

Analysis of Climatic Structure with Karachi Dengue Outbreak

Syed Afrozuddin Ahmed^{1,*}, Junaid Saghir Siddiqi¹, Sabah Quaiser¹ and Afaq Ahmed Siddiqui²

¹Department of Statistics, University of Karachi, Karachi, Pakistan

²Department of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Karachi, Karachi, Pakistan

Abstract: Various studies have reported that global warming causes unstable climate and serious impact on physical environment and public health. The increasing incidence of dengue case is now a priority health issue and has become a health burden for Pakistan. In this study it has been investigated that spatial pattern of environment causes the emergence or increasing rate of dengue fever incidence that effects the population and its health. The climatic or environmental and the Dengue Fever (DF) case data was processed by coding, editing, tabulating, recoding and restructuring and finally applying different statistical methods, techniques and procedures for the analysis and interpretation. Five climatic variables which we have studied are precipitation (**P**), Maximum temperature (**Mx**), Minimum temperature (**Mn**), Humidity (**H**) and Wind speed (**W**) collected from 1980-2012. The data on Dengue Fever cases in Karachi for the period 2010 to 2012 are available and reported on weekly basis. Principal Component 1 (PC1) for all groups of the period can be interpreted as the General atmospheric condition. PC2 the second important climate factor for dengue period (2010-2012) comes out contrast between precipitation and wind speed. PC3 is the weighted difference between maximum temperature and wind speed. PC4 is the contrast between maximum and wind speed. Negative Binomial and Poisson regression model are used to correlate the dengue fever incidence to climatic variable and principal component (PC) score. Due to the problems of over dispersion the Poisson models are not useful for interpretation through Negative Binomial model we found that relative humidity causes an increase on the chances of dengue occurrence by 1.71% times. While maximum temperature positively influence on the chances dengue occurrence by 19.48% times. Minimum temperature affects on the chances of dengue occurrence by 11.51% times. Wind speed is effecting negatively on the weekly occurrence of dengue fever by 7.41%times.

Keywords: Principal component analysis, Dengue Fever, Negative Binomial Regression model, Poisson Regression model.

1. INTRODUCTION

The word of Dengue comes from the Swahili phrase meaning the devil spirit. Record shows that this fever was first mentioned from the Jin Dynasty (265-420 A.D) of Chinese medical encyclopedia meaning take it as poison from water coming from the flying insect. It is also recognized that initial epidemics came in Asia, Africa and North America in 1779 – 1780 and after this diseases was reported in Philippines and Thailand in 1950s [1]. Now this viral disease has spread in most parts of the world [2]. The viral disease which is called dengue is, transmitted to human body through the mosquito (vector) named *Aedes aegypti* [3]. In most of the tropical and sub tropical countries dengue fever has become a serious and dangerous health risk for the public. About two million people are affected by dengue fever in the world [2]. “Den-1, Den-2, Den-3 and Den-4” are used the names of serotypes which causes severe dengue fever [4,5]. In the medical terminology its other name is “break bone fever” and high grade fever, flushing of face, headache, pain in joints (arthralgia) and muscle, (myalgia) headache, rashes in skin (skin erythema) close to the measles, generalized body ache

are its symptoms. Sometime this disease is life threatening as in blood disorder causes the lack of increased Capillary Permeability of platelets and shortage of plasma due to these problems dengue hemorrhagic shock syndrome may be observed. Patients also suffer from high blood pressure. No vaccines are still commercially available for this disease; however prevention is the way through which we may reduce the mosquito habitat and exposure to bites.

The maturity period of mosquito from laying egg to transfer the diseases in human body is approximately four weeks time it leads to outbreak of dengue fever within 7 to 45 days. The maturing period of larvae to adult is about 7 to 12 days. Female mosquitoes take 8 to 15 days to complete life cycle [2].

Dengue viruses are transmitted by mosquitoes called *Aedes Aegypti*, which are highly sensitive to environmental conditions. The climatic variables i.e., precipitation, humidity, temperature and wind speed help in the survival and reproduction of dengue mosquitoes. High temperature helps to decrease the time lag for the replicating of the dengue mosquito virus. This process is called “Extrinsic Incubation Period”, in this period virus is transferred from mosquito to human body through the mosquito’s salivary glands. In the high temperature mosquito become infectious

*Address correspondence to this author at the Department of Statistics, University of Karachi, Karachi, Pakistan; E-mail: safrozuddin@yahoo.com

faster and has higher probability to dangerous for human being before completing its life cycle. Climatic change and environmental factors are helping to flourish the dengue vector but there are some other factors which also play a vital role for the dengue outbreak. In the hot weather incubation time of virus is lesser, which indicates that there is low probability of mosquito to become infectious. Due to global warming mosquito is increasing. So we may say that the critical factor is climate by which the situation may be controlled the mosquito habitat. Due to variation in climate mosquito moves to a suitable habitat. The relation between climatic variable and dengue was studied from 1900 and the researchers have established the links between the climatic variables and the dengue fever. Different studies show that there is a strong association between dengue incidence and rainfall. Researchers [6-13] have studied and reported that there is a positive correlation between the amount of rainfall and relative humidity with dengue occurrence. Similar results have also reported by other study groups [14-15] regarding the correlation between rainfall and infection rate of dengue [16-17] have shown the relationship between climatic factors and dengue fever incidence to indicate or predict variation in dengue incidence [18] have assessing the effects of temperature and precipitation on dengue transmission relation has studied the dengue epidemics and its association with precipitation and temperature.

In Karachi, Pakistan first Dengue fever outbreak was reported in 1994 [19] and one patient out of 145 died. Many researchers have studied and reported epidemics of dengue and its causes from different parts of Pakistan [20-21]. Pakistan has experience a number of dengue fever outbreaks since 1994. In October 2003 dengue outbreaks were detected for the first time in sub-mountainous areas of district Haripur, Khyber Pakhtoonkhawa province and district Khushab, Punjab Province claiming 6 lives among 717 cases. In October 2005, Dengue again hit Karachi after 10 years and 21 deaths out of total 103 confirmed cases were recorded. Since then, the disease has become widely accepted as one of the major public health problems. Over 21,204 people were reportedly infected in the country in 2010. The reasons for this problem are poor hygiene and inefficient sewerage system which provides an ideal habitat for the dengue vector to lay eggs and flourish. For the last few years transmission of dengue virus situation has been high in the country, Karachi and Lahore are the cities which are heavily threatened by dengue epidemics in pre and post monsoon periods [21-22]. In the periods of flood this adverse situation may also be observed. The internally displaced

peoples who migrated from the arm conflicted areas and they were living without shelter and proper healthy environment suffered with the dengue and other viral diseases [23].

2. METHODOLOGY

This study was conducted in Karachi capital of Sindh province, it is the largest and most populous metropolitan city of Pakistan and its main seaport and financial centre. The city has an estimated population of 23.5 million people as of April 2013[21] and a density of nearly 6,000 people per square kilometer (15,500 per square mile) and is the 3rd-largest city in the world by population. The climate of Karachi features an arid climate, albeit a moderate version of this climate. Karachi is located on the coast and as a result has a relatively mild climate. Karachi has two main seasons; Summer & Winter, while spring and autumn are very short [22]. The climate data of Karachi were collected from Pakistan Meteorologist Department, on monthly basis from 1980 -2012 and daily basis from 2010-2012. Names of five climatic variables which we have studied are precipitation (**P**), Maximum temperature (**Mx**), Minimum temperature (**Mn**), Humidity (**H**) and Wind speed (**W**).

The dengue cases in Karachi from 2010 to 2012 are reported on weekly basis. The data on the number of deaths due to Dengue fever are not available with us; therefore our study will only focus on the dengue occurrence cases only. In 2010 Govt. of Sindh, Pakistan established the Dengue surveillance cell for not only keeping the records of Dengue Fever cases, but also providing the information about the DF and the health facilities to the people in this regard.

The dependent and explanatory variables along with the corresponding codes which are used to express these variables in equations and unit of measurement are listed in the Table 1.

Table 1: Climate variables, Dengue Fever their Labels and their Unit of Measurement

S.No	Variables	Code	Unit of measurement
1	Precipitation	P	mm
2	Relative humidity	H	%
3	Maximum temperature	Mx	°C
4	Minimum temperature	Mn	°C
5	Wind speed	W	Knots
6	Dengue Fever	DF	

2.A. PCA to Explore Karachi's Climatic Structure

Principal Component Analysis (PCA) is used to obtain the climatic structure through climatic variables. PCA technique is recognized as reduction and extraction for dimensionality of the data and rating as much of the variation present in the original data set [24-25]. This is one way of identifying patterns in the data [24]. It is difficult to find patterns in the data for this purpose we use PCA which highlights the similarities and difference in the data. The pattern in the data can be found by squeeze the data in other word by suppressing the dimensionality of the data by avoiding the losses of information. This procedure is also use in the image compression or image reorganization. The purpose to apply the PCA is to reduce the manifest variables, in this way the set of components will be reduced [26]. The components are called PC1, PC2, PC3 and so on, (for the first, second and third principal components) are independent of each other then extract the decrease the amount of variance from the original data set. PC1 (the first component) captures most of the variance, PC2 captures the second most of the variance and so on until all the variance is accounted for, in this way very few will be retain for the further study [27].

The components have been treated as climatic factors or climatic structures. The Principle component method transforms correlate observed variables in to uncorrelated variables which, are linear combination of observed data. The required condition to apply this technique is correlation/covariance can be defined.

For the purpose of statistical studies, we have applied PCA and other statistical techniques also fitting different statistical models to find the relations between DF and climatic variables [28-29] have also discussed the relation between climatic variable and dengue incidence.

For the purpose of statistical analysis we have taken the monthly and weekly data dengue period 2010-2012. The reason we have considered the data of

the period 2010-2012 is that dengue data is available for this period only.

We have applied Principal Component Analysis technique for monthly and weekly data of 2010-2012 to explore if there is any difference in the structure of climate in this time period [30-31] and have also applied Principal Component Analysis technique for the characterization of climatic variables.

2.B. Modeling DF with Climatic Variables with Different Lags: Weekly Data

We have applied the negative binomial and Poisson regression models on Karachi weekly climatic variables and climatic factors data of 2010-2012. In both negative binomial and Poisson model the dependent variable is DF cases. Poisson model is found to be not suitable because of very large (high) over dispersion [32] and their AIC's values are very large varying from 45 to 48, so we will not consider it. We found negative binomial models much better than Poisson models. The AIC values in case of negative binomial models are very low as compare to AIC value of Poisson models, the AIC values from 7.44 to 7.51. The values of AIC decreases as lag period increases. We have shown the estimation of Poisson parameter for the purpose of comparison.

3. RESULTS

In Table 2 we presented number of dengue fever cases by months, for the years 2010, 2011 and 2012. The number of cases repeated for each succeeding year has drastically reduced. The number of dengue fever cases repeated in 2012 is only $734/5249 \times 100 \approx 14\%$ of the 2010. This indicates that Government of Sindh has tried to facilitate the people against dengue fever in terms of awareness of DF, health facilities, import medicine and also expert teams to work together with local teams.

Figures 1 and 2 shows the weekly and monthly dengue cases. It can be clearly seen from Figures 1 and 2 that the curve of monthly data is only the

Table 2: No. of Dengue Cases Reported in Karachi (2010 –2012)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
2010	11	3	14	7	7	12	25	224	128	2510	1902	406	5249
2011	18	5	8	4	17	8	32	40	393	415	104	35	1079
2012	11	14	8	12	16	28	39	40	6	141	262	102	734

Source: Dengue surveillance cell Karachi, Govt. of Sindh, Pakistan.

smoothed form weekly data curve. The high peaks of curve Figure 2 give the same picture as we have described in the earliest paragraph.

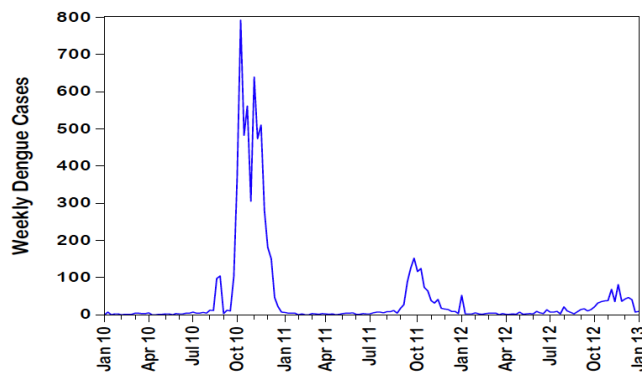


Figure 1: Karachi Weekly Dengue Cases 2010-2012.

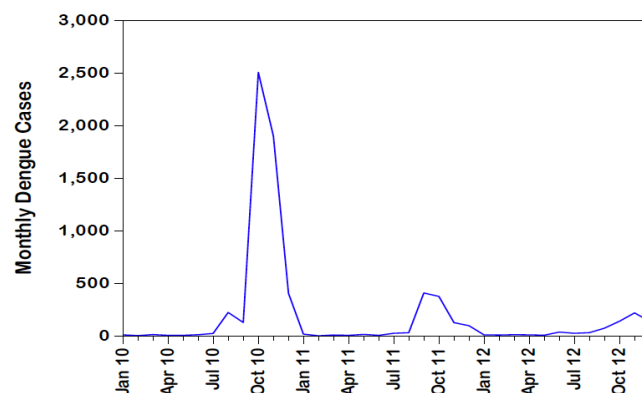


Figure 2: Karachi Monthly Dengue Cases 2010-2012.

3.A. PCA to Explore Karachi's Climatic Structure

For explanation and understanding purposes we present all four PCs for considered period in Table 3. This table contain, in columns; the linear combination (PCs) in approximate and simplified form, percentage of variation explained by respective Principle components and abbreviation of labels of the PC.

According to the elbow rule for scree plot all plots (see Appendix-I) show that two Factor model (PC) is

sufficient to explain the variation in the data. But for the comments and interpretation we consider four PCs: as the percentage explained by 3rd and 4th PC in some case are quite large. The small variation in weather may affect the general atmospheric condition can affect life.

The first four PCs for the monthly and weekly data are almost the same. For weekly dengue period data the percentage of variance of PC1 and PC2 are highly different to each other, they are around 44% and 22%, which shows that there are only 66% of total variation factor of Karachi climate is namely PC1 and PC2 variation in temperature with humidity and wind speed and contradiction of wetness and temperature. Explanation and interpretation of PCs one by one, for all data sets, shown in Table 3.

PC1s for both monthly and weekly data are being constructed as the linear combination of Humidity, Maximum temperature, Minimum temperature and Wind speed, in other words it is the weighted average of these four climatic variables and all the four variables are related to General Atmospheric Condition (GAC). For the reader to understand as to how we have interpreted and labeled the PCs, let us take the example of first PC of the monthly data. It is (PC1) weighted average $[PC1= 0.9(H+Mx)+Mn+W]$ of the Humidity (H), Maximum temperature (Mx), Minimum temperature (Mn) and Wind speed (W). Thus it may be interpreted that major source of variation in the data occurs due to "General Atmospheric Condition". Therefore we label it or name it as General Atmospheric Condition GAC. This PC (GAC) contributes/explain (65.6%) of total variation in the observed data.

PC2 of monthly data can be label as contrast between precipitation and wind speed and its contribution in terms of explanation of variation is

Table 3: PCA for Monthly and Weekly Climatic Variables: (2010 – 2012)

PCs	2010 -2012(Monthly Dengue Period)			2010 -2012 (weekly Dengue Period)		
	Linear combination	Label	%	Linear combination	Label	%
PC1	$0.9H+0.9Mx+Mn+0.7W$	GAC	65.6	$0.6H+0.7Mx+0.8Mn+0.8W$	GAC	44
PC2	$0.9P-0.5W$	CPW	21.8	$0.5P+0.7H-0.4Mx-0.4Mn$	C[PH][T]	22
PC3	$0.5W-0.5Mx$	CWT	10	$0.7P-0.4H+W$	C[P][WH]	18
PC4	$0.2P+0.1Mx-0.3$	C[PT][H]	2.4	$0.5Mx-0.4W$	CTW	11

Where: **GAC** is General Atmospheric condition. **CTP** is Contrast between Temperature & Precipitation. **CTW** is Contrast between Temperature & Wind speed. **CWH** is Contrast between Wind speed & Relative Humidity. **CPW** is Contrast between Precipitation & Wind speed. **CWT** is Contrast between Wind speed & Temperature. **C[PT][H]** is Contrast between the combination of Precipitation and Temperature with Humidity. **C[PH][T]** is Contrast between the combination of Precipitation & Humidity with Temperature. **C[P][WH]** is Contrast between of Precipitation and the combination with Wind speed & Humidity .

21.8%. This contribution may consider as large contribution. Thus we may conclude that this factor plays an important role in the formation of weather of city Karachi.

PC3: The third component in the variation of Karachi climate is “opposite of variation of wetness and wind speed “which constitutes 18% of variation. Whereas the fourth component is the contradiction of “temperature” and “wind speed “which is 11.2% of variation.

PC4 for Dengue (Monthly) period is constructed as contrast of humidity against weighted average of precipitation and maximum temperature, which contributes only 2.4 of the total variation. PC4 for the weekly data is the contrast between Mx and W and it explaining 11.2% of variation.

3.B. Modeling DF with Climatic Variables with Different Lags: Weekly Data

The following expressions are established from negative binomial models from the Tables 4 and 5. In these equations we have considered only those variables whose coefficient are significant at least $\alpha = 0.05$.

$$\begin{aligned} \text{Log (DF)} &= 0.018H + 0.118Mx - 0.115W\text{Lag} & (-1) [7.513] & \text{KNBCV(-1)} \\ \text{Log (DF)} &= 0.022H + 0.139Mx - 0.136W\text{Lag} & (-2) [7.464] & \text{KNBCV(-2)} \\ \text{Log (DF)} &= -4.183 + 0.021H + 0.163Mx + 0.107Mn - 0.123W\text{Lag} & (-3) [7.441] & \text{KNBCV(-3)} \\ \text{Log (DF)} &= -5.1882 + 0.017H + 0.178Mx + 0.109Mn - 0.077W\text{Lag} & (-4) [7.443] & \text{KNBCV(-4)} \end{aligned}$$

$$\begin{aligned} \log(\text{DF}) &= -5.1882 + 0.017(H) + 0.178(Mx) + 0.109(Mn) - 0.077(W) \\ \text{DF} &= (e^{-5.1882}) (e^{0.017H}) (e^{0.178Mx}) (e^{0.109Mn}) (e^{-0.077W}) \end{aligned}$$

We consider negative binomial climatic variables of Karachi [KNBCV (-3)] model as the best predictive model for number of DF cases because its AIC value is the smallest (AIC=7.441) among the above 4 models. But the AIC value at lag three and lag four are almost equal as four weeks lag is equivalent to one month and the variables involved in [KNBCV (-3)] & [KNBCV (-4)] are same. Therefore we are going to consider [KNBCV (-4)]. In all the four lag model relative humidity, maximum temperature and wind speed shows significant effect on the occurrence of ‘DF’ while wind speed ‘W’ has significant effect on prevention of ‘DF’.

3.C. Model with Climatic Factors with Different Lags: Weekly Data

We have applied the Negative Binomial and Poisson regression models on Karachi weekly climatic

Table 4: Negative Binomial and Poisson Regression Models for Karachi Weekly Climatic Variables and PCS Score with Lags (1=1 Week, 2=2 Week)

C.V	PCA	Lag 1=1 week				Lag 2=2week			
		Poisson		Negative Binomial		Poisson		Negative binomial	
		KPCV(-1)	KPPC(-1)	KNBCV(-1)	KNBPC(-1)	KPCV(-2)	KPPC(-2)	KNBCV(-2)	KNBPC(-2)
Constant	Constant	C.V	PCA	C.V	PCA	C.V	PCA	C.V	PCA
P-value	P-value	(0.0004)	(0.0000)	(0.4957)	(0.0000)	(0.0000)	(0.0000)	(0.0944)	(0.0000)
Rainfall	PC1	-0.000960*	0.073784*	0.004614	-0.006991	-0.002222*	0.246229*	0.006360	0.195029
P-value	P-value	(0.2009)	(0.0000)	(0.6073)	(0.9655)	(0.0025)	(0.0000)	(0.4753)	1.221883
Humidity	PC2	0.025574*	0.208700*	0.018027*	0.121896	0.028595*	0.270955*	0.022406*	(0.2218)
P-value	P-value	(0.0000)	(0.0000)	(0.0048)	(0.4056)	(0.0000)	(0.0000)	(0.0004)	(0.0000)
Max-Temp	PC3	0.092952*	-0.093486*	0.118338*	-0.062715	0.083167*	-0.031850*	0.139797*	0.090866
P-value	P-value	(0.0000)	(0.0000)	(0.0470)	(0.6283)	(0.0000)	(0.0083)	(0.0166)	0.615796
Min-Temp	PC4	0.063665*	1.103704*	0.042956	0.809898*	0.104455*	1.068109*	0.079271	0.080677
P-value	P-value	(0.0000)	(0.0000)	(0.3055)	(0.0000)	(0.0000)	(0.0000)	(0.0501)	0.579903
Windspeed		-0.144869*		-0.114828*		-0.36728*		-0.135944*	0.926157*
P-value		(0.0000)		(0.0048)		(0.0000)		(0.0004)	7.677492
DF		150	151	151	154	148	150	151	153
AIC		48.59230	104.1927	7.512931	8.388649	46.94487	107.4916	7.463632	8.366574

factors for the Dengue period 2010-2012 complete results (output) are given in Tables 4 and 5. Below we write the models negative binomial considering those variables whose coefficients are significant with at least $\alpha = 0.05$. Also apart from lag differences, AIC values are given. These expressions are obtained from Tables 4 and 5. Principle components analyses are applied and components are extracted and available in Table 3 and the interpretations are made in earlier section. But in the expressions we have used the notation in term of PCs for the reason of simplicity.

$$\begin{aligned} \text{Log (DF)} &= 3.421 + 0.81 \text{ PC4} & \text{Lag (-1) [8.388]} & \text{KNBPC(-1)} \\ \text{Log (DF)} &= 3.369 + 0.9262 \text{ PC4} & \text{Lag (-2) [8.366]} & \text{KNBPC (-2)} \\ \text{Log (DF)} &= 3.257 + 0.452\text{PC1}+1.003 \text{ PC4} & \text{Lag (-3) [8.310]} & \text{KNBPC (-3)} \\ \text{Log (DF)} &= 3.228 + 0.792 \text{ PC1} + 0.927\text{PC4} & \text{Lag (-4) [8.308]} & \text{KNBPC (-4)} \end{aligned}$$

$$\begin{aligned} \log(\text{DF}) &= -3.228 + 0.792(\text{PC1}) + 0.927(\text{PC4}) \quad \text{--- (A)} \\ \text{DF} &= (e^{-3.228}) (e^{0.792\text{PC1}}) (e^{0.927\text{PC4}}) \end{aligned}$$

4. DISCUSSION

In this study our focus was on exploring the climatic variables and/or the climatic factors (structure) which may influence the increase or decrease in the number of dengue fever cases in Karachi.

The main effecting explanatory variables are humidity, maximum temperature, minimum temperature

and wind speed, where humidity, maximum temperature and minimum temperature effecting positively on the occurrence of dengue fever cases, while the wind speed seems to be causing in the reduction of dengue fever cases. On the basis of the model expressed above following are some comments.

- 1) Weekly relative humidity is effecting the chance of occurrence positively and in the past 4 weeks (28 days) humidity value is estimated to positively influence on the chances of dengue occurrence by 1.71% times.
- 2) Weekly maximum temperature is effecting to increase the chance of occurrence and in the past 4 weeks (28 days) maximum temperature value is estimated to positively influence the chances of dengue occurrence by 19.48% times.
- 3) Minimum temperature is effecting the chance of occurrence DF cases positively for the weekly data. For the past 4 weeks the minimum temperature affects positively on the chances of dengue occurrence by 11.51% times.
- 4) Wind speed is effecting negatively on the weekly occurrence of dengue fever by 7.41%.

[30] Have found minimum temperature having positive significant effect while wind speed and

Table 5: Negative Binomial and Poisson Regression Models on Karachi Weekly Climatic Variables and PCS Score with Lags (3=3 Week, 4=4 Week)

C.V	PCA	Lag 3=3 week				Lag 4=4week			
		Poisson		Negative Binomial		Poisson		Negative binomial	
		KPCV(-3)	KPPC(-3)	KNBCV(-3)	KNBPC(-3)	KPCV(-4)	KPPC(-4)	KNBCV(-4)	KNBPC(-4)
		C.V	PCA	C.V	PCA	C.V	PCA	C.V	PCA
Constant	Constant	2.121113*	3.195835*	-4.183118*	3.256533*	-2.843588*	3.246479*	-5.188223*	3.227487*
P-value	P-value	(0.0000)	(0.0000)	(0.0050)	(0.0000)	(0.0000)	(0.0000)	(0.0003)	26.53814 (0.0000)
Rainfall	PC1	-0.003240*	0.455423*	0.007874	0.451676*	-0.004586*	0.611525*	0.008183	0.791778*
P-value	P-value	(0.0000)	(0.0000)	(0.3774)	(0.0028)	(0.0000)	(0.0000)	(0.3378)	5.166530 (0.0000)
Humidity	PC2	0.030179*	0.259762*	0.020994*	0.036624	0.029054*	0.333114*	0.017063*	0.074926
P-value	P-value	(0.0000)	(0.0000)	(0.0005)	(0.8076)	(0.0000)	(0.0000)	(0.0044)	0.477296 (0.6332)
Max-Temp	PC3	0.077762*	-0.009479	0.163378*	0.212483	0.073605*	-0.064236*	0.177967*	0.289465
P-value	P-value	(0.0000)	(0.4267)	(0.0034)	(0.1682)	(0.0000)	(0.0000)	(0.0011)	1.660872 (0.0967)
Min-Temp	PC4	0.144178*	1.36116*	0.107020*	1.003453*	0.169875*	1.022751*	0.109589*	0.926595*
P-value	P-value	(0.0000)	(0.0000)	(0.0089)	(0.0000)	(0.0000)	(0.0000)	(0.0064)	7.508790 (0.0000)
Windspeed		-0.159716*		-0.123031*		-0.137838*		-0.077030*	
P-value		(0.0000)		(0.0012)		(0.0000)		(0.0432)	
DF		147	150	148	151	148	148	148	150
AIC		45.43362	101.0056	7.441136	8.310884	45.58021	107.5101	7.443659	8.308137

humidity having negative effect on the dengue occurrence for the data from Putrajaya, Malaysia (July 2006 to December 2008). The effect of humidity on the occurrence of DF in Karachi is positive. We do not know exactly why there is such a contrast between Karachi and Putrajaya but we have a hunch that the variation of daily temperature in winter and summer varies differently in Karachi compare to Putrajaya. The other reason may be that the data from Putrajaya also includes cloudiness, may be this variable has aliasing effect on the model.

Note: We feel that there is no need to apply negative binomial and Poisson models for monthly data when applied these models to weekly data, 4-week lag(1 months) shows the best model. Model for monthly data did not show better results, so we have to drop the idea to present in this research.

The main effecting explanatory climatic factors are PC1 and PC4 (see Table 3). PC1 is interpreted as *general atmospheric condition (GAC)* and it effects positively on the occurrence of dengue fever cases. Whereas PC4 which we have labeled as *contrast between temperature and wind speed (CTW)* also seems to be causing in the increases of dengue fever cases. On the basis of expression (A) the following comments are made:

- 1) PC1 is effecting the chance of occurrence positively and in the past 4 weeks (28 days) PC1 value is estimated to positively influence the chances of dengue occurrence by 120.78% times.
- 2) PC4 is affecting the chance of occurrence positively and in the past 4 weeks (28 days) PC4 value is estimated to positively influence the chances dengue occurrence by 152.72% times.

Again the Poisson model is not suitable because of large (high) over dispersion also their AIC's values are very large that is why have avoided to consider the Poisson model [29]. For the comparison we have shown the estimation of Poisson parameter, while the negative binomial models are found to be suitable for all lags where the AIC values are very small [30].

We have used the first four PCs to model, DF data, using Poisson and negative binomials. The advantage of applying PC is that they are perfectly uncorrelated and thus there is no question of collinearity (among the explanatory variable) in the model [33]. The other advantage in this case is that instead of original

variable environmental factors variables are considered.

In this case of negative binomial models, all four models have shown quite good behavior, however for the model KNBPC (-4) is the best with smallest value of AIC of all four models PC4 is the best predictor. We interpret the model KNBPC (-4). This model says that PC1 and PC4 are having significant effect on DF. As PC1 represents general atmospheric condition, it means that variation in general atmospheric condition effects significantly (P-value 0.000) the occurrence of DF. The PC4 is interpreted as contrast between maximum temperature and wind speed has also significant effect (P-value=0.000) on the occurrence of DF, that is, high temperature with low wind speed is suitable for the spread of DF.

5. CONCLUSION

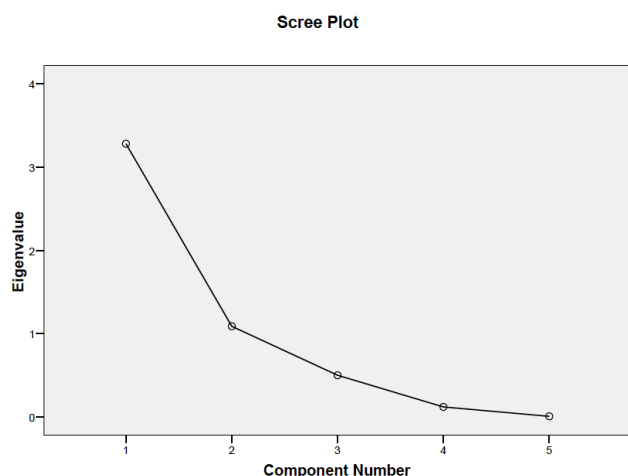
The environmental conditions of city are studied using principal component analysis and negative binomial models. Karachi is the largest city of Pakistan where environmental conditional is changing. Karachi is getting warmer. Negative binomial model for Karachi data with one month lag gives a good fit to DF cases. This is also true for environmental manifest data and environmental factors (PCs). Therefore we may conclude that the cases of dengue fever are affected by the change of month. In this study relative humidity and minimum temperatures 28 days or 4 weeks lag (approximate to one month) identified as significant predictor for the dengue incidence in Karachi. Thus relative humidity and minimum temperature have some impact on the occurrence of dengue fever. If the minimum temperature recorded increases from previous day, the numbers of dengue cases are expected to increase in next 28 days.

6. RECOMMENDATIONS

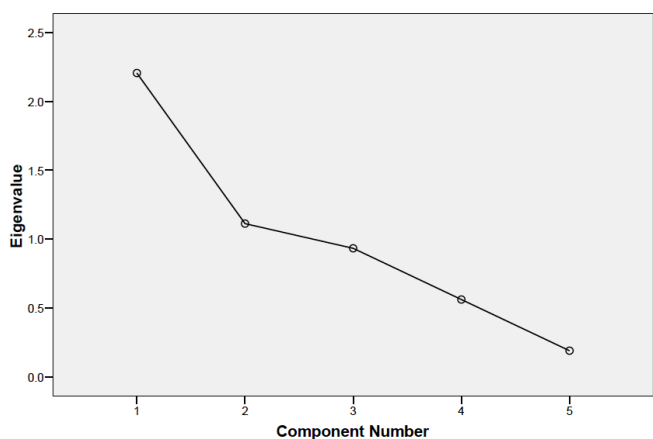
It is suggested that surveillance teams should keep an eye on the variation in relative humidity and minimum temperature. If the high variation in these two climatic variation persists, serious remedial measure should be taken. Since in months of August and September large number of dengue cases are expected so remedial measures should be taken in the months of June and July. The data we have is related to the number of dengue cases reported with respect to time, but we suggest further studies the number of death due to dengue locality from where the patient are coming, measure of awareness of the people are also needed to be studied. We also suggest some Bio-

statistical studies of the patient and the cause of spread of diseases. Effect of increase of population of the country on the spread of several diseases is also requires attention by the researcher and the governments. We find that over population of the cities not only creating social problems but also creates health problems. It is suggested that government should not only control the population and control cities from having over crowded population by making new cities where people can get job and can have health facilities and daily necessities. We also suggest creating dengue surveillance cells not only in provincial's capitals also at district levels whose function should be awareness, medical and fumigation on regular basis.

APPENDIX-I



Scree Plot of Karachi Monthly Climatic Variables (2010-2012).



Scree Plot of Karachi Weekly Climatic Variables (2010-2012).

REFERENCES

- [1] www.news-medical.net/health/Dengue-Fever.aspx.
- [2] World Health Organization. Dengue: Guidelines for diagnosis, treatment, prevention and control. New edition. France 2009.
- [3] Farrar J, Focks D, Gubler D, Barrera R, Guzman MG, Simmons C, *et al.* Towards a global dengue research agenda. *Trop Med Int Health* 2007; 12: 695-9. <http://dx.doi.org/10.1111/j.1365-3156.2007.01838.x>
- [4] www.cdc.gov/dengue/entomologyEcology/climate.html.
- [5] Weaver SC, Vasilakis N. Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. *Infect Genet Evol* 2009; 9: 523-540. <http://dx.doi.org/10.1016/j.meegid.2009.02.003>
- [6] Gubler DJ, Kuno G. Dengue and dengue hemorrhagic fever New York, NY: CAB International 1997.
- [7] Afrozuddin SA. Modelling of Population, Health and Environment Data through Generalized Linear Models (GLMs) Unpublished Thesis for the degree Ph.D, in the Department of Statistics, University of Karachi, and Karachi, Pakistan 2013.
- [8] Hii YL, Rocklöv J, Ng N, Tang CS, Pang FY, *et al.* Climate variability and increase in incidence and magnitude of dengue incidence in Singapore. *Global Health Action* 2009; 2.
- [9] Su GL. Correlation of climatic factors and dengue incidence in Metro Manila, Philippines. *Ambio* 2008; 37: 292-4. [http://dx.doi.org/10.1579/0044-7447\(2008\)37\[292:COCFAD\]2.0.CO;2](http://dx.doi.org/10.1579/0044-7447(2008)37[292:COCFAD]2.0.CO;2)
- [10] Arcari P, Tapper N, Pfueller S. Regional Variability in Relationships between Climate and Dengue/DHF in Indonesia. *Singaporean Journal of Tropical Geography* 2007; 28(3): 251-272. <http://dx.doi.org/10.1111/j.1467-9493.2007.00300.x>
- [11] Johansson MA, Dominici F, Glass GE. Local and Global Effects of Climate on Dengue Transmission in Puerto Rico. *PLoS Negl Trop Dis* 2009; 3(2). World Academy of Science, Engineering and Technology 2010; 38: 908.
- [12] Reiter P. Climate change and mosquito-borne disease. *Environmental Health Perspectives* 2001; 109: 141-61. <http://dx.doi.org/10.1289/ehp.01109s1141>
- [13] Kuhn K, Campbell-Lendrum D, Haines A, Cox J. Using climate to predict infectious disease epidemics. Geneva: World Health Organization 2005.
- [14] Chowell G, Stinchez F. Climate-Based Descriptive Models of Dengue Fever: The 2002 Epidemic in Colima, Mexico. *Journal of Environmental Health* 2006; pp. 41-44.
- [15] Barbazan P, Guiserix M, Boonyuan W, Tuntaprasart W, Pontier D, Gonzalez JP. Modelling the effect of temperature on transmission of dengue. *Med Vet Entomol* 2010; 24: 66-73. <http://dx.doi.org/10.1111/j.1365-2915.2009.00848.x>
- [16] Afrozuddin SA, Junaid SS. Using PCA, Poisson and Negative Binomial Model to Study Climatic Factor and Dengue Fever Outbreak in Pakistan. *proceedings, 12th International Conference on Statistical Sciences: March 24-26, 2014; Vol. 26.*
- [17] Brunkard JM, Cifuentes E, Rothenberg SJ. Assessing the role of temperature, precipitation and ENSO in dengue re-emergence on the Texas – Mexico border region. *Salud Pública de México* 2008; 50: 227-34. <http://dx.doi.org/10.1590/s0036-36342008000300006>
- [18] Amarakoon AMD, Chen AA, Rawlins SC, Taylor MA. Dengue epidemics its association with precipitation and temperature, and its seasonality in some Caribbean countries. *West Indian Medical Journal* 2004; 53(Suppl 2): 60.
- [19] Chan YC, Salahuddin NI, Khan J, Tan HC, Seah CL, Li J. Dengue haemorrhagic fever outbreak in Karachi, Pakistan, 1994. *Trans R Soc Trop Med Hyg* 1995; 89: 619-20. [http://dx.doi.org/10.1016/0035-9203\(95\)90412-3](http://dx.doi.org/10.1016/0035-9203(95)90412-3)

- [20] Hakim ST, Saleem M, Nadeem SG. An Experience with Dengue in Pakistan: An Expanding Problem. *Ibnosina J Med BS* 2011; 3(1): 3-8.
- [21] Jahan F. Dengue Fever (DF) in Pakistan. *Asia Pac Fam Med* 2011; 10(1): 1. <http://dx.doi.org/10.1186/1447-056X-10-1>
- [22] Afrozuddin SA, Siddiqi JS, Quaiser S. Principal Component Analysis to Explore Climatic Variability that Facilitates the Emergence of Dengue Outbreak in Karachi. *Pakistan Journal of Meteorology* 2014; 11(21) 1-10.
- [23] www.karachi.com/v/geography
- [24] David L. Linear algebra and its applications. Addison-Wesley. New York.
- [25] Afrozuddin SA, Siddiqi JS, Quaiser S. Using PCA, Poisson and Negative Binomial Model to Study the Climatic Factor and Dengue Fever Outbreak in Lahore. *Journal of Basic & Applied Sciences* 2015; 11: 8-16.
- [26] Will T. Introduction to the singular value Decomposition. Davidson College 1999. www.Davidson.edu/academic/math/will/svd/index.html
- [27] A tutorial on Principal component analysis by Jonathon Shlens 1-13.
- [28] Edwards Jackson J. A user guide to Principal Components. Wiley series in Probability and Mathematical Statistics 1991.
- [29] Green WH. *Econometric Analysis* 5th ed. Prentice Hall 2003.
- [30] Fairoos W, W.Y, Azaki W, W.H, Alias M, Beewah Y. Modelling Dengue fever (DF) and dengue haemorrhagic fever (DHF) outbreak using Poisson and negative binomial model. *World Academy of Science, Engineering and Technology* 2010; 38: 903-908.
- [31] Sadiq N, Qureshi MS. Climatic variability and linear trend models for the five major cities of Pakistan. *Journal of Geography and Geology* 2010; 2(1): 38-39. <http://dx.doi.org/10.5539/jgg.v2n1p83>
- [32] Afrozuddin SA. Random Effects in Generalized Linear Models. Unpublished Thesis for the degree of M.Phil, in the Department of Mathematical Statistics and Operational Research, University of Exeter, Exeter, England 1993.
- [33] Jolliffe IT. Discarding Variables in a Principal Component Analysis. I: Artificial Data. *Appl Statist* 1972; 21: 160-173. <http://dx.doi.org/10.2307/2346488>

Received on 19-05-2015

Accepted on 14-07-2015

Published on 30-09-2015

<http://dx.doi.org/10.6000/1927-5129.2015.11.72>© 2015 Ahmed *et al.*; Licensee Lifescience Global.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.