Using PCA, Poisson and Negative Binomial Model to Study the Climatic Factor and Dengue Fever Outbreak in Lahore

Syed Afrozuddin Ahmed[1,*], Junaid Saghir Siddiqi[1], Sabah Quaiser[1] and Shahid Kamal[2]

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

[2]College of Acturial Science, University of Punjab, Pakistan

Abstract: Various studies have reported that global warming causes unstable climate and many serious impact to physical environment and public health. The increasing incidence of dengue incidence is now a priority health issue and become a health burden of Pakistan. The study aims to understand, explore and compare the climatic factors of Karachi and Lahore that causes the emergence or increasing rate of dengue fever incidence that effects the population and its health. Principal component analysis (PCA) is performed for the purpose of finding if there is/are any general environmental factor/structure which could be considered as Pakistani climate. We developed an early warning model for the prediction of dengue outbreak in Lahore. This has been done by using Poisson regression and Negative binomial regression model. For this purpose we use daily, weekly and monthly data of Lahore. The negative binomial model with lag (28) for Lahore daily data for climatic variable is best model. Lahore daily and weekly maximum temperature effect negatively and for the past 28 days it is estimated to negatively influence on the dengue occurrence by 26.1% times. Daily wind speed is effecting negatively by 14.7% times and minimum temperature effect positively for the past 28 days by 86.7%times.

Keywords: Dengue Fever, Principal component analysis, Negative Binomial Model.

INTRODUCTION

Climate is a composite term and this refers to weather conditions over a period of time. Climate is an average or general conditions of temperature, humidity, atmospheric pressure, wind speed and rainfall of a place. These elements act as climatic factors. Climate is an important subject of scientific enquiry, particularly as it has such an impact on vegetation, soil and health etc. Its importance increases manifold because of its influence on human life. Natural hazards like storms, floods and desertification are result s of climatic changes. Extreme temperatures, weather hot or cold affect human beings, crops, animals and mosquitoes lifecycles especially Aedes agepty mosquitoes which causes the Dengue Fever [1-3]. In short, there are many ways in which climate affects our daily lives. The viral disease which is called dengue transmitted to human body through the mosquito named Aedes agepty [3]. In the medical terminology its another name is “break bone fever” and the headache, pain in joints and muscle, headache, rashes in skin close to the measles are its symptoms. Sometime this disease is threatening to the life as results in blood disorder causes the lack of platelets and shortage of plasma and due to these problems dengue hemorrhagic shock syndrome may be observed also patients may go in to problem of low 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. This viral disease has spread in most part of the world [2]. 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 [4]. The dengue infectious disease spreads in warm climate [5]. Aedes Agepty mosquitoes are seriously sensitive by environmental conditions. The climatic variables i.e., precipitation, humidity, temperature and wind speed are helping for the survival and reproduction of dengue mosquitoes. High temperature helps to decrease the time lag for the replicate 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 faster and has higher probability to dangerous for human being before completing its life cycle. So we may say that the critical factor is climate by which the situation may be controlled to the mosquito habitat. We can see vast amount of information from the 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 studied groups [14, 15] regarding the correlation between rainfall and infection rate of dengue. The relationship between climatic factors and dengue fever incidence to indicate or predict variation in dengue incidence also shown by [16, 17]. The

*Address correspondence to this author at the Department of Statistics, University of Karachi, Pakistan; Tel: 0333-2165107;
E-mail: safrozuddin@yahoo.com

ISSN: 1814-8085 / E-ISSN: 1927-5129/15 © 2015 Lifescience Global

effects of temperature and precipitation on dengue transmission relation have assessed by [18] has studied the dengue epidemics and its association with precipitation and temperature.

In Karachi, Pakistan first Dengue fever outbreak reported in 1994 [18] 1 patient out of 145 died. Many researchers have studied and reported epidemics of dengue and its causes from different parts of Pakistan [19-21]. Pakistan has experience a number of dengue fever outbreaks since 1994. The reasons why and how the dengue epidemics become endemic diseases in Pakistan, different reasons are for this problems like poor hygiene and inefficient sewerage system as results these are the ideal habitat for the dengue vector to lay eggs and flourish. From the last few years transmission of dengue virus situation is high in the country, Karachi and Lahore are the cities which are heavily threatened by dengue epidemics in pre and post monsoon periods [22, 23]. 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 [22].

STATISTICAL MODEL

A. Generalized Linear Model

The generalized linear model is used [24-26] in the situations where an independent sequence of random variables is taken and a function of their means is linked to a linear predictor based on known covariates [27]. This allows regression techniques to be used for non-normal (but exponential family) distributions on a non-linear scale. These techniques may conveniently be used [28] more general and some function of μ_i is assumed to be a linear combination of β's [29]. The generalized linear model formulation is the usual method of fitting factors and variables to data which involve an exponential family error structure [30]. Thus, the GLM allows the extension of regression techniques to non-normal sampling models with effects on a suitable non-linear scale.

Use of Poisson and negative binomial regression models depends on the nature of the distribution of the dependents variables [30]. Researchers prefer to use Negative binomial regression because of the assumption of Poisson models which are hard to observe with the data, some researchers also prefer to use both Poisson and negative binomial model within the same study [29]. The model parameters give the same information in both Poisson and Negative binomial regression model. Poisson regression model are commonly used for count data [31]. In our study the "events" nature of severe and non severe dengue cases. We have used the both Poisson and negative binomial regression models in our study where we have use the count data of dengue occurrence. Poisson and negative binomial regression models are different in regards to their assumptions of the mean and variance of the dependents variables, in the case of Poison model the assumption is the mean and variance of the distributions are equal i.e. equi dispersion whereas the negative binomial regression models we do not use the assumptions of equal mean and variance and is particularly correct for using this model when the over dispersion is in the data. Which shows that mean is less than variance [30].

Some researchers have noted that count data for occurrence of diseases often have over dispersion so they have preferred to use negative binomial regression in those types of studies. We have also preferred to use negative binomial regression models on our data because it also shows the over dispersion.

B. Principal Component Analysis

In this study our focused is to exploring the climatic variables and/or the climatic factors (structure) which may influence in the increase or decrease in the number of dengue fever cases in Karachi and Lahore. To find the climatic structure we used Principle component analysis method. 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 [32]. This is one of the way of identifying patterns in the data [33]. It is difficult to find patterns in the data for this purposed we use PCA which highlight 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 [34]. The new components are called PC1, PC2, PC3 and so on, (for the first, second and third principal components) are independent and 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

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 temperatrure	Mn	°C
5	Wind speed	W	Knots
6	Dengue Fever	DF	

Table 2: No. of Dengue Cases Reported in Lahore (2011-2012)

		Jan	Feb	March	April	May	June	July	Aug	Sep	Oct	Nov	Dec	Total
Lahore	2011	0	0	4	5	15	3	6	1124	8254	5805	1531	55	1680
	2012	0	0	0	0	5	1	0	6	174	58	15	1	260

Source: Dengue surveillance cell Lahore, Govt. of Punjab, Pakistan.

on until all the variance is accounted for, in this way very few will be retain for the further study [35]. 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. So we don't find any problem in the application of principle component analysis.

DATA AND METHODOLOGY

This study was conducted in Lahore capital of Punjab Province. The climate of Lahore features an arid climate, albeit a moderate version of this climate. The features of Lahore are clearly based on five season semi-arid climate with five seasons. The climate data of Lahore 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 while dengue cases in Lahore are reported in daily basis from 2011 to 2012. 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 2011 Govt. of Punjab, Pakistan has established the Dengue surveillance cell for not only keeping the records of Dengue Fever reported 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, tables etc and unit of measurements are listed in the Table **1**. In Table **2** we presented number of dengue fever cases by months, for the years 2011 and 2012.

Graphs shown in Figures **1** to **2** also reveal this fact, in these graphs high peak is during September October 2011, while this is negligible for the year 2012. The

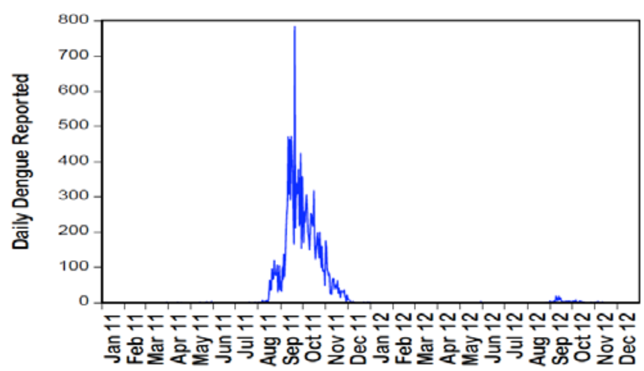

Figure 1: Lahore Daily Dengue Cases.

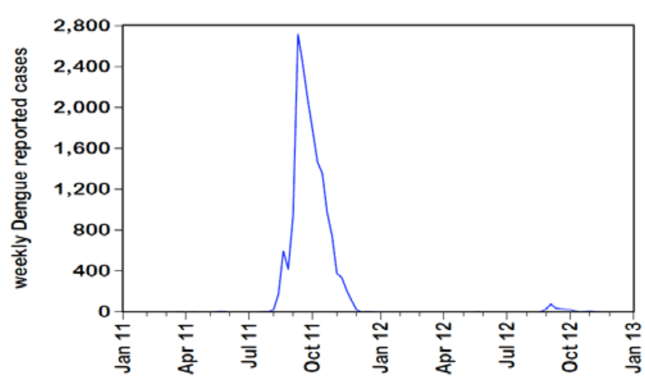

Figure 2: Lahore Weekly Dengue Cases.

graph for daily, weekly and monthly data are similar, spikes can be seen in the daily graph but, it is completely smooth in the monthly data graph. The continuous efforts of the government have reduced the dengue outbreak which can be observed from the year 2012 in the given Table **2**.

RESULTS AND DISCUSSION

A. PCA to Explore Climatic Structure

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 [7, 8] have also discussed the relation between climatic variable and dengue incidence.

For the purpose of statistical analysis we have taken the Lahore daily and weekly data from 2011-2012. The reason why we have considered the data of these periods is that dengue data is available for these periods only.

We have applied Principal Component Analysis technique for daily and weekly data of Lahore 2010-2012 to explore if there is any difference in the structure of climate in these time period. For explanation and understanding purposes we present all four PCs for considered period in Table **3**. These tables contains, 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, all plots 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. Table **3** is prepared for the daily, weekly and monthly basis for the period 2011-2012.

The Interpretation and Labeling of Principal Component are done on the basis of Table **3**. The PC1s for Lahore daily data of 2011-2012 are constructed as linear combination of same climatic variables, that are Mx, Mn and W. Therefore this component can be interpreted or labeled as "Windy and hot "(HtW)". While in PC1 of weekly period of 2011-2012, one more climatic variable P has been added and these can be interpreted as " Windy, hot and rainy" (WHtR). Temperatures, Mx and Mn Contributes highly in PC1 of all data. The percentage contribution of all PC1s varies from 43% to 48% wh. One can generally conclude that major climatic factor is variation in temperature together with high wind.

PC2 of Lahore daily period of 2011-2012 are linear combination of H and P thus, this factor can be named as "Wetness" (Wt).

PC3 for Lahore weekly period 2011-2012 is the contrast between Wind and average temperatures, we can label these factors "windy with low temperature" (WLT). PC3 of daily period of 2011-2012 are found to be contrast between W and P that we may label as "windy and dry" (WD).

PC4 for Lahore of two sets of data are contradiction between H and P that indicate in Lahore there is a climatic factor which is "humid but no rain" (HnR). The percentage of contribution varies from 6% to 14%. It is interesting to see that for daily period of 2011-2012 the contribution is quite high i.e. 14% and for weekly data of 2011-2012 is only 6%.

B. Modeling DF with Climatic Variables with Different Lags

Lahore Daily Data

We have applied the negative binomial and Poisson regression models on Lahore daily climatic variables

Table 3: PCA for Lahore Daily and Weekly Climatic Variables: (2011 – 2012

PCs	2011-2012 (Daily Dengue period)			2011-2012 (Weekly Dengue period)		
	Linear combination	***Label***	***%***	***Linear combination***	***Label***	***%***
PC1	0.9(Mx+Mn) + 0.5W	HtW	43	.9(Mx+Mn)+.7W+.5P	WHtR	50
PC2	0.75(P+H)	Wt	27	0.9H+0.8P	Wt	31
PC3	0.7W-0.5P	WD	15	0.6W-0.3(Mx+Mn)	WLT	12
PC4	0.5(H-P)	HnR	14	0.3H-0.4P	HnR	6

Where:
WLT is windy with low temperature; **HnR** is humid but no rain; **Wt** is wetness; **WD** is wind and dry; **HtW** is windy and hot; **WHtR** is windy hot and rainy.

and climatic factors (principal component score) for the Dengue period 2011-2012. In both negative binomial and Poisson model the dependent variable is DF cases.

Poisson model is not found suitable because of large (high) over dispersion and all their AIC's values are very large the AIC value varies between 55.421 to 68.49, when using climatic variable, and it varies from 63.93 to 79.34 in case when using PC's as explanatory variables. Indicating how poor these models are so we will not consider the Poisson model. We have shown the estimation of Poisson parameter for the purpose of comparison, while negative binomials model are suitable for all lags as the value of AIC are very small.

The following expressions are established from negative binomial models from the Tables **4** and **5**. In these equation we have only consider only those variables whose coefficient are significant at least α = 0.05.

Log (DF) = 5.012- 0.398Mx + 0.628Mn -0.227W   Lag (-7) [3.894] LNBCV (-7)

Log (DF) = 3.258 -0.394Mx + 0.675 Mn – 0.215 W   Lag (-14) [3.855] LNBCV (-14)

Log (DF) = - 0.426Mx + 0.736Mn – 0.187W   Lag (-21) [3.819] LNBCV (-21)

Log (DF) = - 0302Mx + 0.624Mn – 0.159W   Lag (-28) [3.815] LNBCV (-28)

We consider LNBCV (-28) model as the best representative because its AIC value is the smallest among the 4 models. In all the four lag model maximum temperature and wind speed shows significant effect on prevention of "DF" whereas the minimum temperature has significant effect on the occurrence of "DF". We write the best fitted model, i.e., LNBCV (-28), transforming it to DF form:

$\log(DF) = -0.302\,(Mx) + 0.624\,(Mn) - 0.159(W)$

$DF = (e^{-0.302Mx})\,(e^{0.624Mn})\,(e^{-0.159W})$

The main effecting explanatory variable are maximum temperature, minimum temperature and wind speed, where minimum temperature effecting positively on the occurrence of dengue fever cases, while the remaining two variables seems to be causing in the reduction of dengue fever cases,

1) Daily maximum temperature is effecting negatively and in the past 28 days maximum temperature value is estimated to negatively influence on the dengue occurrence by 26.1%.

2) Daily wind speed is effecting negatively and in the past 28 days wind speed value is estimated to negatively influence on the dengue occurrence by 14.7%.

3) Minimum temperature is effecting very highly on the daily DF occurrence. For the past 28 days the minimum temperature affects positively on the dengue occurrence by as high as 86.7%.

Lahore Weekly Data

The same models negative binomial and Poisson regression models are applied to Lahore weekly climatic data of the Dengue period 2011-2012. The following equations are established from negative binomial models from the Tables **4** and **5**.

Log (DF) = 22.993- 0.164H -1.056Mx + 1.223Mn - 0.610W   Lag (-1) [6.346]

Log (DF) = 19.030 -0.146H-0.983Mx + 1.232Mn – 0.545 W   Lag (-2) [6.260]

Log (DF) =14.766 -0.948Mx + 1.252Mn – 0.521W   Lag (-3) [6.198]

Log (DF) = 14.136 -0.116H -0.849Mx+1.149Mn -0.442W   Lag (-4) [6.214]

The climatic variables for the Dengue period 2011-2012 are used in negative binomial and Poisson regression model with different lags. Again the Poisson model is not suitable because of large (high) over dispersion also their AIC's values are very large so we have avoided to consider the Poisson model.

$\log(DF) = 14.76 - 0.948(Mx) + 1.252(Mn) - 0.521(W)$

$DF = (e^{14.1766})\,(e^{-0.948Mx})\,(e^{1.25Mn})\,(e^{-0.521W})$

We have shown the estimates of Poisson model parameter in the Tables **4** and **5** for the purpose of comparison while the negative binomial models are suitable for all lags where the AIC values are very small (see Tables **4** and **5**)

However the best model is negative binomial model with climatic variables with three weeks i.e. 21 days lag with AIC value 6.198. This model says that maximum temperature and Wind speed negatively effect on dengue occurrence whereas minimum temperature cause positively on occurrence of dengue (see Tablse **4** & **5**). This model and the model LNBCV (-28) (for daily data) give the same information, i.e. Mx and W make negative effect while Mn effects positively.

The models with lag 1, 2 and 4 contain variables H, Mx, Mn and W while the model with smallest value of AIC (6.198) missed H.

The main effecting explanatory variable are maximum temperature, minimum temperature and wind speed, where minimum temperature effecting positively on the occurrence of dengue fever cases, while the remaining two variables seems to be causing in the reduction of dengue fever cases,

1) Weekly maximum temperature is effecting negatively and in the past 4 weeks (28 days) maximum temperature value is estimated to negatively influence on the dengue occurrence by 61.2%.

2) Weekly wind speed is effecting negatively and in the past 4 weeks (28 days) wind speed value is estimated to negatively influence on the dengue occurrence by 40.6%.

3) Minimum temperature is effecting very highly on the daily DF occurrence. For the past 28 days the minimum temperature affects positively on the dengue occurrence by as high as 249.06%.

C. Modeling DF with Climatic Factor (PCs) with Different Lags

Lahore Daily Data

We have applied the negative binomial and Poisson regression models on Lahore daily climatic factors (PCs) for the dengue period 2011-2012. We have used the first four PCs as independent variables to explore if they (PCs) are making only significant effect on daily cases of dengue fever in increasing or decreasing the number of cases using Poisson and negative binomial models. In this case Poisson model could not be found suitable to explain the variation on the dengue fever data. The advantage of applying PC is that they are perfectly uncorrelated and thus there is no question of collinearity in the model. The other advantage in this case is instead original variable environmental factors variables are used to see which climatic factor may be seen as cause of increasing or decreasing number of daily dengue fever cases. The interpretation of PCs can be found from the Table **3**.

The following equations are established from negative binomial models from the Tables **4** and **5**. The following equations are constructed form negative binomial Model with different lags using PCs as explanatory variables from table, only those variable (factors) are presented whose coefficient are significant with p-value less than 0.05.

Log (DF) = 2.011 + 2.341 PC2 – 0.442 PC4 Lag (-7) [3.961] LNBPC (-7)

Log (DF) = 1.672 + 2.665 PC2 – 0.417 PC4 Lag (-14) [3.942] LNBPC (-14)

Log (DF) = 1.336 + 2.908 PC2 – 0.364 PC4 Lag (-21) [3.920] LNBPC (-21)

Log (DF) = 1.053 + 2.965 PC2 – 0.430 PC4 Lag (-28) [3.901] LNBPC (-28)

All four models show small values of AIC 3.901 to 3.961. The negative binomial model LNBPC (-28) has smallest value of AIC that is 3.901, so we choose the model LNBPC (-28) as best model and the model shows that the positive effect of *wetness* (PC2) and the negative effect of *humidity but no rain* (PC4) on dengue occurrence. Corresponding AIC values for Poisson models varies from 63.93 to 79.34 indicating how much the negative binomial model is better than Poisson model. In four models PC2 has positive sign but PC4 has negative sign, thus interpretation will be same for all models. It is interesting to note that when data is transformed, in other words considering climatic factors instead of original variables. The model says that humid rainy weather provide opportunity for the survival of dengue mosquito, while when the weather is humid with no rain its prevent dengue fever. We write the best fitted model, i.e., LNBPC (-28), transforming it to DF form to explore the amount of effect of the two PC's on DF cases.

$$\log(DF) = 1.053 + 2.965(PC2) - 0.430(PC4)$$

$$DF = (e^{1.053})\,(e^{2.965PC2})\,(e^{-0.430PC4})$$

This model is highly skewed no linear trend exists therefore we cannot make comments as we have done previously in case of LNBCV (-28). Again the Poisson models are not suitable because of large (high) over dispersion values are very large so we have avoided to consider the Poisson model. For the comparison we have shown the estimation of Poisson parameter, while the negative binomial model is suitable for all lags where the AIC values are very small.

Lahore Weekly Data

After the application of negative binomial and Poisson regression models using weekly climatic factors from Lahore on the Dengue period 2011-2012.The following equations are established from negative binomial models from the Tables **4** and **5**. The climatic factors for the Dengue period 2011-2012 are used in negative binomial and Poisson regression model with different lags. The following equations are performed form Negative binomial Model with different lags using PC'S as explanatory variables from Tables **4** & **5**. Again we have only included in these equations those variables whose coefficients are significant at α = 0.05.

Log (DF) = 3.680 + 1.460 PC1 – 1.574PC3 Lag (-1) [6.519]

Log (DF) = 3.259 + 2.019PC1 –1.626 PC3 +0.920 PC4 Lag (-2) [6.474]

Log (DF) = 2.835 + 2.110 PC1 -1.837 PC3 + 1.452 PC4 Lag (-3) [6.399]

Log (DF) = 2.5492+2.287PC1+1.133PC2-1.614PC3+0.898PC4 Lag (-4) [6.382]

Again the Poisson model is not found suitable because of large (high) over dispersion also their AIC values are very large so we have avoided considering the Poisson model. For the comparison we have shown the estimation of Poisson parameter, while the negative binomial model is suitable for all lags where the AIC values are very small.

As we have stated earlier that the advantage of applying PC is that they perfectly uncorrelated and thus there is no question of collinearity (among the explanatory variable) in the model. The other advantage in this case is instead original variable environmental factors variables are used.

The interpretation of PC can be found from the Table **3**. In this case of negative binomial models, all four models are shown quite good behavior, but for the model LNBPC (-4) is the best with smaller value of AIC. We interpret the model LNBPC (-4). This model says that PC1 i.e. general atmospheric condition. PC2 is the combine effect of precipitation and humidity that explains 31% of variation whereas the PC4 which is the contradiction of precipitation and humidity that explains only 6% of variation having positive effect on the dengue outbreak however the PC3 which is the contrast of wind speed with temperatures having negatively affected on dengue occurrence i.e. prevention of dengue outbreak.

Note

We feel that there is no need to NB/Poisson models for monthly data because when applied these model to weekly data, 4-week lag shows the best model. Therefore model for monthly data may not show better results.

CONCLUSIONS AND RECOMMENDATIONS

The environmental conditions of Lahore are studied using PC and negative binomial models. Lahore is the largest cities of Pakistan there environmental

Table 4: Lahore Weekly Dengue, Climate Variable and PCA with Different Lags (2011-2012)

		Lag 1=1 week				**Lag 2=2 week**			
		Poisson		***Negative Binomial***		***Poisson***		***Negative binomial***	
C.V	***PCA***	**LPCV(-1)**	**LPPC(-1)**	**LNBCV(-1)**	**LNBPC(-1)**	**LPCV(-2)**	**LPPC(-2)**	**LNBCV(-2)**	**LNBPC(-2)**
		C.V	**PCA**	**C.V**	**PCA**	**C.V**	**PCA**	**C.V**	**PCA**
Constant	**Constant**	25.6303*	4.208996*	22.99307*	3.679522*	24.11909*	4.030225*	19.02968*	3.259218*
P-value	**P-value**	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Rainfall	**PC1**	0.151803*	0.971068*	0.075160	1.459905*	0.125212*	1.203691*	0.104581	2.019065*
P-value	**P-value**	(0.0000)	(0.0000)	(0.5516)	(0.0189)	(0.0000)	(0.0000)	(0.3552)	(0.0005)
Humidity	**PC2**	-0.197634*	0.218382*	-0.164177*	0.637811	-0.186571*	0.335563*	-0.145725*	0.649315
P-value	**P-value**	(0.0000)	(0.0000)	(0.0074)	(0.3168)	(0.0000)	(0.0000)	(0.0046)	(0.22850)
Maxtemp	**PC3**	-1.153103*	-1.073113*	-1.056105*	-1.573685*	-1.183044*	-1.029282*	-0.982974*	-1.625876*
P-value	**P-value**	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Mintemp	**PC4**	1.302340*	-0.023035*	1.223284*	0.980951	1.360018*	0.036117*	1.231869*	0.920636*
P-value	**P-value**	(0.0000)	(0.0016)	(0.0000)	(0.0435)	(0.0000)	(0.0000)	(0.0000)	(0.0337)
Windspeed		-0.607758*		-0.610050*		-0.485962*		-0.545110*	
P-value		(0.0000)		(0.0000)		(0.0000)		(0.0000)	
D.F		**98**	**99**	**99**	**101**		**98**	**98**	**99**
AIC		**222.90**	**464.47**	**6.346**	**6.519**		**435.52**	**6.260**	**6.474**

Table 5: Lahore Weekly Dengue, Climate Variable and PCA with Different Lags (2011-2012)

		Lag 3=3 week				Lag 4=4 week			
C.V	***PCA***	***Poisson***		***Negative Binomial***		***Poisson***		***Negative binomial***	
		LPCV(-3)	**LPPC(-3)**	**LNBCV(-3)**	**LNBPC(-3)**	**LPCV(-4)**	**LPPC(-4)**	**LNBCV(-4)**	**LNBPC(-4)**
		C.V	**PCA**	**C.V**	**PCA**	**C.V**	**PCA**	**C.V**	**PCA**
Constant	**Constant**	18.01813*	3.767791*	14.76573*	2.834607*	16.05917*	3.423748*	14.13624*	2.592102*
P-value	**P-value**	(0.0000)	(0.0000)	(0.0059)	(0.0000)	(0.0000)	(0.0000)	(0.0065)	(0.0000)
Rainfall	**PC1**	0.039401*	1.486695*	-0.030708	2.109989*	0.166028*	1.668099*	0.163236	2.286586*
P-value	**P-value**	(0.0000)	(0.0000)	(0.7771)	(0.0002)	(0.0000)	(0.0000)	(0.1065)	(0.0000)
Humidity	**PC2**	-0.130130*	0.483574*	-0.080481	1.277328*	-0.118620*	0.597242*	-0.116017*	1.132829*
P-value	**P-value**	(0.0000)	(0.0000)	(0.1502)	(0.0119)	(0.0000)	(0.0000)	(0.0273)	(0.0019)
Maxtemp	**PC3**	-1.007152*	-0.953186*	-0.947992*	-1.837171*	-0.861272*	-1.005102*	-0.848551*	-1.613902*
P-value	**P-value**	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Mintemp	**PC4**	1.218032*	0.205335*	1.251960*	1.452142*	1.043985*	0.059440*	1.148540*	0.898104*
P-value	**P-value**	(0.0000)	(0.0000)	(0.0000)	(0.0006)	(0.0000)	(0.0000)	(0.0000)	(0.0053)
Windspeed		-0.274910*		-0.521240*				-0.441874*	
P-value		(0.0000)		(0.0000)				(0.0001)	
D.F		**96**	**97**	**98**	**97**	**97**	**96**	**96**	**96**
AIC		**235.87**	**403.45**	**6.198**	**6.399**	**208.42**	**325.14**	**6.214**	**6.382**

conditional is changing. Lahore is getting cooler. Negative binomial model for Lahore data with one month lag give 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 effected by the change of month. In this study relative humidity and minimum temperatures at 28 days or 4 weeks lag (approximate to one month) identified as significant predictor for the dengue incidence in Lahore. 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 number of dengue cases is expected to increases in next 28 days. It is suggested that surveillance teams should keep eyes on the changes in relative humidity and minimum temperature and if the minimum temperature for several take serious remedial measures. 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.

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] 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

[5] Gubler DJ, Kuno G. Dengue and dengue hemorrhagic fever New York, NY: CAB International 1997.

[6] Afrozuddin SA, Junaid SS. Principal Component Analysis to Explore Climatic Variability that facilitates the Emergence of Dengue Outbreak in Karach. Will be Published in 21st issue (January 2015) of the Pakistan Journal of Meteorology 2014.

[7] 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.

[8] 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

[9] 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

[10] 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.

[11] Reiter P. Climate change and mosquito-borne disease. Environmental Health Prospectives 2001; 109: 141-61. http://dx.doi.org/10.1289/ehp.01109s1141

[12] Kuhn K, Campbell-Lendrum D, Haines A, Cox J. Using climate to predict infectious disease epidemics. Geneva: World Health Organization 2005.

[13] 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.

[14] 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

[15] 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

[16] 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(Supp 2): 60.

[17] 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

[18] Hakim ST, Saleem M, Nadeem SG. An Experiencewith Dengue in Pakistan: An Expanding Problem. Ibnosina J Med BS 2011; 3(1): 3-8.

[19] Humayoun MA, Waseem T, Jawa AA, Hashmi MS, Akram J. Multiple dengue serotypes and high frequency of dengue hemorrhagic fever at two tertiary care hospitals in Lahore during the 2008 dengue virus outbreak in Punjab, Pakistan. Int J Infect Dis 2010; 14S3: e54-e59.

[20] 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

[21] Riaz MM, Mumtaz K, Khan MS, Patel J, Tariq M, Hilal H, Siddiqui SA, Shezad F. Out break of Dengue Fever in Karachi 2006: a clinical perspective. J Pak Med Assoc 2009; 59(6): 9-4.

[22] Pakistan-economic-survey-2012-2013.

[23] www.karachi.com/v/geography.

[24] Afrozuddin SA. Random Effects in Generalized Linear Models. Unplished Thesis for the degree of M.Phil, in the Department of Mathematical Statistics and Operational Research, University of Exeter, Exeter, England 1993.

[25] Afrozuddin SA. Normaility Assumptions in Generalized Linear Models. Second International Seminar on Islamic Economics. Organized by department of Statistics, University of Peshawar, Pakistan 1995; 127-142.

[26] Hardin JW, Hilbe JM. Generalized Linear Models and Extensions. A Stata Press Publication. Texas, USA 2012.

[27] Bradley EL. The equivalence of maximum likelihood and weighted least squares in the exponential family. Journal of the American Statistical Association 1973; 68: 199-200.

[28] Knudsen DC. Generalizing Poisson regression: Including a priori information using the methods of offsets. The Professional Geographer 1992; 44: 202-208. http://dx.doi.org/10.1111/j.0033-0124.1992.00202.x

[29] Bock RD, Lieberman M. Fitting a response model for n dichotomously scored items. Psychometrika 1970; 35: 179-197. http://dx.doi.org/10.1007/BF02291262

[30] Lovett A, Flowerdew R. Analysis of count data using Poisson regression. The Professional Geographer 1989; 41: 190-198. http://dx.doi.org/10.1111/j.0033-0124.1989.00190.x

[31] MacDonald J, Lattimore P. Count models in criminology. In Piquero A, Weisburd D, Eds. Handbook of Quantitative Criminology. Springer 2010.

[32] David L. Linear algebra and its applications. Addison-Wesley. New York.

[33] Will T. Introduction to the singular value Decomposition. Davidson College 1999. www.Davidson.edu/academic/math/will/svd/index.html

[34] A tutorial on Principal component analysis by Jonathon Shlens 1-13.

[35] Edwards J. A user guide to Principal Components. Wiley series in Probability and Mathematical Statistics 1991.

Received on 18-08-2014 Accepted on 23-12-2014 Published on 07-01-2015

http://dx.doi.org/10.6000/1927-5129.2015.11.02

© 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.