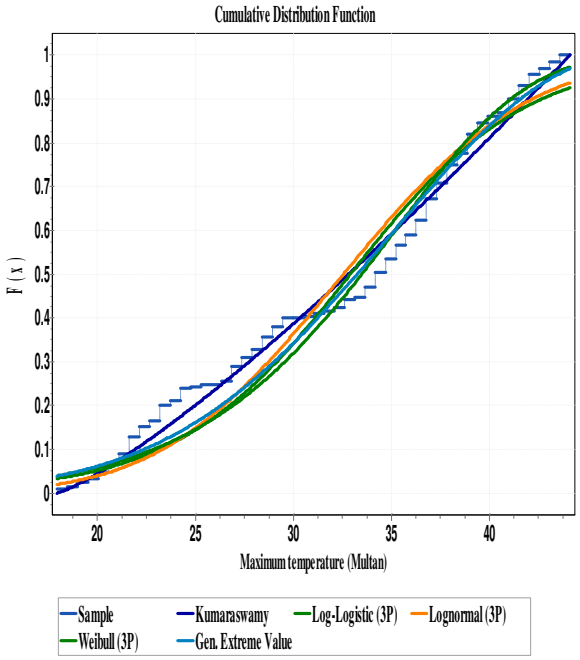
(b)


Probability Density Function
f (x)
Maximum temperature (Lahore)
Kumaraswamy
Inv. Gaussian (3P)
Log-Logistic (3P)
Weibull (3P)
Gen. Extreme Value

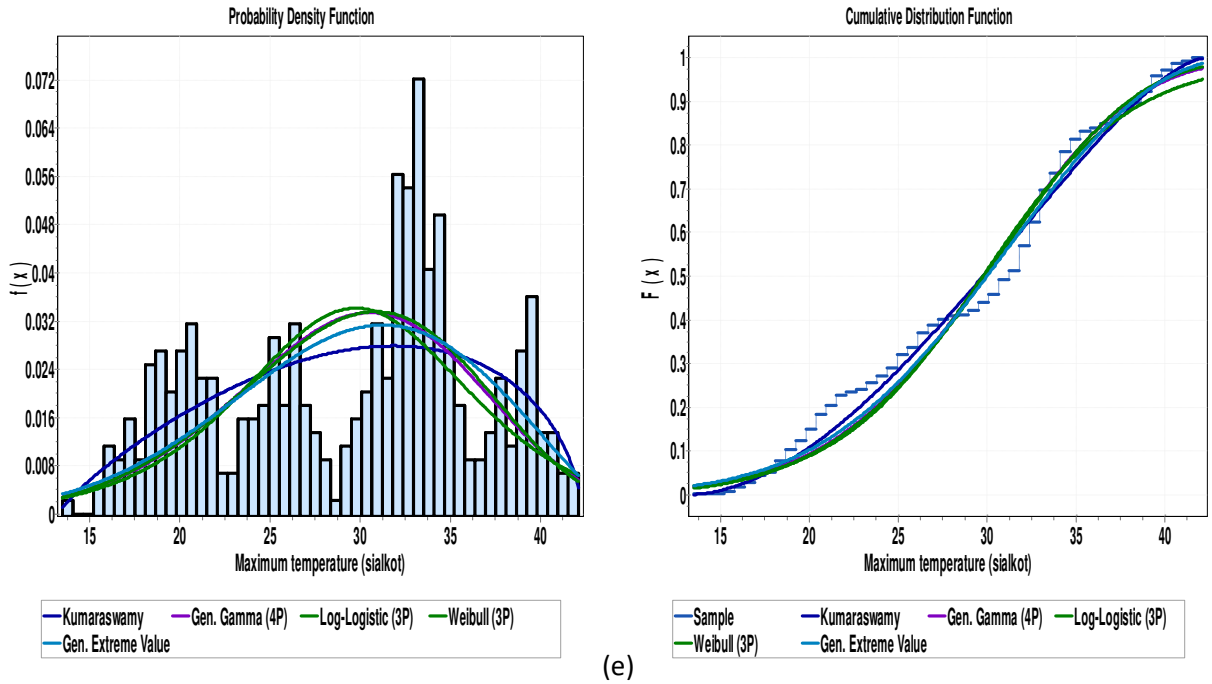

Cumulative Distribution Function
F (x)
Maximum temperature (Lahore)
Sample
Kumaraswamy
Inv. Gaussian (3P)
Log-Logistic (3P)
Weibull (3P)
Gen. Extreme Value

1 (c)


Probability Density Function
f (x)
Maximum temperature (Multan)
Kumaraswamy
Log-Logistic (3P)
Lognormal (3P)
Weibull (3P)
Gen. Extreme Value


Cumulative Distribution Function
F (x)
Maximum temperature (Multan)
Sample
Kumaraswamy
Log-Logistic (3P)
Lognormal (3P)
Weibull (3P)
Gen. Extreme Value

2 (d)



(e)

Fig.2 Histograms with fitted distributions and cumulative distribution functions for: (a) Hyderabad (b) Karachi (c) Lahore (d) Multan and (e) Sialkot

Table. Probability distributions (PDF) and cumulative distribution (CDF) for each distribution.

Distributions	**PDF**	**CDF**
Exponential (2P)	$f(x) = \lambda \exp(-\lambda(x-\lambda)$	$F(x) = 1 - exp(-\lambda(x-\gamma))$
Normal		
Gamma	$f(x) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} \exp(-x/\beta)$	$F(x) = \frac{\Gamma x/\beta(\alpha)}{\Gamma(\alpha)}$
Inv. Gaussian	$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right)$	$F(x) = \emptyset\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right) + \emptyset\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}+1\right)\right)\exp(2\lambda/\mu)$
Logistic		
Log-Logistic 3	$f(x) = \frac{\alpha}{\beta}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}\left(1+\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{-2}$	$F(x) = \left(1+\left(\frac{\beta}{x-\gamma}\right)^{\alpha}\right)^{-1}$

Log normal	$f(x) = \frac{\exp\left(-\frac{1}{2}\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right)^2\right)}{(x-\gamma)\sigma\sqrt{2\pi}}$	$F(x) = \emptyset\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right)$
Gamble,	$f(x) = \frac{1}{\sigma}\exp(-z-\exp(-z))$	$F(x) = \exp(-\exp(-z))$
KUM GEV	$f(x) = \begin{cases} \frac{1}{\sigma}\exp\left(-(1+kz)^{-\frac{1}{k}}\right)(1+kz)^{-\frac{1}{k}} & k \neq 0 \\ \frac{1}{\sigma}\exp(-z-exp(-z)) & k = 0 \end{cases}$	$F(x) = \begin{cases} \exp\left(-(1+kz)^{-\frac{1}{k}}\right)(1+kz)^{-\frac{1}{k}} & k \neq 0 \\ \exp(-z-exp(-z)) & k = 0 \end{cases}$
Gamma (3P)	$f(x) = \frac{(x-\gamma)^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}\exp(-(x-\gamma)/\beta)$	$F(x) = \frac{\Gamma(x-\gamma)/\beta(\alpha)}{\Gamma(\alpha)}$
Pearson type 3	$f(x) = \frac{1}{X\lvert\beta\rvert\Gamma(\alpha)}\left(\frac{\ln(X)-\gamma}{\beta}\right)^{\alpha-1} exp\left(-\frac{\ln(X)-\gamma}{\beta}\right)$	$F(x) = \frac{\Gamma(\ln(x)-\gamma)/\beta(\alpha)}{\Gamma(\alpha)}$
Generalize gamma	$f(x) = \frac{k(x-\gamma)^{k\alpha-1}}{\beta^{k\alpha}} exp(-(x-\gamma)/\beta^{k})$	$F(x) = \frac{\Gamma((x-\gamma)/\beta)^{k(\alpha)}}{\Gamma\alpha}$
Weibull (3p)	$f(x) = \frac{\sigma}{\beta}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}\exp\left(-\left(\frac{x-\gamma}{\beta}\right)\right)^{\alpha}$	$F(x) = 1 - exp\left(-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)$
GAMMA	$f(x) = \frac{(X-\gamma)^{\alpha-1}}{\beta^{\alpha}\Gamma\alpha}\exp(-\frac{x-\gamma}{\beta})$	$F(x) = \frac{\Gamma(x-\gamma)/\beta(\alpha)}{\Gamma\alpha}$
LN	$f(x) = \frac{\exp\left(-\frac{1}{2}\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right)^2\right)}{(x-\gamma)\sigma\sqrt{2\pi}}$	$F(x) = \emptyset\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right)$

Table 2. Descriptive analysis of five stations

Variable	SD	Skewness	Kurtosis	Mean	Max Temp	Latitude	Longitude	Altitude (m)
Hyderabad	0.78	0.03	0.14	41.32	43.10	**25°23'**	**68°22'**	7
Karachi	0.76	1.08	4.16	36.23	39	24°54'	67°4'	4
Lahore	1.25	-0.46	1.25	40.30	43.10	31°33'	74°19'	215
Multan	1.06	-0.26	-0.61	42.32	44.10	30°11'	71°28'	123
Sialkot	1.55	-0.14	-0.30	40.04	43.23	32°30'	74°31'	256

Table 3. Achieved statistical indicators of different distribution functions for the selected Stations.

Stations	Models	KS	AD	AIC	BIC	Ranking
Hyderabad	EXP	0.1624	3.9981	2776.031	2785.011	12
	N	0.0671	0.4159	2735.534	2744.677	3
	IG	0.06762	0.4161	2735.999	2745.908	4
	L	0.08922	0.7404	2755.989	2764.908	10
	LL3	0.07071	0.4874	2737.679	2746.980	5
	LN3	0.07155	0.4541	2744.210	2752.401	7
	GUM	0.11081	1.1564	2757.399	2765.590	11
	KUM	0.03591	0.2765	2730.076	2741.980	1
	GEV	0.05475	0.2988	2732.890	2743.066	2
	Wei3	0.08661	0.6363	2752.534	2761.920	9
	PE3	0.07356	0.4596	2750.456	2759.098	8
	GG (4)	0.0713	0.4497	2741.089	2750.098	6
Karachi	EXP	0.28726	70.427	2314.958	2323.149	**12**
	N	0.1191	10.686	2259.021	2267.032	6
	IG	0.11835	10.599	2254.753	2261.098	5
	L	0.11953	12.821	2264.123	2272.701	7
	LL3	0.09089	9.2624	2252.382	2261.890	4
	LN3	0.12672	11.496	2265.067	2273.032	8
	GUM	0.18769	38.955	2303.642	2311.833	**11**
	KUM	0.07047	3.7466	2233.042	**2241.021**	**2**
	GEV	0.06033	10.164	2221.620	2229.042	1
	Wei3	0.07542	3.6572	2235.021	2242.324	3
	PE3	0.13562	12.357	2279.034	2287.435	10
	GG (4)	0.12716	11.561	2276..067	2284.032	9
Lahore	EXP	0.23172	57.062	3054.008	**3067.098**	**12**
	N	0.13854	9.5406	3005.043	3012.760	8
	IG	0.13415	9.6267	2954.231	2964.098	6
	L	0.15459	13.645	3011.890	3020.005	8
	LL3	0.11894	9.5841	2875.078	2885.675	4
	LN3	0.13616	10.041	2999.910	3008.101	7
	GUM	0.20152	29.918	3021.575	3029.767	11
	KUM	0.0125	4.5895	2800.981	2812.109	1
	GEV	0.0939	4.7563	2823.081	2835.142\	2
	Wei3	0.09681	6.4167	2856.052	2864.160	3
	PE3	0.15904	12.143	3017.879	3024.009	10
	GG (4)	0.13815	9.5734	2912.098	2922.564	5

Multan	EXP	0.21595	42.866	3075.670	**3084.003**	**11**	1
	N	0.12852	10.988	3023.006	3031.890	6	2
	IG	0.13224	11.123	3033.045	3041.008	7	
	L	0.14184	15.788	3041.678	3049.014	8	
	LL3	0.10831	10.906	3005.352	3013.008	3	
	LN3	0.12375	11.209	3073.808	3082.000	5	
	GUM	0.19464	30.893	3056.536	**3063.002**	**10**	
	KUM	0.09133	3.7197	2956.870	**2967.054**	**2**	
	GEV	0.08861	6.0058	2901.065	2911.045	1	
	Wei3	0.10279	7.9024	2978.076	2988.078	4	
	PE3	0.14867	12.568	3044.081	3052.061	9	
Sialkot	EXP	0.22985	58.414	3148.089	3153.079	12	
	N	0.12617	7.929	3011.484	3019.676	7	
	IG	0.12081	8.2295	2983.072	2992.560	6	
	L	0.14319	11.702	3042.014	3048.019	10	
	LL3	0.11209	8.2	2957.546	2968.745	5	
	LN3	0.12929	8.3351	3032.168	3040.360	8	
	GUM	0.18772	26.492	3134.089	3142.190	11	
	KUM	0.09057	3.4658	2899.043	2901.043	1	
	GEV	0.09113	4.6078	2913.060	2920.609	2	
	Wei3	0.10358	6.3469	2925.012	2932.081	3	
	PE3	0.14089	9.5884	3034.073	3042.009	9	
	GG (4)	0.10862	6.4591	2934.053	2945.019	4	