Original Research Paper

Is Air Pollution Controlled Enough? "Tapered Probabilities" Answer

Ramalingam Shanmugam

Honorary Professor of International Studies, School of Health Administration,
Texas State University, San Marcos, TX 78666, USA

Email: rs25@txstate.edu

Abstract: Air pollution is dangerous sespecially to the residents in cities, as it causes lung, heart, or respiratory illness. Public health professionals ponder whether the air pollution is significantly controilled. The amount of existing air pollution is measurable, but not its regulated level, which is therefore as a parameter in a novel manner in this article. However, there is no suitable methodology in the literature to extract and test whether the an estimated data evidence about the regulated air pollution level is significant. To fulfill the need, this article constructs a methodology based on tapered probability model for the air pollution data. Our new methodology is demonstrated using the air pollution data in African, American, Asian, European and Oceanic continents.

Introduction

Motivation

Air pollution is too serious menace to the residents in cities worldwide, when they are vulnerable to respiratory related illness (see Dominici *et al*., 2000 for details). More often than not, air pollution is life threatening to patients with asthma or respiratory complications (Louwies *et al*., 2013). Air pollution data are meticulously collected but are not well utilized to learn their intricate patterns. A reason is detected and it is the lack of suitable analytic methodology. After a thorough literature search, we notice that no satisfactory underlying probability model for the air pollution data seems to have been identified in the literature. We propose and construct a versatile model called Tapered Probability Function (TPF) to address the significance of the regulated level of the air pollution in this article.

The TPF is not popular among the healthcare professionals but it is popular among the geologists who deal with the earthquakes. See Kagan and Schoenberg (2001) for details on the earthquake application of the TPF. Estimation and hypothesis testing of the parameters in TPF are still technical challenges. To ease the mathematical and computational difficulties of the TPF, an innovative approach is devised, presented and explained in this article. New expressions for the methodology are derived, displayed and interpreted in the context of discussion about air pollutions. This new approach is utilized to assess the significance of an estimated regulated level of the air pollution in major cities across the African, Asian, American, European and Oceanic continents. The contents of this article are made easier for the applied healthcare in their practice to test the significance of an estimated regulated level of the air pollution.

A threat to live healthy in this advanced electronic age of 21st century is the air pollution related hazards. See Mateen and Brook (2011) for details about the health impacts of air pollution. What is air pollution? Air pollution refers the existence of harmful visible or invisible substances including particulates and molecules in the earth's atmosphere. Air pollution has adverse effects on human health. The World Health Organization warned that the air pollution in the year 2014alone caused the premature death of 7 million people worldwide with the highest death rate in India compared to other nations. See Mateen and Brook (2011) for details.

The developing nations aspire to improve their standard of life. Their industrial or agricultural activities are the source of air pollutions. The developed nations tend to maintain their higher standard of living by continuing on their automobile emissions caused air pollutions. For a different reason, the developing and developed nations do pollute air. What do the air pollution data suggest about who among the developed and developing nations pollute more? Shanmugam and Hertelendy (2011) answered this question. Stronger policies to reduce the emission of air pollutants globally would help.

Science Publications

More and more industrial wastes are dumped to air, land and water sources and they cause tremendous illnesses. In this process, a significant amount of oxygen in the natural environment is depleted with an increase of adverse toxics in the air. Mainly recognized three culprits in the air pollution are PM_{10}, SO_2 and NO_2.

How does PM_{10} affect our health? When the PM_{10} is inhaled, the particles evade the respiratory system's natural defense and lodge deep deposit in the lungs (Pope *et al.*, 2002). Health problems begin to arise as the body reacts to these foreign particles. The "sensitive populations" to PM_{10} are children, elders, exercising adults and those suffering from asthma or bronchitis. The presence of PM_{10} increases the severity of asthma attacks, cause bronchitis, inflict lung diseases and reduce the body's ability to fight infections (see Weinmayr *et al.*, 2010 for details). See Figure 1 through Figure 3 for the trend of PM_{10} in global major cities in Africa, America, Asia, Europe and Oceanic continents.

The SO_2 denotes sulfur dioxide and it is a toxic chemical gas with a pungent, irritating smell. More often, it is by volcanos. In ancient times, Romans injected SO_2 into wine bottles to keep wine fresh and free from bad smell. Inhaling SO_2 causes respiratory difficulties and related lung diseases.

The NO_2 is a reddish-brown chemical gas with odor and is often a by-product of human activities, volcanos and lightning (see Figure 10, Figure 11 and Figure 12). The NO_2 absorbs sunlight and ozone's thickness. Abundant NO_2 in air triggers skin cancer, mild irritation in nose and throat, bronchitis, pneumonia and lung diseases.

It appears from the data (see Table 2 and Figure 13 for details) that the level of PM_{10} is more when the level of SO_2 is more and vice versa. Hence, this article selects and focuses just the pollutant PM_{10} for data analytics and continental comparisons (see Figure 1 through Figure 5 for the trend in African, American, Asian, Europe and Oceanic continents).

Data are collected in general worldwide to understand the patterns, sources and practical ways to reduce air pollutions if not their total elimination. However, the efforts are not satisfactory enough always. Why is it so? It is perhaps due to a lack of suitable data analytic methodology, which requires finding a suitable underlying model for the collected data. What is model? The model is an abstraction of the reality. The model has to be simple and yet, powerful to capture the essence in reality. One such powerful model, which could help to capture and assess the regulated level of the air pollution is TPF.

Returning to the discussion of air pollutants, what is particulate matter, PM_{10}? It refers to fine suspended (less than 10 microns in diameter which is 1/7th thickness of the human air) capable of penetrating deep into the respiratory tract and causing health damage.
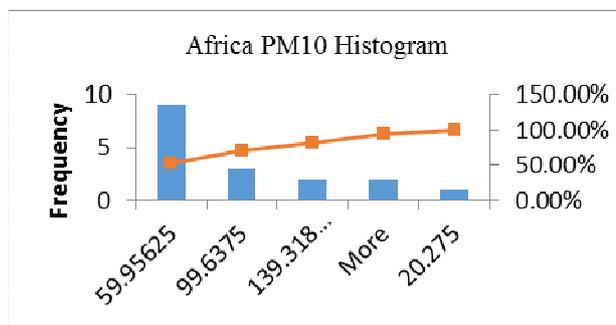
Fig. 1. Africa's PM_{10}
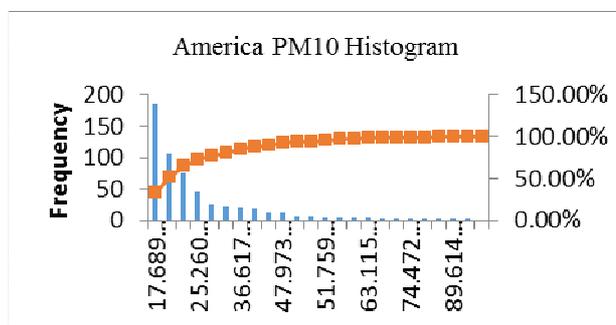
Fig. 2. America's PM_{10}
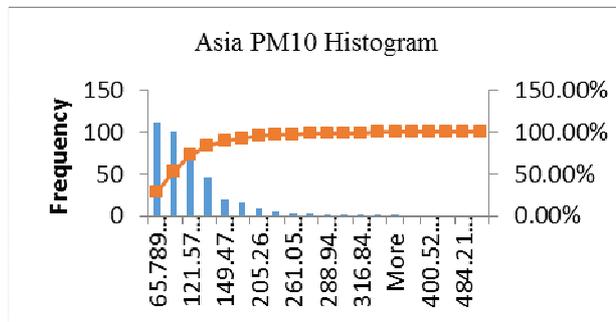
Fig. 3. Asia's PM_{10}
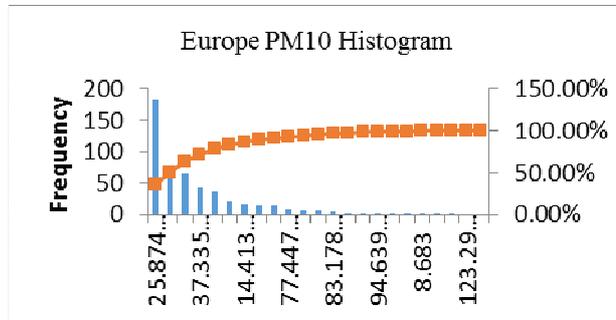
Fig. 4. Europe's PM_{10}

Table 1. Average population, PM_{10}, SO_2 and NO_2 among countries

Continent (#nations)	Average Population in 1,000	PM_{10}		SO_2		NO_2	
		Ave	Min	Ave	Min	Ave	Min
Asia (38)	6,438	75.52	31	81	6.00	54.79	13
America (11)	9,045	28.81	10	24.54	1.0	62.54	5.00
Europe (33)	2,533	28.60	11	19.51	3.00	55.09	20
All (82)	5,069	48.57	10	46.75	1.00	55.29	5.00

Table 2. Correlation (with p-value) among PM_{10}, SO_2 and NO_2, controlling population

$variable \downarrow / variable \rightarrow$	PM_{10}	SO_2	NO_2
PM_{10}	1	0.47 (0.000)	0.02 (0.83)
SO_2		1	0.18(0.10)
NO_2			1
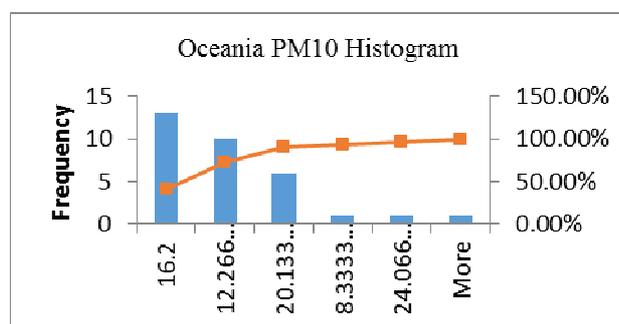
Fig. 5. Oceanic's PM_{10}

See Pope *et al*. (2016) and Malley *et al*. (2017) for details on how the particulate matter destroys health. The PM_{10} is a mixture of materials such as smoke, soot, dust, salt, acids and metals. Where does PM_{10} come from? It appears in the atmosphere when gases are emitted from the motor vehicles or industries. Other sources include stoves, fireplaces, dust from construction, landfills, wildfires, brush burning, industrial sources and windblown dust from the open lands.

The other two troubling air pollutants to a healthy living are sulfurdioxide (SO_2) and nitrogendioxide (NO_2). What is SO_2? It is an airpollutant produced when fossil fuels containing sulfur are burned. It contributes to the acid rain and it can damage human health, particularly that of the young and the elders (see Counter and Buchanan, 2004 for details). The SO_2 is a toxic gas with a pungent, irritating smell, released naturally by the volcanic activity. In the ancient days, the SO_2 was entered by burning sulfur candles inside empty wine vessels to keep the wine fresh and free from vinegar smell. The SO_2 is a noticeable in the regular atmosphere. The SO_2 is a precursor to the acid rain and all health problems.

What is NO_2? The NO_2 is a chemical compound due to industrial synthesis of the nitric acid. At a higher temperature, the NO_2 appears as a reddish-brown gas with a biting odor. It is a paramagnetic, bent molecule. It is introduced into the environment by several natural causes. It is used in the space vehicles such as the Titan rockets by the NASA and space agencies. The NO_2 is a poisonous, pungent gas formed when nitric oxide combines with hydrocarbons and sunlight. See Weinmayr *et al*. (2010) for details.

What are the tolerance levels of the air pollutions? The World Health Organization (WHO) recommends no more than20 micrograms per cubic meter for PM_{10}, no more than 40micrograms forNO_2 and no more than20 micrograms per cubic meter for SO_2 (http://apps.who.int). In accordance with the US Clean Air Act, the Environmental Protection Agency (EPA) often reviews the national air quality standards for PM_{10}, SO_2 and NO_2. The remedies to control air pollutions are dust control in roads, building constructions, landfills, landscaping, barrier, fencing to reduce windblown dust, programs to reduce emission from wood stoves and fireplaces, cleaner-burning gasoline and diesel fuels, emission control devices for motor vehicles, controls for industrial facilities etc.

Extracting pertinent evidence from the collected air pollution data and learning the realities using appropriate analytics to formulate healthy environmental policies or to execute regulating procedures are too technical stumbles for the practicing environmentalists on daily basis. To alleviate such stumbles, this article is prepared, articulated and demonstrated with real data on air pollution. One unique environmental policy for all worldwide cities is a feasible target.

This article first introduces formally in the next section a versatile and powerful model called Tapered Probability Function (TPF) with its statistical properties. The aims of this article are to construct a data analytic methodology to use for a statistical comparison of the cities across three ambitiously industrializing continents (namely African, Asian, American, European and Oceanic) with respect to their air pollution levels. Neat new expressions are derived in this article to make a hypothesis testing of an estimated regulated level of the air pollution. The p-values and statistical powers for the three continents are calculated and compared (see Abramowitz and Stegun, 1972 for details). By emulating our new methodology of this article, practicing environmentalists could analyze the air pollution data of their own, rank and classify the places in different groups.

How Tapered Frequency Distribution Canassess Air Pollution Level?

To begin with, let *Y* be a random variable indicating the air pollution level in a city at a specified time. With the parameters $\theta > 0, \tau \geq 0$ and $\phi \geq 0$ denoting respectively the severity level, a threshold level and an regulated level of the air pollution pattern, let the probability density function:

$$f(y|\tau,\theta,\phi) = \left(\frac{\phi}{y} + \frac{1}{\theta}\right)\left(\frac{\tau}{y}\right)^{\phi} e^{-\left(\frac{y-\tau}{\theta}\right)};\tag{1}$$

$$0 \leq \tau \leq y < \infty; \theta > 0; \phi \geq 0;$$

with a survival function:

$$S(y|\tau,\phi,\theta) = \Pr(T \geq y|\tau,\phi,\theta)$$
$$= \int_{\tau}^{y} S(t|\tau,\phi,\theta)dt = \left(\frac{\tau}{y}\right)^{\phi} e^{-\left(\frac{y-\tau}{\theta}\right)}\tag{2}$$

be the underlying model for the data on *Y*. The model (1) is called Tapered Probability Function (TPF). A reason for selecting the TPF (1) is for its versatility to portray the regulated level of the air pollution. The TPF is quite familiar to the geologists who unravel the mysteries behind the earthquakes. The TPF is not that much familiar to the environmentalists or healthcare professionals who deal with air pollution or its impact on health. Some heuristic introductions of the TPF are worthwhile.

First, is f (*y*|τ, θ, φ) in (1) a bona fide probability function? To be a bona fide, a function should be non-negative and the area under the functions should add up to one. For the specified parametric space $0 \leq \tau; \theta > 0; \phi \geq 0;$ and the sample space $0 \leq \tau \leq y < \infty$, the function $f(y|\tau,\theta,\phi)$ is non-negative. In addition, the area under the non-negative function is one because
$$\int_{\tau}^{\infty}\left(\frac{\phi}{y}\right)\left(\frac{\tau}{y}\right)^{\phi} e^{-\left(\frac{y-\tau}{\theta}\right)}dy = 1 - \int_{\tau}^{\infty}\left(\frac{1}{\theta}\right)\left(\frac{\tau}{y}\right)^{\phi} e^{-\left(\frac{y-\tau}{\theta}\right)}dy.$$

Hence, the function $f(y|\tau,\theta,\phi)$ in (1) is indeed a bona fide probability function. Some heuristic interpretations of the TPF (1) now follows. First, the Tail Value at Risk (*TVaR*) of air pollution function is:

$$E(Y|Y > y_p) = \frac{-\int_{y_p}^{\infty} y dS(y|\tau,\phi,\theta)}{S(y_p|\tau,\phi,\theta)}\tag{3}$$
$$= \phi\theta e^{\frac{y_p}{\theta}}$$

What does the *TVaR* (3) refer? Given the air pollutionexceeds a tolerance mark y_p (recall that asthma and other respiratory patients in cities like Beijing, Denver, Delhi, Albuquerque among other global cities are advised not to leave home when the pollution level is too dangerous), the average pollution level (see Table 1) on the day is what the expression $E(Y|Y > y_p)$ in (3) indicates. Notice that it increases as the regulated parameter $\phi \geq 0$ increases. Of course, the severity parameter $\theta > 0$ also plays a role. Furthermore, in such a day for asthma and other respiratory patients not to leave their home, will the pollution not worsen by staying within an extra level "s"? This is called hope probability and it is:

$$\Pr(Y \leq m + s | Y > m) = \frac{S(m + s|\tau,\phi,\theta)}{S(m|\tau,\phi,\theta)}\tag{4}$$
$$= (1 + \frac{s}{m})^{-\phi} e^{-\frac{s}{\theta}}$$

See Figure 6 for the dynamics of the hope probability in z-axis, θ in x-axis and ϕ in y-axis with m = 10 and s=1. Two things need to be noticed in hope probability (4). They are: (1) when the regulated level, ϕ diminishes to zero due to complete control or as the severity parameter, θ diminishes, the hope probability is lesser and (2) the hope probability is lesser when the added amount "s" is larger.

However, if the air pollution exceeds the tolerance mark "t", will it remain within the safety mark "m"? This is called safety probability and it is:

$$\Pr(Y \leq m | Y > t) = \frac{S(m + s|\tau,\phi,\theta)}{S(m|\tau,\phi,\theta)}\tag{5}$$
$$= 1 - (\frac{t}{m})^{\phi} e^{-(\frac{m-t}{\theta})}$$

The safety probability (5) vanishes, when the tolerance and safety marks are closer. When the regulated parameter ϕ approaches zero (due to strict environmental regulations), thesafety probability (5) converges to only a finite non-zero amount $1 - e^{-(\frac{m-t}{\theta})}$, but not zero unless the severity rate θ is extremely large. For a larger pollution severity (that is, $\theta \to \infty$), the safety probability (5) converges to a larger amount $1 - (\frac{t}{m})^{\phi}$ depending only on the regulated parameter ϕ. Such observations warn us the importance of dealing simultaneously both the severityand the regulated parameters in discussions of air pollution.

How much memory of the past air pollution pattern is kept by the nature? That is, if the air pollution has crossed a tolerance mark *m*, how probable for the pattern to continue so the air pollution will exceed an additional allowance "*S*"? An answer to this question resides in the memory function (6). That is:

$$memory_{m,s,\tau,\theta,\phi}$$
$$= \Pr(Y \geq m + s | Y \geq m, \tau, \phi, \theta) - \Pr(Y > s)\tag{6}$$
$$= e^{-\left(\frac{m-\tau}{\theta}\right)}\left(\frac{\tau}{s}\right)^{\phi}\{1 - \left(1 + \frac{s}{m}\right)^{-\phi} e^{-\left(\frac{s}{\theta}\right)}\}.$$

17
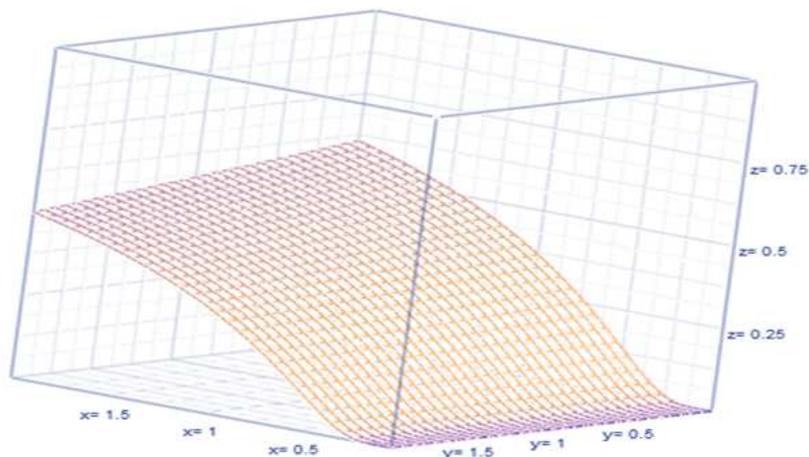
Fig. 6. Hope probability in z-axis, s in x-axis and m in y-axis
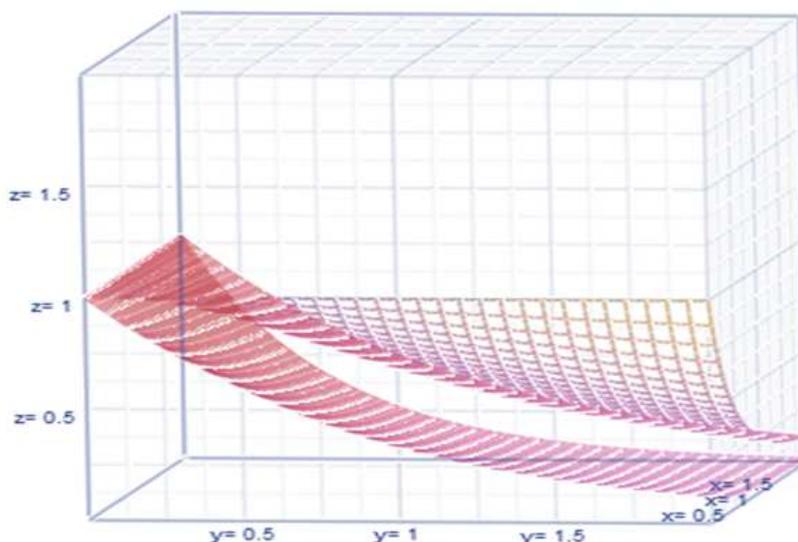
Fig. 7. Monotonic decreasing memory in z-axis, s in x-axis and m in y-axis

From the memory function (6), realize that the nature's decreasing memory is expedited by two different co-factors. The first factor $\left(1+\dfrac{s}{m}\right)^{-\phi}$ accommodates the influence by the regulated parameter ϕ. The second factor accommodates the influence by the severity parameter θ. See Figure 7 for how the dynamics of the nature's memory changes as both the severity and regulated parameters shift. We would have missed the above observations without using TPF (1).

So far, we have witnessed several advantages of using the TPF (1) to capture and interpret all about air pollutions. The mathematical difficulties of the TPF do continue as we proceed to construct a data analytic methodology. To begin realizing it, note that the mean $E(y|\tau,\theta,\phi)$ and variance, $Var(y|\tau,\theta,\phi)$ of the TPF (1) are:

$$E(y|\tau,\theta,\phi) = \tau + Shift_{\tau,\theta,\phi}, \tag{7}$$

where, $\phi\theta = \tau$ and hence, the pollution level shifts from the threshold, τ and the shift is:

$$\begin{aligned} &Shift_{\tau,\theta,\phi} \\ &\approx \{\frac{(\tau e)^{\phi}\Gamma(1-\phi)\Pr(\chi^2_{\max(1,2)|1-\phi|df} > 2\phi)}{(2\theta)^{\phi-1}}\}, \\ &\rightarrow 2\theta, if\, \phi \rightarrow 0 \end{aligned} \tag{8}$$

And:

$$Var(y|\tau,\theta,\phi)$$

$$= 2Shift_{\tau,\theta,\phi}\left|[\tau + 2\theta(1-\phi) - Shift_{\tau,\theta,\phi}]\right| \quad (9)$$

The chi-squared distribution function, $\Pr(\chi^2_{mdf} > a)$ has been widely tabulated and it is made available. Usually the Moment Estimators (ME) are easier, in general, compared to the Maximum Likelihood Estimators (MLE). However, for the TPF (1), even the MLEs are difficult as they are seen in (7) and (8) when the regulated parameter approaches one.

Let us now find the MLE. For this purpose, consider a random sample $y_1, y_2, .. y_n$ from the TPF (1). The obvious Maximum Likelihood Estimator (MLE) of the threshold parameter τ is:

$$\hat{\tau} = \min(y_1, y_2, .. y_n) = y_{(1)} \quad (10)$$

for the likelihood function to be non-zero. Then, the log-likelihood function is:

$$\ln L(\hat{\tau},\theta,\phi) = \sum_{i=1}^{n}\{\ln\left(\frac{\phi}{y_i} + \frac{1}{\theta}\right)$$

$$+ \phi\ln(\frac{y_{(1)}}{y_i}) - \left(\frac{y_i - y_{(1)}}{\theta}\right)\}$$

$$\approx -n\{\ln\theta + \frac{(\bar{y} - y_{(1)})}{\theta} \quad (11)$$

$$- \phi[\theta\bar{y}_{HM} + \ln y_{(1)} - \bar{y}_{LM}]\}$$

$$- \frac{(\phi\theta)^2}{2}\sum_{i=1}^{n}\frac{1}{y_i^2}$$

Now, with the notation $\partial_{\hat{\tau}}$ indicating the derivative with respect to $\hat{\tau}$, equating the score function $\partial_{\hat{\tau}}\ln L(\hat{\tau},\theta,\phi) = 0$ provides an estimated balancing expression:

$$\hat{\tau} = y_{(1)} = \hat{\phi}\hat{\theta} \quad (12.a)$$

where, $\hat{\phi}$ and $\hat{\theta}$ are MLEs of the regulated level and severity level of the air pollution and are yet to be found. See Figure 8 about the triangular relation among parameters. Reversing their estimates, we notice a non-estimated balancing expression:

$$\phi\theta = \tau \quad (12.b)$$

The expression (12.b) may be heuristically interpreted in a sense as follows. Recall that the parameters θ and ϕ respectively portray the severity and regulated level of the air pollution. In practice, when the air pollution's severity is more, the regulated level must have been lesser and vice versa.

The expression (12) asserts that for a specified threshold $\hat{\tau}$, the efforts are to reduce the estimated regulated level $\hat{\phi}$ for a higher estimated severity $\hat{\theta}$ and vice versa. The MLEs are the simultaneous solutions of the nonlinear score functions $\partial_{\phi}\ln L(\hat{\tau},\theta,\phi) = 0$ and $\partial_{\theta}\ln L(\hat{\tau},\theta,\phi) = 0$ of the log-likelihood (11). Each of the $\hat{\phi}_{MLE}$ and $\hat{\theta}_{MLE}$ are likely to incrementally change depending on the current value of the other in the iterations. These are nonlinear interconnected estimating equations. To ease the process, we differentiate and take then take expectation and substitute the sample variance and mean in appropriate places. For this purpose, we note that the inverse moments $E[\frac{1}{y}]$ and $E[\frac{1}{y^2}]$ need to be obtained. Note that:

$$E[\frac{1}{y}] = \int_{\tau}^{\infty}\frac{1}{y}\left(\frac{\phi}{y} + \frac{1}{\theta}\right)\left(\frac{\tau}{y}\right)^{\phi}e^{-\left(\frac{y-\tau}{\theta}\right)}dy$$

$$= \tau^{\phi}e^{\tau/\theta}\int_{1/\tau}^{0}(\phi x^{\phi+2} + \frac{x^{\phi+1}}{\theta})\partial_x(-\theta e^{-\frac{1}{x\theta}}) \quad (13)$$

$$\approx \frac{1}{\tau}(1 + \frac{\phi\theta}{\tau}) \approx \frac{2}{\tau}$$

And:

$$E[\frac{1}{y^2}] = \int_{\tau}^{\infty}\frac{1}{y^2}\left(\frac{\phi}{y} + \frac{1}{\theta}\right)\left(\frac{\tau}{y}\right)^{\phi}e^{-\left(\frac{y-\tau}{\theta}\right)}dy$$

$$= \tau^{\phi}e^{\tau/\theta}\int_{1/\tau}^{0}(\phi x^{\phi+3} + \frac{x^{\phi+2}}{\theta})\partial_x(-\theta e^{-\frac{1}{x\theta}}) \quad (14)$$

$$\approx \frac{1}{\tau^2}(1 + \frac{\phi\theta}{\tau}) \approx \frac{2}{\tau^2}$$

because of (12). To find the initial value of the $\hat{\phi}_{MLE}$, we consider a limiting scenario $\ln L(\hat{\tau},\theta,\phi\to\infty)$ of the likelihood function. After algebraic simplifications, the MLEs are then obtained and they are:

$$\hat{\phi}_{MLE} \approx \left\{\frac{1}{\bar{y}_{LM} - \ln y_{(1)}}\right\} \quad (15)$$

And:

$$\hat{\theta}_{MLE,\phi} \approx \frac{\left|s_y^2 + \bar{y}^2 - y_{(1)}^2\right|}{2\left|\bar{y} - \phi(\bar{y} - y_{(1)})\right|} \quad (16)$$

The vector of the MLE, $\begin{pmatrix}\hat{\theta}\\\hat{\phi}\end{pmatrix}$ is asymptotically (with the sample size) bivariate Gaussian distributed around the mean $\begin{pmatrix}\theta\\\phi\end{pmatrix}$ with variance-covariance matrix:

$$\Sigma = I_{2x2}^{-1} = \{E\begin{pmatrix}-\partial_{\theta\theta}^2\ln L & -\partial_{\theta\phi}^2\ln L\\-\partial_{\theta\phi}^2\ln L & -\partial_{\phi\phi}^2\ln L\end{pmatrix}\}^{-1}$$
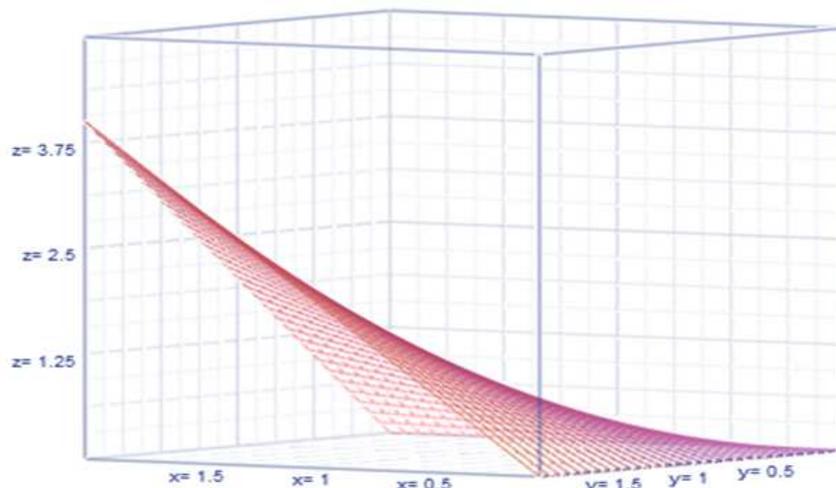
Fig. 8. Nonlinear balancing relation $\hat{\tau}$ in z-axis, $\hat{\phi}$ in x-axis and $\hat{\theta}$ in y-axis

where, $I_{2\times2}$ is the expected Fisher's information matrix, ∂_{ij}^2 is the second derivative with respect to i and j. That is:

$$I \approx n \begin{pmatrix} \dfrac{(E[y]-y_{(1)})+\theta}{\theta^3} & \dfrac{1}{y_{(1)}} \\ \dfrac{1}{y_{(1)}} & (\dfrac{\theta}{y_{(1)}})^2 \end{pmatrix} \qquad (17)$$

The $E(Y)$ is derived in (7). The determinant of the Fisher's information matrix is:

$$D \approx \left| \frac{E(y)-y_{(1)}}{\theta y_{(1)}^2} \right| \qquad (18)$$

What is hazard function for the TPF? The hazard function is the ratio of the probability function over the survival function. In the survival analysis, the hazard function is recognized as the failure rate function. The hazard function for the TPF (1) is much simpler expression:

$$h(y|\tau,\theta,\phi) = \frac{f(y|\tau,\theta,\phi)}{S(y|\tau,\theta,\phi)} = \left(\frac{\phi}{y} + \frac{1}{\theta} \right) \qquad (19)$$

See Figure 7 for the graphics of (19). The hazard function (19) is impacted by both the air pollution's severity and regulated level. See Patel and Schoenberg (2011) for details about the graphical approaches. In a scenario of higher air pollution's severity (that is, $\theta \rightarrow \infty$), the air pollution's regulated level has a full advantage marginally on the hazard (that is, $\lim_{\theta \rightarrow \infty} h(y|\tau,\theta,\phi) = \left(\frac{\phi}{y} \right)$).

In a scenario of completely controlled and reduced air pollution by the regulated (that is, $\phi \rightarrow 0$), the air pollution's severity has a full advantage marginally on the hazard function (that is, $\lim_{\phi \rightarrow 0} h(y|\tau,\theta,\phi) = \left(\frac{1}{\theta} \right)$).

To advise asthma and other respiratory patients, the healthcare professionals wonder, at times, whether an estimated air pollution's regulated level, $\hat{\phi}$ is statistically significant. This amounts to testing the null hypothesis $H_o : \phi = 0$ against an alternative hypothesis $H_1 : \phi = \phi_1 \neq 0$, where ϕ_1 is a specified value. For this purpose, the Wald's (1943) likelihood ratio approach is resorted. The Wald's approach is a powerful data analytic methodology.

In our context, to test the null hypothesis $H_o : \phi = 0$ against an alternative hypothesis $H_1 : \phi \neq 0$, we will use the log-likelihood ratio test statistic (20) below:

$$\begin{aligned} -\ln \Lambda_{\phi_0} &= \ln L(\hat{\tau}, \hat{\theta}_{\phi=\hat{\phi}}, \hat{\phi}) - \ln L(\hat{\tau}, \hat{\theta}_{\phi=\phi_0}, \phi_0) \\ &\approx n\{\ln(\frac{\hat{\theta}_{\phi=\phi_0}}{\hat{\theta}_{\phi=\hat{\phi}}}) + (\bar{y}_{AM}-y_{(1)})(\frac{1}{\hat{\theta}_{\phi=\phi_0}} - \frac{1}{\hat{\theta}_{\phi=\hat{\phi}}}) \\ &\quad + \hat{\phi}(\hat{\theta}_{\phi=\hat{\phi}}\bar{y}_{HM} - \bar{y}_{LM} + \ln y_{(1)})\} \end{aligned} \qquad (20)$$

which follows a non-central chi-squared probability distribution with one degrees of freedom (df) and the non-centrality parameter $\hat{\delta}_o = \left(\frac{\hat{\phi}}{v\hat{a}r(\hat{\phi})} \right)$, where $v\hat{a}r(\hat{\phi})$ is the MLE of diagonal element in the inverse of the variance-covariance matrix of the MLEs $\bar{y}_{AM} = \sum_{i=1}^{n} y_i / n$ $\bar{y}_{HM} = \sum_{i=1}^{n} \frac{1}{y_i} / n$ and $\bar{y}_{LM} = \sum_{i=1}^{n} \ln y_i / n$ are respectively the arithmetic, harmonic and logarithmic

means of the data (see Johnson *et al*., 1995 for details). Note that:

$$\hat{\theta}_{\hat{\phi}} \approx \frac{\left|s_y^2 + \bar{y}^2 - y_{(1)}^2\right|}{2\left|\bar{y} - \hat{\phi}(\bar{y} - y_{(1)})\right|} \qquad (21.a)$$

$$\hat{\theta}_{\phi=0} \approx \frac{\left|s_y^2 + \bar{y}^2 - y_{(1)}^2\right|}{2\bar{y}} \qquad (21.b)$$

And:

$$\hat{\phi} \approx \left|\frac{1}{\bar{y}_{LM} - \ln y_{(1)}}\right| \qquad (21.c)$$

Stuart and Ord(1994) for properties of the non-central chi squared probability distribution. Recall that the MLE, $\hat{var}(\hat{\phi})$ of variance-covariance matrix which is the inverse of the estimated Fisher's information matrix $\hat{I}_{2x2}$, in (18). Using the determinant (18) of the Fisher's information matrix I_{2x2}, the variance-covariance matrix is obtained and it is:

$$\Sigma = \begin{pmatrix} \text{var}\,\hat{\theta} & \text{cov}\{\hat{\theta},\hat{\phi}\} \\ \text{cov}\{\hat{\theta},\hat{\phi}\} & \text{var}\,\hat{\phi} \end{pmatrix}$$

$$\approx \frac{\theta y_{(1)}^2}{n\left|E(y) - y_{(1)}\right|} \begin{pmatrix} (\frac{\theta}{y_{(1)}})^2 & -\frac{1}{y_{(1)}} \\ -\frac{1}{y_{(1)}} & \frac{(E[y] - y_{(1)}) + \theta}{\theta^3} \end{pmatrix} \qquad (22.a)$$

Hence, the variance is $\text{var}\,\hat{\phi} \approx \frac{\phi}{n}\{1 + \frac{\theta}{\left|E(y) - y_{(1)}\right|}\}$.

Incidently, the correlation between the MLEs of the air pollution's severityrate and the regulated level is:

$$corr(\hat{\theta},\hat{\phi}) \approx -\sqrt{\frac{\theta}{\{\theta + \left|E(y) - y_{(1)}\right|\}}} \qquad (22.b)$$

Interestingly, the correlation (22.b) points out that when the air pollution's regulated level (that is, ϕ) is low, the air pollution's severity (that is, θ) is higher and it makes sense. Also, we notice from (22.b) that such a correlation is inversely proportional to the detected variability, $Var(\hat{\phi})$ of the estimated regulated levels, $\hat{\phi}$. In other words, the more variability among the estimated regulated levels, $\hat{\phi}$ lessens the correlation between the estimated air pollution's regulated and the severityrate.

We now return to the hypothesis testing. Under null hypothesis, $H_o : \phi = 0$, note that the noncentrality parameter is:

$$\hat{\delta}_o = \left(\frac{\hat{\phi}}{\hat{var}(\hat{\phi})}\right) \approx \frac{n}{\{1 + \frac{\hat{\theta}_{\phi=0}}{\left|\bar{y} - y_{(1)}\right|}\}}$$

where, $\hat{\theta}_{\phi=0} \approx \frac{\left|s_y^2 + \bar{y}^2 - y_{(1)}^2\right|}{2\bar{y}}$. It is known that the non-central chi squared probability distribution with one df and non-centrality parameter δ approximately follows $\left(1 + \frac{\delta}{1+\delta}\right)$ times a central chi squared probability distribution with $\left(\frac{\{1+\delta\}^2}{1+2\delta}\right)$ df (Stuart and Ord, 1994). This suggests that the null hypothesis $H_o : \phi = 0$ could be rejected in favor of an alternative hypothesis $H_o : \phi = 0$ if:

$$-\ln \Lambda_o > \left(1 + \frac{\delta_o}{1+\delta_o}\right)\chi^2_{\left(\frac{\{1+\delta\}^2}{1+2\delta}\right)df,\alpha}$$

where, the right side is the critical value based on the $100(1-\alpha)^{th}$ percentile of the central chi squared probability distribution with $\left(\frac{\{1+\delta\}^2}{1+2\delta}\right)$df and a significance effect $\alpha \in (0,1)$. We now write the p-value (23) for rejecting the null hypothesis in favor of an alternative hypothesis and it is:

$$p - value$$
$$= \Pr[\chi^2_{\left(\frac{\{1+\delta_o\}^2}{1+2\delta_o}\right)df,\alpha} > -\ln \Lambda_{\phi_0} / \left(1 + \frac{\delta_o}{1+\delta_o}\right) \qquad (23)$$

The statistical power of the test statistic (20) can be calculated with a selection of a specific value for $\phi = \phi_1$ in the alternative hypothesis. The statistical power is the probability of rejecting the null hypothesis $H_o : \phi = 0$ in favor of an alternative hypothesis $H_1 : \phi = \phi_1 = 0.5$. Under the alternative hypothesis, the minus log likelihood ratio is:

$$-\ln \Lambda_{\phi_1} \approx \ln L(\hat{\tau},\hat{\theta}_{\phi=\hat{\phi}},\hat{\phi}) - \ln L(\hat{\tau},\hat{\theta}_{\phi=\phi_1},\phi_1)$$
$$\approx n\{\ln(\frac{\hat{\theta}_{\phi=\phi_1}}{\hat{\theta}_{\phi=\hat{\phi}}}) + (\bar{y}_{AM} - y_{(1)})(\frac{1}{\hat{\theta}_{\phi=\phi_1}} - \frac{1}{\hat{\theta}_{\phi=\hat{\phi}}}) \qquad (24)$$
$$+ \hat{\phi}([\hat{\theta}_{\phi=\hat{\phi}} - \hat{\theta}_{\phi=\phi_1}]\bar{y}_{HM} - \bar{y}_{LM} + \ln y_{(1)})\}$$

The expression (24) follows a non-central chi-squared probability distribution with one df and non-centrality parameter:

$$\hat{\delta}_1 = \left(\frac{\hat{\phi} - \phi_1}{\hat{var}(\hat{\phi})}\right) \approx \frac{n\{|1 - \frac{\phi_1}{\hat{\phi}}|\}}{\{1 + \frac{\hat{\theta}_{\phi=\phi_1}}{\left|\bar{y} - y_{(1)}\right|}\}} \qquad (25a)$$

Where:

$$\hat{\theta}_{\phi=\phi_1} \approx \frac{\left| s_y^2 + \bar{y}^2 - y_{(1)}^2 \right|}{2\left| \bar{y} - \phi_1(\bar{y} - y_{(1)}) \right|} \qquad (25b)$$

The non-central chi squared probability distribution with one df and non-centrality parameter $\hat{\delta}_1$ is approximately $(1 + \frac{\hat{\delta}_1}{1+\hat{\delta}_1})$ times a central chi squared score with $\frac{(1+\hat{\delta}_1)^2}{(1+2\hat{\delta}_1)}$ df. The power is the probability of accepting a true alternative hypothesis $H_1 : \phi = \phi_1 \neq 0$. That is:

$$Power = \Pr[\chi^2_{\frac{(1+\hat{\delta}_1)^2}{(1+2\hat{\delta}_1)})df}$$

$$< \frac{\left(1 + \frac{\delta_o}{1+\delta_o}\right)(-\ln \Lambda_{\phi=\phi_1})\chi^2_{\frac{(1+\hat{\delta}_0)^2}{(1+2\hat{\delta}_0)})df,\alpha}}{\left(1 + \frac{\delta_1}{1+\delta_1}\right)(-\ln \Lambda_{\phi=o})}] \qquad (26)$$

We now turn to discuss two extreme scenarios of the above mentioned data analytics. The first scenario is one in which the air pollution severity is highest (that is, $\theta \to \infty$). The second scenario is one in which the air pollution's regulated level is smallest (that is, $\phi \to 0$).

Scenario 1 (Pareto Distribution)

When the air pollution severity rate is highest (that is, $\theta \to \infty$), the TPF (1) becomes Pareto PDF:

$$f(y|\theta \to \infty, \tau, \phi)$$

$$= \left(\frac{\phi}{y}\right)\left(\frac{\tau}{y}\right)^\phi; \qquad (27)$$

$$0 \leq \tau \leq y < \infty; \phi \geq 0;$$

as a special case with the survival function $S(y|\tau, \phi, \theta \to \infty) = \left(\frac{\tau}{y}\right)^\phi$. The mean and variance are:

$$E(y|\tau, \theta, \phi) = \frac{\phi}{(\phi-1)\tau^{\phi-1}} \qquad (28)$$

And

$$Var(y|\tau, \theta, \phi) = \frac{\phi[(\tau^\phi - 1)\phi(\phi-2) + \tau^\phi]}{(\phi-2)[(\phi-1)\tau^{(\phi-1)}]^2} \qquad (29)$$

What is Pareto distribution? Pareto (1897) introduced a power-law (later it was called Pareto distribution in his

name) to describe the differential allocation of wealth in society. It provides a better-fitting alternative when other (such as the lognormal, half-normal, exponential, Frechet) distributions fails to fit less heavy tail frequency pattern. The hope probability (4) becomes $\Pr(Y \leq m + s | Y > m) = (1 + \frac{s}{m})^{-\phi}$.

However, the safety probability (5) becomes $\Pr(Y \leq m | Y > t) = 1 - (\frac{t}{m})^\phi$.

How much memory (6) of the past air pollution pattern is kept by the nature? That is:

$$memory_{m,s,\tau,\theta,\phi} = \left(\frac{\tau}{s}\right)^\phi \left\{\left(1 + \frac{s}{m}\right)^{-\phi} - 1\right\}.$$

The MLE of the threshold parameter τ is for $\hat{\tau} = \min(y_1, y_2, .. y_n) = y_{(1)}$ the likelihood function to be non-zero. Also, the MLE is:

$$\hat{\phi}_{MLE} = \frac{n}{\sum_{i=1}^{n} \ln(\frac{y_i}{\hat{\tau}})}$$

The hazard function in (19) is $h(y|\tau, \theta, \phi) = \left(\frac{\phi}{y}\right)$.

Scenario 2 (Guaranteed Exponential Distribution):

When the air pollution's regulated level is negligible (that is, $\phi \to 0$), the TPF (1) approaches guaranteed exponential PDF:

$$f(y|\theta, \phi) = \left(\frac{1}{\theta}\right)e^{-\left(\frac{y-\tau}{\theta}\right)}; \qquad (30)$$

$$0 \leq \tau \leq y < \infty; \theta > 0$$

as a special case with the survival function $S(y|\tau, \phi \to 0, \theta) = e^{-\left(\frac{y-\tau}{\theta}\right)}$.

First, the tail value at risk ($TVaR$) in (3) of air pollution function becomes $E(Y|Y > y_p) \to 0$. The hope probability in (4) becomes $\Pr(Y \leq m + s | Y > m) = e^{-\frac{s}{\theta}}$.

Thesafety probability (5) becomes $\Pr(Y \leq m | Y > t) = 1 - e^{-\left(\frac{m-t}{\theta}\right)}$.

How much memory of the past air pollution pattern is kept by the nature in this scenario? The memory function (6) is $memory_{m,s,\tau,\theta,\phi} = e^{-\left(\frac{m-\tau}{\theta}\right)}\left\{e^{-\left(\frac{s}{\theta}\right)} - 1\right\}$.

The mean, $E(y|\tau, \theta, \phi)$ in (7) and variance, $Var(y|\tau, \theta, \phi)$ in (8) become respectively

$E(y|\tau,\theta,\phi) = \tau + Shift_{\tau,\theta,\phi}$, where the pollution level shifts from the threshold, τ is $Shift_{\tau,\theta} = 2\theta e^{\tau/\theta} \Pr(\chi^2_{2df} > \frac{2\tau}{\theta})$ depending on just the air pollution's severity θ and

$Var(y|\tau,\theta) + [E(y|\tau,\theta)]^2$

$= \tau^2 + 8\theta^2 e^{\tau/\theta} \Pr(\chi^2_{4df} > \frac{2\tau}{\theta})$.

The MLE of the threshold parameter $\hat{\tau} = \min(y_1, y_2, ..y_n) = y_{(1)}$ is for the likelihood function to be non-zero. Then, the log-likelihood function is

$\ln L(\hat{\tau}, \theta) = \sum_{i=1}^{n} \leq \left\{ \ln\left(\frac{1}{\theta}\right) - \left(\frac{y_i - y_{(1)}}{\theta}\right) \right\}$.

The MLE of the air pollution's severity is

$\hat{\theta}_{MLE, \phi \neq 0} \approx \frac{\left| s_y^2 + \bar{y}^2 - y_{(1)}^2 \right|}{2\left| \bar{y} - \hat{\phi}(\bar{y} - y_{(1)}) \right|}$.

Pollution in Africa, American, Asian, European and Oceania

In this section, the data analytic methodology of the Section 2 is illustrated using the air pollution data of cities in American, Asian and European continents. Their minimum (Min) average (Ave) numbers are displayed in Table 1. The Min and Ave of PM_{10} as well as SO_2 are more in Asia and it is a clue that the regulated is low in Asia. The American continent has higher average population in cities. A comparison of Fig. 8-10 reveals that in America and Europe, the NO_2 is more than the PM_{10} and SO_2. The variations in the levels of NO_2, PM_{10} and SO_2 are about the same in Europe. In Asia, the levels of SO_2 has a higher variation. In America, the levels of NO_2 has a higher variation. It is interesting that the Fig. 11 reveals that Africa is quite an outlier compared to the continents America, Asia, Europe and Oceania (including Australia and New Zealand) with respect to the pollution PM_{10}.

Note that higher level of SO_2 means a higher level of PM_{10} (see the correlations and their p-values in Table 2). The variations in the multivariate data are analyzed using the principal components. The first two principal components explained 66.7% of the total variations and the proximity among the variables NO_2, PM_{10} and SO_2. The NO_2 is more closely connected to the urban population. More SO_2 is associated with a higher PM_{10} (see Fig. 9). Hence, this article focuses on the analysis and interpretations of only the PM_{10}.

The minimum level $y_{(1)}$ of PM_{10} is lowest in American continent and keeps rising in other continents in the order of Oceania, Europe, Asia and Africa (see Table 3). The order of the estimated severity, $\hat{\theta}$ is lowest in Oceania and keep rising in the other continents America, Europe, Asia and Africa in that order (see Table 3). The estimated regulated level $\hat{\phi}$ of the air pollution is lowest in Asia but is better in other continents Europe, America, Africa and Oceania in that order (see Table 3). The expected shift $Shift_{\hat{\tau},\hat{\theta},\hat{\phi}}$ of the air pollution is lowest in Africa but it increases in other continents America, Oceania, Europe and Asia in that order (see Table 3). The tail value at risk (*TVaR*) of the air pollution function is $E(Y|Y > y_p)$ and it is the expected average in the tail of the frequency trend of the air pollution. The *TVaR* is lowest in America and it increases in the other continents Europe, Asia, Oceania and Africa in that order (see Table 3). The hope probability portrays the chance for the air pollution to remain within a tolerance level if it exceeded a warning level. The hope probability is best in Asia but declines in other continents Africa, Europe, America and Oceania in that order (see Table 3). If we think of nature as a chance oriented system for emitting air pollution, the system's memory is lowest in Oceania but increases in other continents Asia, Africa, America and Europe in that order (see Table 3).
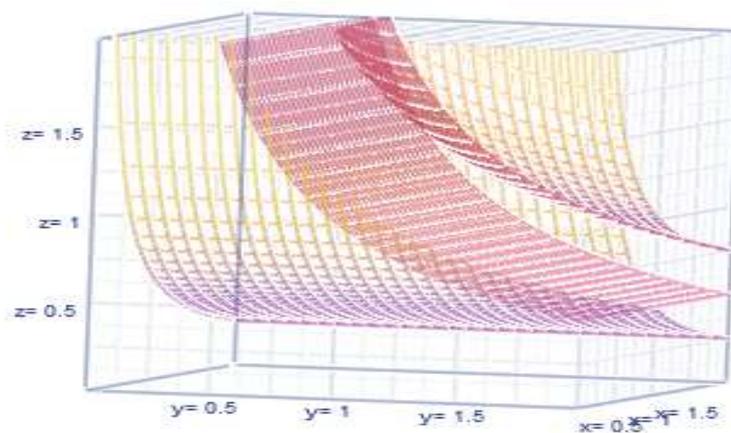
Fig. 9. Hazard function $h(y)$ in z-axis, ϕ in x-axis, and θ in y-axis
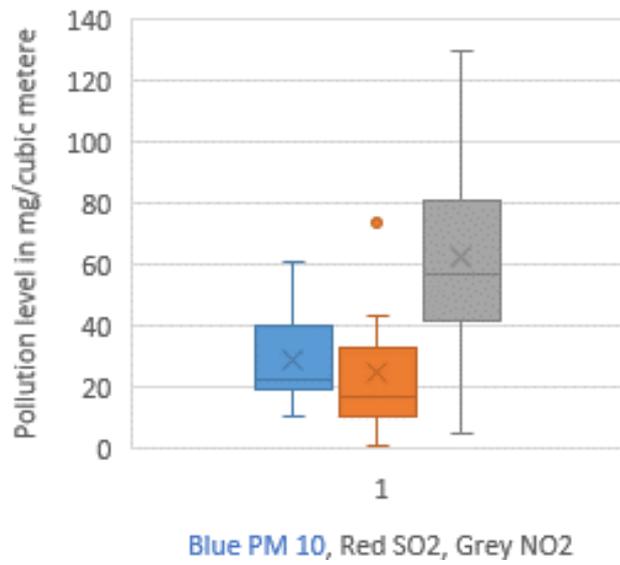
Fig. 10. American cities
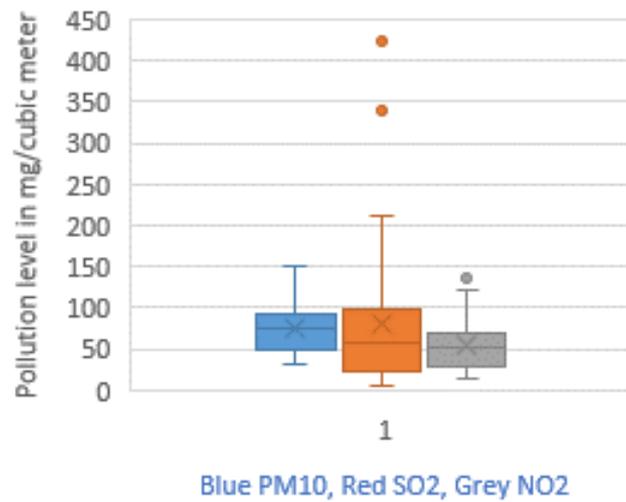
Fig. 11. Asian cities
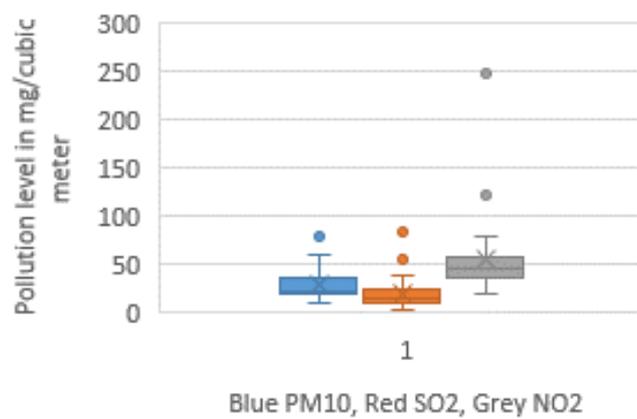
Fig. 12. European cities
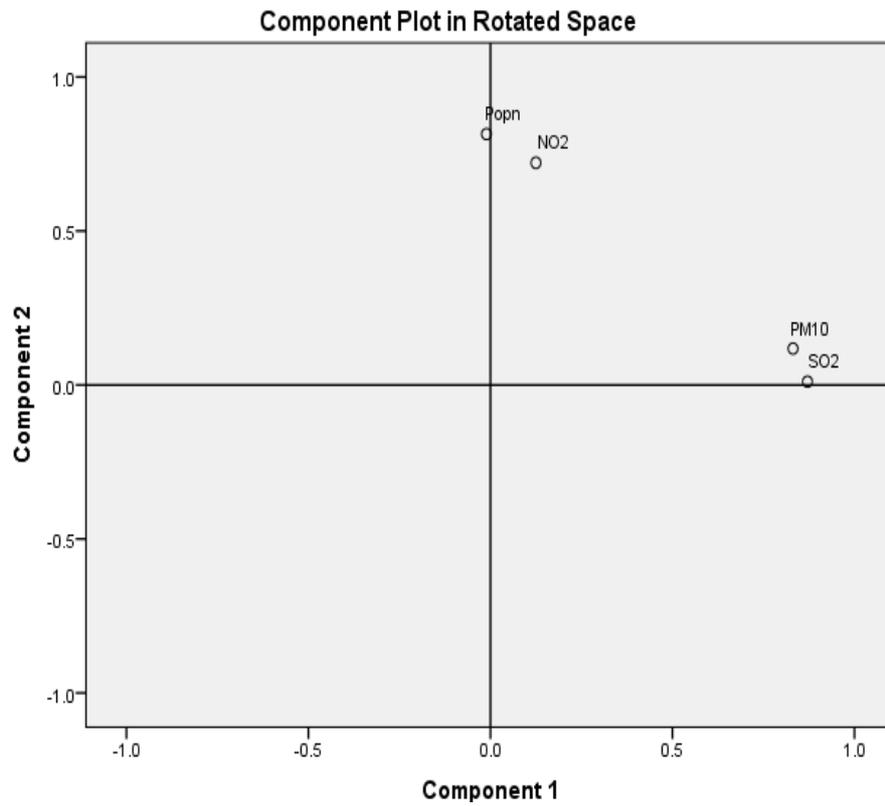
Fig. 13. Proximities among PM_{10}, SO_2 and NO_2
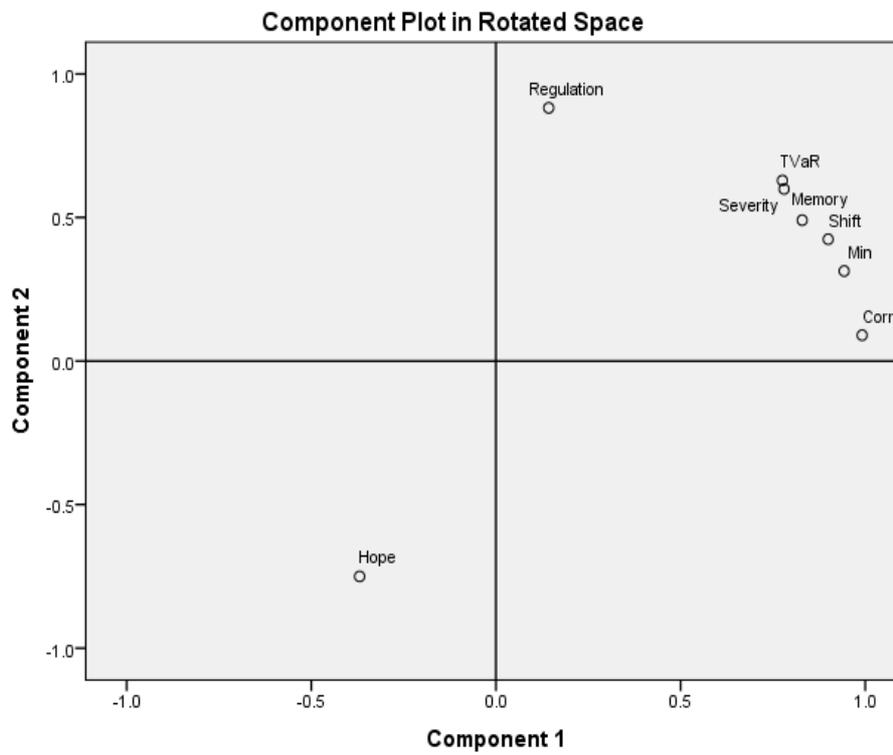
Fig. 14. Proximity among the characteristic values

25
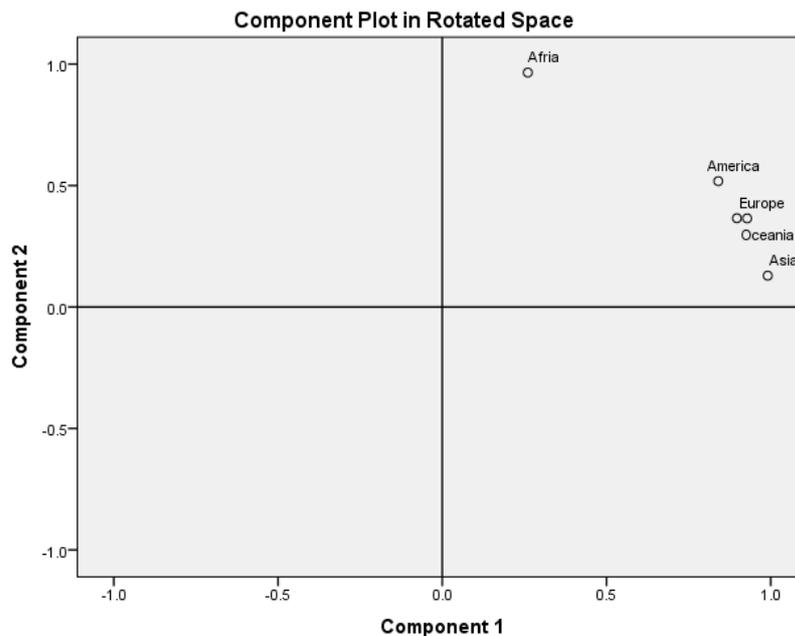
Fig. 15. Proximity among the continents in terms of PM_{10}

Table 3. A comparison of $PM10$ related matters among continents: Africa, America, Asia, Europe and Oceania

$\dfrac{\text{continent} \rightarrow}{\text{Attribute} \downarrow}$	Africa (17 cities)	America (568 cities)	Asia (396 cities)	Europe (511 cities)	Oceania (32 cities)
$\hat{\tau} = y_{(1)}$	20	6	10	9	8
Severity $\hat{\theta}$	159.79	40.06	116.89	51.50	34.54
regulated $\hat{\phi}$	0.977	0.904	0.504	0.83	2.07
$Shift_{\tau,\theta,\phi}$	0.0001	0.02	14.54	13.25	11.63
$TVaR$ $E(Y\|Y>20)$	177.09	59.68	69.93	63.24	127.72
Hope probability $Pr(Y \leq 30\|Y>20)$	0.63	0.54	0.75	0.59	0.32
$memory_{m=20,s=10,\tau,\theta,\phi}$	0.13	0.13	0.12	0.29	0.06
$corr(\hat{\theta},\hat{\phi})$	-0.87	-0.84	-0.77	-0.82	-0.92
p-value for $H_o:\phi=0$	0.004	0.0001	0.0001	0.98	0.0002
Power for $H_1:\phi=0.5$	0.001	0.0001	0.999	0.98	0.564

Remember that the estimated severity of the air pollution is inversely proportional to the estimated regulated level of the air pollution. This inverse relation is exhibited in their negative correlation. The estimated negative correlation is lowest in Oceania but increases in other continents Africa, America, Europe and Asia in that order (see Table 3). The null hypothesis $H_0:\phi=0$ portrays the insignificant regulated level of the air pollution and the p-value chance of rejecting the null hypothesis. The p-value is highest only in Europe but is smaller in all other continents Oceania, Africa, America and Asia. The power is the probability of accepting the alternative hypothesis $H_1:\phi=\phi_1$ in a hypothetical scenario in which the true value of the regulated level of

the air pollution is $\phi_1=0.5$. The power is best in Asia and Europe, is reasonable in Oceania but is worst in African and American continents (see Table 3).

The data analytics based statistical results identify the geometric distance among the results regulated, severity, $TVaR$, memory, shift, minimum and correlation levels of the continents (see Fig. 14) and it confirms the importance of the regulated level in the discussion of air pollution. Consequently, some continents are in close proximity than others with respect to air pollution (see Figure 13, Figure 14 and Figure 15). African continent is alone as an outlier. The American continent is in the opposite spectrum of the Asian continent, while in the middle are the European and Oceania with very similar in the air pollution (see Fig. 15).

Comments and Conclusions

This article has developed new statistical properties of the tapered probability function and demonstrated their utility to analyze and interpret air pollution trend on PM_{10}. The derived analytical expressions helped to compare Africa, America, Asia, Europe and Oceanic continents. In particular, the extreme occurrences of the PM_{10} in these continents were captured and explained using the tail value at risk (*TVaR*), a probabilistic measures named hope probability and safety probability, stochastic measure of system's memory level, how much the pollution has shifted on the average and how it controls the volatility of the pollution etc. The maximum likelihood estimators for the severity level, a threshold level and regulated level of the air pollution are obtained and applied in the data to capture and interpret the air pollution proximities among the continents. An expression for the hazard function of the air pollution is derived and demonstrated for all the continents (see Fig. 9).

An expression to quantify the correlation between the estimated the severity level and regulated level of the air pollution has been derived and explained for all the continents. The locally most powerful likelihood ratio test methodology has been developed to examine whether the statistical estimate of the air pollution's regulated level in a continent is significant and to estimate the statistical power of the test in an event having the true regulated level. With all these new expressions, the proximities among the continents with respect to the air pollution. Their graphical illustrations portray the importance of the computed results for the continents. Only the African continent is seen to be outlier compared to the other continents. Special expressions are identified for the Pareto and guaranteed exponential distributions as particular cases of our general results in the article.

All these add to our better understanding the air pollution in cities across the continents. This understanding will assist governing agencies to formulate and implement regulated policies. The environmental professionals can easily mimic the approach of this article in their practice of controlling the air pollution in their jurisdiction. The healthcare professional are helped by the analytic results to better prepare to deal with the health impacts of worst air pollution before it ever occurs. Needless to mention is that the residents who have been suffering in major cities with severe air pollution are helped by the approach and methodology of this article eventually.

Acknowledgment

Author's Contributions

The contents of this article are the opinion and thoughts of the author.

Ethics

There is no violation of ethical standards in the contents.

References

Abramowitz, M. and I. Stegun, 1972. Handbook of mathematical functions. U. S. Department of Commerce, National Bureau of Standards.

Counter, S.A. and L.H. Buchanan, 2004. Mercury exposure in children: A review. Toxicol. Appl. Pharmacol., 15: 209-230.

Dominici, F., J. Samet and S. Zeger, 2000. Combining evidence on air pollution and daily mortality from the largest 20 US cities: A hierarchical modeling strategy. J. Royal Statistical Society (Series A), 163: 263-302. DOI: 10.1111/1467-985X.00170

Johnson, N.L., S. Kotz and N. Balakrishnan, 1995. Continuous Univariate Distributions, 2nd Edn., Wiley, NY.

Kagan, Y.Y. and F. Schoenberg, 2001. Estimation of the upper cutoff parameter for the tapered Pareto distribution. J. Applied Probab., 38A: 158-175. DOI: 10.1239/jap/1085496599

Louwies, T., L. Int Panis, M. Kicinski, P. De Boever and T.S. Nawrot, 2013. Retinal microvascular responses to short-term changes in particulate air pollution in healthy adults. Environ. Health Perspectives, 121: 1011-6. DOI: 0.1289/ehp.1205721

Malley, C.S., J.C. Kuylenstierna, H.W. Vallack, D.K. Henze and H. Blencowe *et al*., 2017. Preterm birth associated with maternal fine particulate matter exposure: A global, regional and national assessment. Environ. Int., 101: 173-182. DOI: 10.1016/j.envint.2017.01.023

Mateen, F.J. and R.D. Brook, 2011. Air pollution as an emerging global risk factor for stroke. J. Am. Medical Association, 305: 1240-1241. DOI: 10.1001/jama.2011.352

Pareto, V., 1897. Cours d'Économie Politique, Tome Second, Lausanne, F. Rouge, quoted by V. Pareto, 1964, Œuvres Complètes, Publ. by de Giovanni Busino, Genève, Droz, v. II.

Patel, R.D. and F.P. Schoenberg, 2011. Graphical test for local self-similarity in Univariate data. J. Applied Statistics, 38: 2547-2562. DOI: 10.1080/02664763.2011.559211

Pope, C.A., 2002. Cancer, cardiopulmonary mortality and long-term exposure to fine particulate air pollution. J. Am. Med. Assoc., 287: 1132-1141. DOI: 10.1001/jama.287.9.1132

Pope, C.A., A. Bhatnagar, J.P. McCracken, W.T. Abplanalp and D.J. Conklin *et al*., 2016. Exposure to fine particulate air pollution is associated with endothelial injury and systemic inflammation. Circulation Res., 119: 1204-1214.
DOI: 10.1161/CIRCRESAHA.116.309279

Shanmugam, R. and A. Hertelendy, 2011. Do Developing or Developed Nations Polluteair More? An Assessment of Health Consequences. In: The Impact of Air Pollution on Health, Economy, Environment and Agricultural Sources, M.K. Khallaf (Eds.). In Tech Publisher, Croatio. pp: 423-444.

Wald, A., 1943. Tests of statistical hypotheses concerning several parameters when the number of observations is large. Trans. Am. Math Soc., 54: 426-482.
DOI: 10.1090/S0002-9947-1943-0012401-3

Weinmayr, G., E. Romeo, M. De Sario, S.K. Weiland and F. Forastiere, 2010. Short-term effects of PM_{10} and NO_2 on respiratory health among children with asthma or asthma-like symptoms: A systematic review and Meta-Analysis. Environ. Health Perspectives, 118: 449-457.
DOI: 10.1289/ehp.0900844