Asian Journal of Business and Management (ISSN: 2321 – 2802)
Volume 04– Issue 02, April 2016

Environmental Change and Natural Disasters: Asian Experience

Jayadevan CM

Charles Stuart University
Australia
Email: Jdevcm [AT] outlook.com

ABSTRACT---- *Environmental change plays an important role in natural disasters. Increasing emission of carbon dioxide, methane, nitrogen, Sulphur dioxide, ozone depleting substances, greenhouse gas, PM10 pollutionand falling share of forest area can cause natural disasters like windstorms, flood, severe draught, heat waves etc which in turn cause substantial loss to the people in the form of death, injury, homelessness and other damages. In this paper, an attempt has been made to find relationship between environmental change and natural disasters with the help of method of canonical correlation, linear discriminatory function analysis and multinomial logistic regression using the data for Asian countries during the period 1991-2013. The study finds a statistically significant relationship between environmental change and natural disasters using multivariate statistical tests. Using univariate regression tests, emission of CO2 and PM10 pollution are found to be statistically significant factors contributing to the natural disasters. LDA technique and multinomial logistic regression show in addition to CO2 and PM10, deforestation also contributed to the natural disasters.*

Keywords--- Environment, natural, disasters,carbon, dioxide

---

1. INTRODUCTION

Asia is home to more than 60 percent of the world's population, produces well over a third of global gross domestic product, has two super powers in waiting in both China and India, and presents a full range of the world's most challenging tradional and nontraditional security concerns(Freeman and Green, 2010).Asia-Pacific partnership on clean Development and Climate (APP), which includes Australia, Canada, China, India, Japan, ROK, and the United States, collectively account for over 50 % of the world's energy use and greenhouse emissions (Freeman and Searight, 2010).In terms of climate change, Asia includes four countries with significant carbon emissions, making the region integral to any global efforts to combat climate change. Two are megacountries with populations of more than 1 billion and rapidly growing economies: China became the world's largest emitter of greenhouse gases(GHG) within the past year, and India's overall emissions now place it in the top countries. Japan is the most idustrialized country in the region and the fifth largest GHG emitter in the world. Indonesia is the third largest GHG emitter and is home to some of the world's major tropical forest resources. Deforestation and forest degradation, as well as peat forest fires, release large amounts of carbon into the atmosphere. These factors make Indonesia's potential for reducing emissions completely different from that of the other significant emitters, as their emissions are derived from a different source(Schaffer,2010).Climate change may not be responsible for the recent skyrocketing cost of natural disasters, but it is very likely that it will impact future catastrophes. Climate models provide a glimpse of the future, and while they do not agree on all of the details, most models predict a few general trends. First, according to the Intergovernmental Panel on Climate Change, an increase of greenhouse gases in the atmosphere will probably boost temperatures over most land surfaces, though the exact change will vary regionally(NASA).

Environmental change plays an important role in natural disasters. Increasing emission of carbon dioxide, methane, nitrogen, Sulphur dioxide, ozone depleting substances, greenhouse gas, PM10 pollution and deforestation can cause natural disasters like floods, volcanic eruptions, earthquakes, tsunamis,windstorms, severe draught, heat waves and other geologic processeswhich in turn cause substantial loss to the people in the form of death, injury, homelessness and other damages. Climate change is predicted to have a range of serious consequences, some of which will have impact over the longer term, like spread of disease and sea level rise, while some have immediately obvious impacts, such as intense rain and flooding (Jason,A and Camilla, B,2015).The main purpose of this paper is to analyze the relationship between environmental change and natural disasters in Asia during 1991-2013.Hypothesis of interest is environmental deterioration cause natural disasters. A canonical correlation analysis has been performed in order to determine if there is a relationship between two sets of variables, one measuring environmental variables and the other measuring natural disasters. The study also makes an attempt to find the factors responsible for the variation in average number of natural disaster events using linear discriminatory function approach.and logistic regression.

2. MATERIALS AND METHODS

The main source of data for this study is taken from online statistical database published by United Nations ESCAP.
In this paper we will use a canoncal correlation analysis (CCA) as a technique for determining if there is a relationship between two sets of variables, one measuring natural disasters and the other measuring environmental change. CCA is a multivariate analysis of correlation between two sets of variables. In CCA, we study interrelationships between sets of multiple predictor variables and multiple response variables. Hypothesis of interest is environmental deterioration cause natural disasters. The null hypothesis is equivalent to testing the hypothesis that all p canonical variate pairs are uncorrelated, or the hypothesis of interest is: $H_\circ : \rho * 1 = \rho * 2 = \cdots = \rho * p = 0$; $H_a: Not\ all\ \rho_i\ equal\ zero$. Response variables representing natural disasters are :1)eventNo- Number of unforeseen and often sudden event that causes great damage, destruction and human suffering.2)deathNo-the number of recorded deathsfrom natural disasters, 3)pAffected-total people affected [Thousands] are sum of injured, homeless, and affected people as a result of a natural disaster, 4)damagUS$ - Economic consequences of a disaster, usually direct (e.g., damage to infrastructure, crops and housing) and indirect (e.g., loss of revenues, unemployment and market destabilization). In each case, the registered figure represents the value of damage at the moment of the event; i.e., the figures are true for the year of the event. Data are converted from millions of US dollars to 2005 US dollars millions.Predictor variables representing environmental change are :1)GHG**:**Greenhouse gasemission, total [Million metric tons of CO2 equivalent]**,** 2)SO2: Sulphur dioxideemission [Thousand tons],3)ODP: Consumption of ozone-depleting substances [ODP metric tons], 4)PM10: Concentration of PM10 in urban area [Micrograms per m3], 5)CH4:Methane (CH4) emission [Thousand tons], 6) CO2- Emission of Carbon dioxide (kg per 2005 US$ of GDP multiplied by GDP), 7)NO2:Nitrous oxide emission [Thousand tons], 8)Forest_Km2- Total forest area in km^2.

In addition to the CCA, we have also made an attempt find factors responsible for the natural distaster events using linear discriminant analysis (LDA). Number of natural disaster events has been has been considered as the response variable. Since this is a discrete variable, this has been classified into three categories, that is 1)6 - 25, 2)2 – 6 and 3) 1 - 2. LDA analysis attempts to use the predictor variables to distinguish among the groups of the response variable. If LDA is able to distinguish among groups, it must have a strong relationship to at least one of the predictor variables. Using LDA,a series of statistical tests are conducted to test the overall relationship among the predictor variables and groups defined by the response variable. Using LDA, this paper is also concerned with an analysis to determine if there is a significant effect of factors like CO2, ODP, PM10 and forest area on the natural disaster events.There are four predictor variables.The hypothesis of interest is: $H_0: \beta_1 = \beta_2 = \beta_4 = \beta_4 = 0; H_a: Not\ all\ \beta_i\ equal\ zero$. The test statistic used for LDA and CCA is $Wilk's\ Lambda\ \wedge = \prod_i \frac{1}{1+\lambda_i}$where λ_i are the eigen values of the corresponding design matrices. There are three main assumptions for LDA and CCA: they are 1)Multivariate Normality (MVN): To test for MVN, we begin by examining the marginal distributions of each univariate variable using box plots. If any of these plots show non-normality, then MVN is suspect and we use a procedure based on Mahalanobis distance, in which we construct a χ^2 probabilities to determine confirmity with multivariate normality. 2)Equality of covariances:the test for equality of covariances is based on Box's M-test and 3)Independence of observations: This test is a function of the experimental design, or data collection method and hence is not tested. For the purposes of this paper we assume that it is true.

3. EMPIRICAL RESULTS

The average annual number of events, death, people affected, damage cost, per capita emission and % of forest area are presented in Table 1. The largest average number of natural disaster events occurred in China, India, Philippines, Indonesia and Bangladesh (Figure 1). The largest average number of natural disaster deaths occurred in Indonesia, Bangladesh, China, India and Pakistan (Figure 2).Countries like China, India, Bangladesh, Philippines, Thailand and Pakistan had the largest average number of people affected by the natural disasters (Figure 3). The average damage cost of natural disasters in terms of US$ at 2005 prices was very high for countries like China, Japan, India, Turkey, Thailand and Indonesia (Figure 4).

Figure 1:Average Number of Events

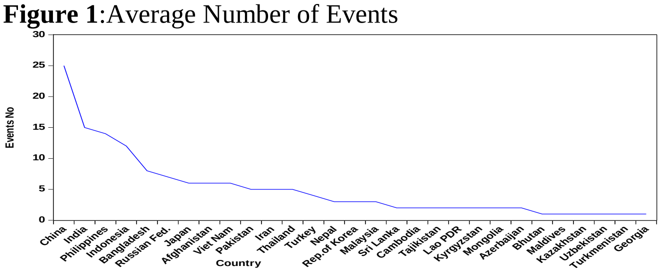


Figure 2:AverageNumber of Deaths

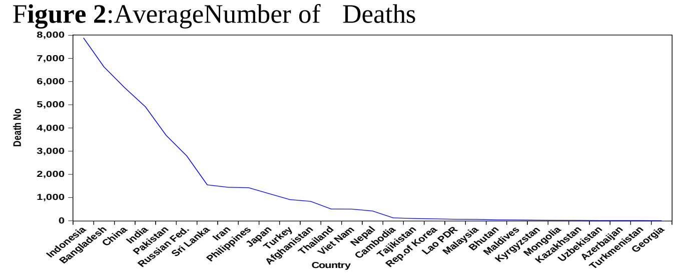


Figure 3:Average Number of People Affected [Thousands]

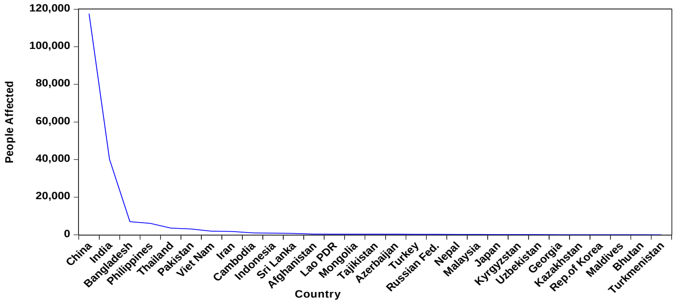


Figure 4:Average Damage Cost in Million US$

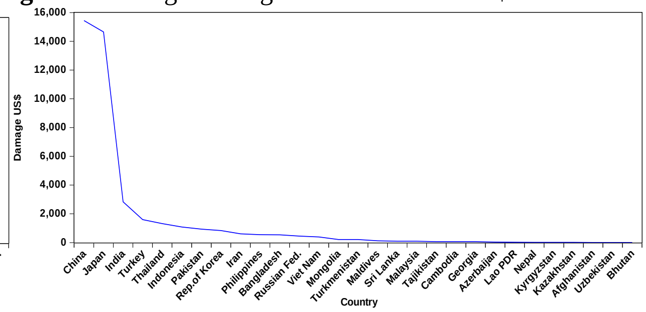


Figure 5:AverageCO2 (kg per 2005 US$ of GDP)

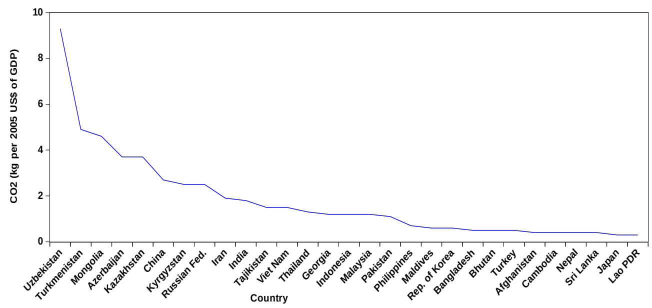
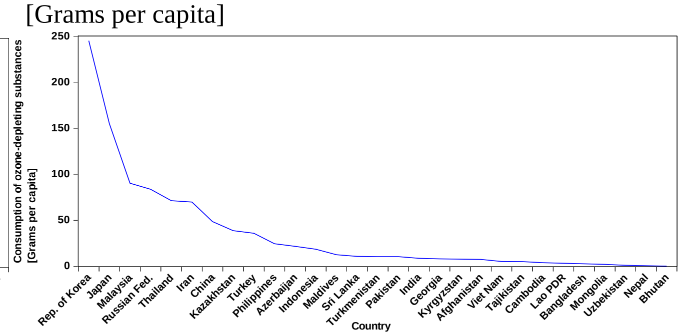


Figure 6:Average Consumption of ozone-depleting Substances [Grams per capita]



The largest average CO2 emission per GDP has been observed for Uzbekistan, Turkmenistan, Mongolia, Azerbaijan, Kazakhstan, China and Russian Fed (Figure 5). The average per capita consumption of ozone-depleting substances has been observed higher for Rep.of Korea, Japan, Malaysia, Russian Fed., Thailand, Iran and China (Figure 6). Forest area as % of total land for countries like Turkmenistan, Mongolia, Uzbekistan, Iran, Kyrgyzstan, Maldives, Tajikistan, Pakistan, Afghanistan and Kazakhstan are less than 10% (Figure 7). Average concentration of PM10 in urban areas is higher for countries like Mongolia, Pakistan, Iran, Nepal, India and Bangladesh (Figure 8).

Figure 7:Average AnnualForest Area as % of Total Land Area

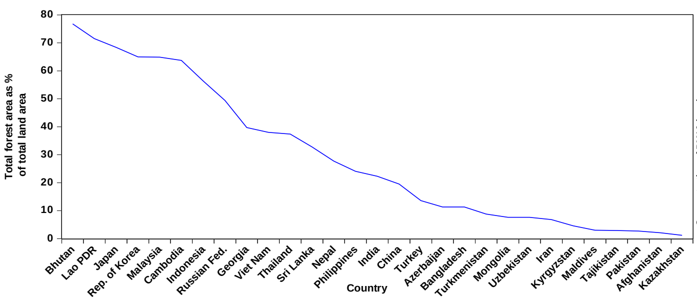


Figure 8:Average AnnualConcentration of PM10 in urban area [Micrograms per m3]

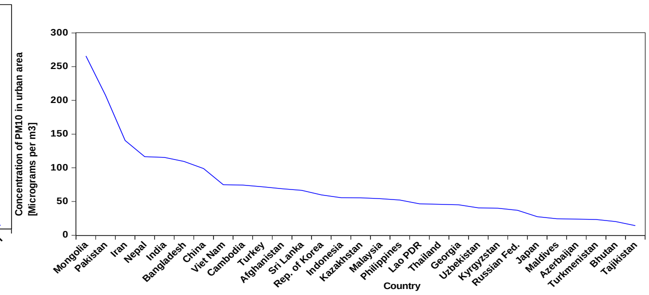


Table 1:Average Number of events, death, people affected, damage cost, Per capita Emission and % of Forest Area

Country	events No	deathNo	PeopleAffected	DamageUS$ (at US 2005 prices)	Odp(Grams per capita)	Forest(as percentage of total land area)	PM10(Micrograms per m3)	CO2((kg per 2005 US$ of GDP)
Afghanistan	6	838	356	9.6	7.6	2.1	69.0	.4
Azerbaijan	2	8	286	32.1	21.8	11.3	23.8	3.7
Bangladesh	8	6,633	6,946	543.3	2.7	11.3	109.4	.5
Bhutan	1	38	12	.6	.2	76.8	20.3	.5
Cambodia	2	124	1,021	63.6	3.9	63.7	74.4	.4
China	25	5,738	117,601	15,444.2	48.6	19.5	98.8	2.7
Georgia	1	2	58	60.2	8.1	39.7	45.2	1.2
India	15	4,917	40,249	2,835.3	8.7	22.3	115.4	1.8
Indonesia	12	7,883	847	1,080.8	18.6	56.4	55.7	1.2
Iran	5	1,445	1,787	603.8	69.9	6.8	140.5	1.9
Japan	6	1,164	136	14,651.4	155.0	68.4	27.5	.3
Kazakhstan	1	16	52	12.6	38.8	1.2	55.4	3.7
Kyrgyzstan	2	26	133	13.4	7.8	4.6	40.1	2.5
Lao PDR	2	63	334	28.3	3.4	71.5	46.6	.3
Malaysia	3	61	153	94.8	90.4	64.9	54.2	1.2
Maldives	1	37	13	128.5	12.7	3.0	24.4	.6
Mongolia	2	23	326	217.0	2.2	7.6	265.8	4.6
Nepal	3	421	192	15.7	.7	27.7	116.5	.4
Pakistan	5	3,687	3,069	936.2	10.5	2.7	207.4	1.1
Philippines	14	1,427	6,084	548.3	24.6	24.1	52.3	.7
Rep. of Korea	3	81	41	831.1	245.2	65.0	59.8	.6
Russian Fed.	7	2,798	213	450.7	83.8	49.4	37.0	2.5
Sri Lanka	2	1,548	698	96.1	10.8	32.8	66.5	.4
Tajikistan	2	95	293	65.8	5.2	2.9	14.3	1.5
Thailand	5	508	3,560	1,318.0	71.3	37.4	45.9	1.3
Turkey	4	911	265	1,596.1	35.9	13.6	71.9	.5
Turkmenistan	1	6	0	211.5	10.6	8.8	23.3	4.9
Uzbekistan	1	9	130	8.0	1.1	7.6	40.5	9.3
Viet Nam	6	504	1,878	394.0	5.3	38.0	74.9	1.5

Table 2:Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	Kurtosis
eventsNo	541	1	42	6	7	2.2	6.0
deathNo	538	0	166604	1806	11194	11.3	143.3
PeopleAffected	539	0	342029	8263	33767	6.2	42.7
DamageUS	539	0	166528	1848	9527	12.1	182.1
Odp(tons)	541	0	171588	6185	18799	5.3	32.4
Forest(Km^2)	541	9	8092685	552277	1609791	4.2	16.8
CO2 (1000 tons)	541	1	115526232	2728543	10940798	7.1	57.6
PM10(Micrograms per m3)	541	12	297	76	51	1.7	3.4

The summary statistics for the response and predictor variables are reported in Table 2. In this case, the degree of skewness is significantly skewed because the numerical value of skewness is greater than 2. So, we conclude that the distribution is significantly non-normal and in this case is significantly positvely skewed. In order to make the data normal, variables like eventsNo, deathNo, PeopleAffected, DamageUS$,Odp, Forest_km2, CO2, PM10 have been converted into logarithms such as leventsNo, ldeathNo, lPAffected, lDamageUS,lOdp, lForest, lCO2, lPM10 respectively. Boxplot for ldop, lpm10 and lforest_km2 show the presence of fewoutliers. Median for lCO2 is much higher than other variables. Variance for lCO2 and lodp are higher than lpm10 and lforest_km2 (Figure 9).

Figure 9: Box Plot for lodp, lpm10, forest and lCO2

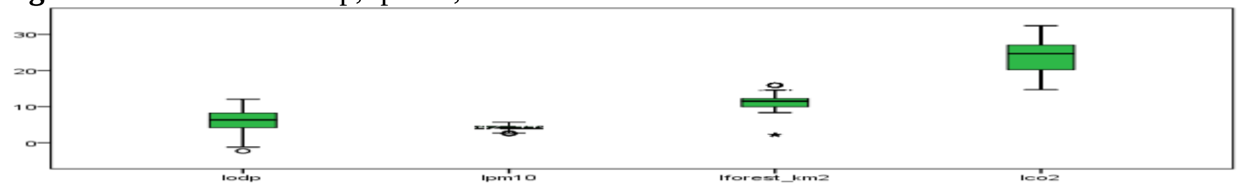


Table 2a:Correlations Among the Natural Disasters Variables

	leventsNo	ldeathNo	LpAffected	ldamageUS
leventsNo	1.00	.68**	.61**	.52**
ldeathNo	.68**	1.00	.60**	.55**
LpAffected	.61**	.60**	1.00	.52**
ldamageUS	.52**	.55**	.52**	1.00

**. Correlation is significant at the 0.01 level (2-tailed).

Table 2b:Correlations Among the Environment Variables

	lCH4	lodp	lpm10	lforest_km2	lco2	lghg	ln2o	lso2
lCH4_e	1.00	.70**	.41**	.80**	.83**	.92**	.97**	.87**
lodp	.70**	1.00	.19**	.60**	.79**	.77**	.74**	.82**
lpm10	.41**	.19**	1.00	.16**	.19**	.24**	.35**	.25**
lforest_km2	.80**	.60**	.16**	1.00	.66**	.81**	.79**	.74**
lco2	.83**	.79**	.19**	.66**	1.00	.93**	.87**	.92**
lghg	.92**	.77**	.24**	.81**	.93**	1.00	.96**	.95**
ln2o	.97**	.74**	.35**	.79**	.87**	.96**	1.00	.90**
lso2	.87**	.82**	.25**	.74**	.92**	.95**	.90**	1.00

**. Correlation is significant at the 0.01 level (2-tailed).

Table 2c: Correlations Among the Natural Disasters and Environmental Variable

Variable	leventsNo	ldeathNo	LpAffected	ldamageUS$	lodp	lpm10	lforest_km2	lCO2
leventsNo	1	.678**	.613**	.520**	.529**	.265**	.467**	.596**
ldeathNo	.678**	1	.596**	.545**	.329**	.340**	.250**	.345**
LpAffected	.613**	.596**	1	.517**	.348**	.351**	.310**	.359**
ldamageUS	.520**	.545**	.517**	1	.469**	.189**	.338**	.569**
lodp	.529**	.329**	.348**	.469**	1	.192**	.600**	.794**
lpm10	.265**	.340**	.351**	.189**	.192**	1	.156**	.193**
lforest_km2	.467**	.250**	.310**	.338**	.600**	.156**	1	.664**
lCO2	.596**	.345**	.359**	.569**	.794**	.193**	.664**	1

**. Correlation is significant at the 0.01 level (2-tailed).

The correlation between the variables of natural disasters are moderate, the largest being 0.678 between the eventsNo and deathNo (Table 2a). The correlations between the environment variables show the presence of multicollinearity between CO2 and CH4, CO2 and GHG, CO2 and N2o, and CO2 and So2. So this study has excluded the environmental variables like CH4, GHG, NO2 and SO2 (Table 2b). The correlations between the variables of natural disasters and environment variables are moderate (Table 2c).

3.1 CCA Results

In our example, we have multiple regressions predicting the y=4 natural disaster variables from the x=4 environment variables. We wish to test the null hypothesis that these regression coefficients are all equal to zero. This would be equivalent to the null hypothesis that the first set of variables is predictor from the second set of variables.$H_o = \beta_{ij} = 0; i = 1,2, \dots, y; j = 1,2, \dots, x$. This is carried out using Wilk's lambda. The results of this is found in Table 3. SAS reports the Wilk's lamda Λ=0.41260, F=34.22; p < 0.0001; Wilk's lamda is ratio of two variance-covariance matrices. If the value of the statistics is too small, it indicates rejection of the null hypothesis. Here we reject the null hypothesis that there is no relationship between the two sets of variables, and can conclude that the two sets of variables are dependent. It is worth noting that the above null hypothesis is equivalent to testing the null hypothesis that all p canonical variate pairs are uncorrelated, or$H_o = p^*1 = p^*2 = p^*p* = 0$;Since the canonical correlations are ordered from the largest to smallest and since Wilk's lambda is significant, we can conclude that at least p*1 $\neq$ 0.We may also wish to test the null hypothesis that may be the second or the third canonical variate pairs are correlated. We can do this in successive tests. Next test whether the second and third canonical variate pairs are correlated.$H_o = p^*2 = p^*3 = p^*4 = 0$;Looking at the second row of table 3, the likelihood ratio test statistic Λ=0.7781; F=15.68, df=(9,1299.8); p < 0.0001. From the test we can conclude that the second canonical variate pair is correlated, p*2$\neq$ 0. Next, we can test the significance of the canonical variate $H_o = p^*3 = 0$.Third row of Table 3 shows the likelihood ratio test statistic Λ=0.99389; F=0.82; df(4,1070); p=0.5120. This is not significant, so we can conclude that the third canonical variate pair is not correlated. Similarly, the fourth row of the table shows the likelihood ratio test statisticΛ=0.99843; F=0.84; df=(1,536);p=0.3592. This is also not significant, so we can conclude that the fourth canonical variate pair is not correlated.Only the first two canonical variate pairs are significantly correlated and response on one another. This indicates that we would want to go ahead and summarize for two pairs.

Table 3:
Test of H0: The canonical correlations in the current row second and all that follow are zero

	Likelihood Ratio	Approximate to F Value	Num DF	Den DF	Pr > F
1	0.4126	34.22	16	1629	<.0001
2	0.7781	15.68	9	1299.8	<.0001
3	0.9938	0.82	4	1070	0.5120
4	0.9984	0.84	1	536	0.3592

Table 3a: Multivariate Statistics and F Approximations

S=4 M=-0.5 N=265.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.4126	34.22	16	1629	<.0001
Pillai's Trace	0.6929	28.08	16	2144	<.0001
Hotelling-Lawley Trace	1.1693	38.88	16	1060	<.0001
Roy's Greatest Root	0.8858	118.71	4	536	<.0001

Multivariate test statistics are presented in Table 3a. Bay far the most common method used is Wilk's lamda(λ). As it tends to have the most general applicability. In our example, the model was statistically significant, with a Wilk's lamda of 0.413, F=34.22, df=(16, 1629) and p< 0.0001. On the basis of this, we can reject the null hypothesis that there was no relationship between the variable sets and conclude that there probably was a relationship. Using Wilk's lamda, $1 - \lambda = 1 - 0.413 = 0.587 = r^2$ for the model. All other test statistics are also significant. This means that the model is significant.Now that we have tested the hypothesis of independence and have rejected them, the next step is to obtain estimates of canonical correlation. The estimated canonical correlations are reported in Table 4. The squared values of the canonical variate pairs, found in the last column, can be interpreted much in the same way as r^2 values are interpreted. We see that 46.97% of the variation in U1 is explained by the variation in v1, 21.7% of the variation in U2 is explained by V2, but only less than 1% of the variation in U3 is explained by V3. These first two are high canonical correlation and implies that only the first two canonical correlation are important.

Table 4:Canonical Correlation

	Canonical Correlation	Adjusted Canonical Correlation	Approximate Standard Error	Squared Canonical Correlation	Eigenvalues of Inv(E)*H = CanRsq/(1-CanRsq) Eigenvalue	Difference	Proportion	Cumulative
1	0.685376	0.680751	0.022819	0.469740	0.8859	0.6085	0.7576	0.7576
2	0.465955	0.461418	0.033690	0.217114	0.2773	0.2728	0.2372	0.9948
3	0.067427	.	0.042838	0.004546	0.0046	0.0030	0.0039	0.9987
4	0.039603	.	0.042966	0.001568	0.0016		0.0013	1.0000

The SAS output provides the estimated canonical coefficients (a_{ij}) for the natural disasters (ND) which are reported in the Table 5. Thus using the coefficient values in the first column, the first canonical variable for natural disasters can be determined using a formula.$U1=e^{0.7270lEventNo} - e^{0.0208lDeathNo} - e^{0.0545\ lPAffected} + e^{0.1831lDamageUS\$}$.The corresponding standardized canonical coefficients for ND variables are reported in Table 6. EventsNo and DamageUs$ have high positive correlations with ND1.We can't consider ND2 because we probably have multicollinearity issues.

Table 5: Raw Canonical Coefficients for the ND Measurements

	ND1	ND2	ND3	ND4
lEventsNo	0.7270	-0.4623	1.0868	-0.9407
lDeathNo	-0.0207	0.4097	-0.3930	-0.1874
lPAffected	-0.0545	0.1981	0.2121	0.3630
lDamageUs$	0.1830	-0.2427	-0.2378	0.1468

Table 6: Standardized Canonical Coefficients for the ND Measurements

	ND1	ND2	ND3	ND4
lEventsNo	0.6950	-0.4420	1.0390	-0.8993
lDeathNo	-0.0530	1.0442	-1.0017	-0.4778
lPAffected	-0.1724	0.6262	0.6704	1.1473
lDamageUs$	0.5936	-0.7871	-0.7712	0.4762

Table 7: Raw Canonical Coefficients for the ENV Measurements

	ENV1	ENV2	ENV3	ENV4
lOdp	0.08560	-0.04308	-0.39096	0.3781
lPm10	0.06088	1.59201	-0.02890	-0.0689
lForest_Km2	0.08006	0.00437	0.51847	0.4350
lCO2	0.16069	-0.03582	0.06301	-0.3981

Table 8: Standardized Canonical Coefficients for the ENV Measurements

	ENV1	ENV2	ENV3	ENV4
lOdp	0.2538	-0.1277	-1.1592	1.1212
lPm10	0.0391	1.0221	-0.0186	-0.0443
lForest_Km2	0.1577	0.0086	1.0212	0.8568
lCO2	0.6653	-0.1483	0.2609	-1.6483

To interpret each component, we must compute the correlation between each variable and the corresponding canonical variate. The correlations between the variables of natural disasters and the canonical variables are found inTable 9. Looking at first canonical variable for natural disasters, we see that all correlations are uniformly large. We had decided earlier not to look at the third and fourth canonical variate pairs. Second canonical variable has high correlations with all variables except lDamageUS$.

Table 9: Correlations Between the ND Measurements and Their Canonical Variables

	ND1	ND2	ND3	ND4
lEventsNo	0.8955	0.2576	0.3106	-0.1876
lDeathNo	0.6684	0.6859	-0.2606	-0.1216
lPAffected	0.6172	0.5242	0.2638	0.5241
lDamageUs$	0.8750	-0.0881	-0.3104	0.3610

\

Table 10: Correlations Between the ENV Measurements and Their Canonical Variables

	ENV1	ENV2	ENV3	ENV4
lOdp	0.8751	-0.0276	-0.3825	0.2951
lPm10	0.2449	0.9683	-0.0493	0.0039
lForest_Km2	0.7481	-0.0020	0.5387	0.3874
lCO2	0.9775	-0.0459	0.0212	-0.2047

Table 11: Correlations Between the ND Measurements and the Canonical Variables of the ENV Measurements

	ENV1	ENV2	ENV3	ENV4
lEventsNo	0.6138	0.1200	0.0209	-0.0074
lDeathNo	0.4581	0.3196	-0.0176	-0.0048
lPAffected	0.4230	0.2442	0.0178	0.0208
lDamageUs$	0.5997	-0.0411	-0.0209	0.0143

Table 12: Correlations Between the ENV Measurements and the Canonical Variables of the ND Measurements

	ND1	ND2	ND3	ND4
lOdp	0.5998	-0.0129	-0.0258	0.0117
lPm10	0.1678	0.4512	-0.0033	0.0002
lForest_Km2	0.5127	-0.0010	0.0363	0.0153
lCO2	0.6700	-0.0214	0.0014	-0.0081

Similar interpretation can take place with the environment variables. The correlation between the environment measurements and the canonical variables for environment variables are found in the Table 10. Since all correlations are large for the first canonical variable, this can be thought of as an overall measure of environmental variables as well, however, it is most strongly correlated with odp and CO2. Most of the correlations with the second canonical variable are small. There is some suggestion that this variable is highly positively correlated with PM10.Correlations between the ND measurements and the canonical variables of the ENV measurements are reported in Table 11. Since all correlations are large for the first canonical variable, it is most strongly correlated with lEventsNo and lDamageUS$.Correlations between the ENV measurements and the canonical variables of the ND measurements are reported in Table 12. Since all correlations are large for the first canonical variable, it is most strongly correlated with lodp and lCO2. Putting together, we see that the best predictor for natural disasters is CO2 and odp.

3.2 Univariate Regression Results

Univariate regression results are reported in Table 13. CO2 and PM10 appear to be significant factors in determining eventsNo, deathNo, PAffected and damageUS$. Odp was another important factor contributing to the eventsNo.

Table 13:Regression analysis for WITHIN CELLS error term

COVA RIATE	B	Beta	Std Error	T-value	Sig.of t		COVA RIATE	B	Beta	Std Error	T-value	Sig.of t
Response variable .. leventsNo							Response variable .. ldeathNo					
lCO2	.0852	.3642	.0246	4.944	.000		lCO2	.1110	.1921	.0471	2.357	.019
lodp	.0497	.1400	.0271	2.019	.044		lodp	.0693	.0789	.0674	1.028	.304
lforest	.0724	.1603	.0632	2.671	.008		lforest	.0654	.0585	.0741	.882	.378
lpm10	.1839	.1290	.0632	2.910	.004		lpm10	.1337	.3797	.1727	7.741	.000
Response variable .. ldamageUS							Response variable .. LpAffected					
lCO2	.3803	.5677	.0506	7.504	.000		lCO2	.1393	.1820	.0643	2.1663	.031
lodp	.0955	.0937	.0725	1.317	.189		lodp	.0163	.0140	.0920	.1774	.859
lforest	-.1428	-.1103	.0797	-1.790	.074		lforest	.1194	.0807	.1012	1.1797	.239
lpm10	.5374	.1317	.1858	2.892	.004		lpm10	1.6838	.3610	.2358	7.1382	.000

3.3LDA Results:

The minimum ratio of valid cases to predictor variables for LDA is 5 to 1. In this case, it is 541/4 $\approx$ 135 to 1, which satisfies the minimum requirement. It also does satisfy the preferred ratio of 20 to 1 (Table 13a).The number of cases in the smallest group in this problem is 157, which is larger than the number of predictor variables (4), satisfying the minimum requirement. In addition, the number of cases in the smallest group satisfies the preferred minimum of 20 cases (Table 13b).

Table 13a:Classification Processing Summary

Processed		555
Excluded	Missing or out-of-range group codes	0
	At least one missing discriminating variable	14
Used in Output		541

Table 13b:Prior Probabilities for Groups

Event_rank	Prior	Cases Used in Analysis	
		Unweighted	Weighted
1	.333	212	212.000
2	.333	157	157.000
3	.333	159	159.000
Total	1.000	528	528.000

Table 14:Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.32[a]	91.05	91.05	.75
2	.13[a]	8.95	100.00	.34

a. First 2 canonical discriminant functions were used in the analysis.

Table 15:Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.382	504.597	6	.000
2	.885	63.875	2	.000

In this analysis there were 3 groups defined by category of number of natural disaster events. Fourpredictor variables, so the maximum possible number of discriminant functions was 2. The canonical correlations for thedimensions one and two are 0.75 and 0.33, respectively (Table 14). In the table of Wilk's lambda which tested functions for statistical significance, the stepwise analysis identified 2 discriminiant functions that were statistically significant. The Wilk's lambda statistic for the test of function 1 through 2 functions (chi-square=504.60) had a probability of 0.000 which was less than the level of significance of 0.05. The Wilk's lambda statistic for the test of function 2 (chi-square=63.88) had a probability of 0.000 which was less than the level of significance of 0.05. The significance of the maximum possible number of discriminant functions supports the interpretation of a solution using 2 discriminant functions(Table 15).

Table 16 shows unstandardized canonical discriminant functions evaluated at group means. Function 1 separates the number of natural disasters events category 3(the negative value of 1.743) from number of natural disasters events category 1(positive value of 0.796) and category 2 (positive value of 0.691). Function 2 separates the number of events category 2(the positive value of 0.508) from events category 1(negative value of -0.360) and events category 3 (negative value of -0.021).Based on the structure matrix, the predictor variables strongly associated positively with discriminant function 1 which distinguished between events number categories are CO2(r=0.892). Based on the structure matrix, the predictor variable strongly associated positively with discriminant function 2 which distinguished between events number categories is pm10 (r=0.669). Other predictor variable strongly associated with discriminant function 2 which were strongly associated negatively with events number categories is forest area (r= -0.705) (Table 17). Using Wilk's lambda and step-wise LDA, the variables that minimizes the overall Wilk's lambda is entered. In our case, CO2, odp and forest area are significant (Table 18).The number of discriminant dimensions is the number of groups minus 1. However, some discriminant dimensions may not be statistically significant. In this example, there are two discriminant dimensions, both of which are statistically significant. The coefficients of linear discriminants are reported in Table 19. The equations of the linear discriminante function are:1) $discriminant\_score\_1=e^{0.458*lPM10} - e^{0.053*lforest\_km} + e^{0.976*CO2}$, 2) $discriminant\_score\_2 = e^{0.718*lpm10} - e^{0.830\overline{l}forest\_\overline{km2}} + e^{0.187*lCO2}$.

Table 16:Functions at Group Centroids

Event_rank	Function	
	1	2
1	.796	-.360
2	.691	.508
3	-1.743	-.021

Table 17:Structure Matrix

	Function	
	1	2
lco2	.892	-.352
lodp[b]	.593	-.255
lforest_km2	.494	-.705
lpm10	.340	.669

b. This variable not used in the analysis.

Table 18:Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
lodp	.619	162	2	525	.000
lpm10	.826	55.1	2	525	.000
lforest_km2	.721	101	2	525	.000
lco2	.484	280	2	525	.000

Table 19:Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
lpm10	.458	.718
lforest_km2	-.05	-.830
lco2	.976	.187

As you can see, the number of events categories 1 and 2 tend to be more at the lCO2 and lp10 (positive) end of dimension 1. The number of events category 3 tend to be at the opposite end in the dimension one. On dimension 2, all events number categories tend to be lower on forest area (Fig 10).

Figure 10:Canonical Discriminant Functions

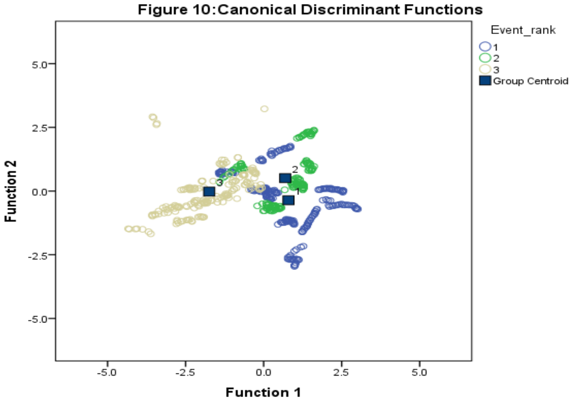


Table 20: Classification Results[a,c]

		Event _rank	Predicted Group Membership 1	2	3	Total
Original	Count	1	144	46	24	214
		2	50	91	17	158
		3	0	26	1...	169
	%	1	67.3	21	11	100
		2	31.6	58	11	100
		3	.0	15	85	100
Cross-validated[b]	Count	1	144	46	24	214
		2	51	86	21	158
		3	0	26	1...	169
	%	1	67.3	21	11	100
		2	32.3	54	13	100
		3	.0	15	85	100

a. 69.9% of original grouped cases correctly classified.

b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

c. 68.9% of cross-validated grouped cases correctly classified.

The cross validated accuracy rate computed by SPSS was 69.9% which was greater than the proportional by chance accuracy criteria of 41.6% (1.25*35.0=41.6). The criteria for classification accuracy is satisfied (Table 20). The proportional by chance accuracy rate was computed by squaring and summing the proportion of cases in each group from the table of prior probabilities for groups (0..333^2 + 0..333^2 + 0..333^2 =33.3).

Main assumptions of LDA and CCA are:1)MVN errors: The first assumption can be checked using Mahalanobis plot although symmetry is probably more important. If normality can not be induced by transformation or if the data are seriously non normal ie categorical, then the alternative of logistic regression should be used. It is worth pointing out that if all the assumptions are satisfied, lda is the optimal procedure and so should be used.

Figure.11:Normal Q-Q Plot for Multivariate Data

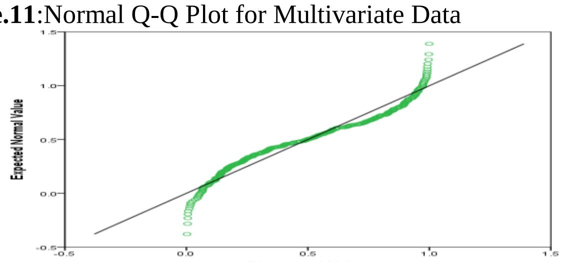


Table 21:Box's Test Results

Box's M		448.461
F	Approx.	37.056
	df1	12
	df2	1151849.499
	Sig.	.000

Tests null hypothesis of equal population covariance matrices.

The plot of ordered Mahalanobis distances against their expected values under the assumption of Multivariate Normality clearly shows slight deviation from the straight line. However, we conclude that the assumption of multivariate normality is approximately upheld (Figure 11).

2)Box's Test of Equality of Covariance Matrices:For the second assumption there is a test of equality of covariances matrices, Box's M test. Violation of this assumption can affect significance tests of classification results. The significance level can be inflated (false positives) when the number of variables is large and the sample sizes of the groups differ. Quadratic methods can be used if the covariance matrices are unequal but a large number of parameters are involved and lda is thus superior for small sample sizes. Overall lda is robust to both the assumption of MVN and equality of covariance matrices, especially if the sample sizes are equal. The formal hypothesis for Box's M test for Equality of covariance would be: $H_0: \sum 1 = \sum 2 = \sum 3, H_a: \sum 1 \neq \sum 2 \neq \sum 3$

$\alpha = 0.05$, $Fobs = \frac{MS_{Regression}}{MS_{Residual}}$, Reject H_0 if p-value <0.05.Tests null hypothesis of equal population covariance matrices.

Do not reject H0 as p-value = 0.000> 0.05

Test Statistic

$$M = \sum n_i \ln|s| - \sum_{i=1}^{k} n_i \ln|s_i|$$

$$C^{-1} = 1 - \frac{2p^2+3p-1}{6(p+1)(k-1)} \left(\sum_{n-1}^{k} \frac{1}{n_i} - \frac{1}{\sum n_i}\right)$$

Sampling Distribution

$$MC^{-1} \sim \frac{\chi^2\,(k-1)(p)(p+1)}{2} \; if\ k, p < 5\ and\ n_i \approx 20 \text{ else } F \text{ distribution}$$

To test the assumption of Equality of co-variances, we use Box's M-test. If the Box's M Test shows p <.05, the covariances are significantly different and the null hypothesis is NOT rejected. If the Box's M Test shows p >.05, the covariances are not significantly different and the null hypothesis is rejected. The value of Box's M is 448.46, with a p-value of 0.000, indicating that the assumption of equal co-variances is not satisfied and null hypothesis is not rejected (Table 21). So the assumption of homoscedasticity is violated. That is we do not reject the null hypothesis of $H_0: \sum 1 = \sum 2 = \sum 3$.Thus, the assumption of multivariate normality is satisfied but the assumption of equality of covariance matrices is not satisfied. In this case, we have used quadratic discriminatory function approach. The estimates of quadratic discriminate function show that the total error rates and the error rate in each group are all smaller than the rate if assigned randomly (69.9%), which indicates that quadratic discriminate function can be properly used to discriminate the event groups.It is found that total error rate without cross-validation classification is slightly lower than that computed from cross-validation method.

3.4 Multinomial Logistic Regression Results

We also ran multinomial logistic regression using the variables used in LDA. Here, we see model fit is significant, $\chi 2$ (8)=461.96, p<0.001. which indicates our full model predicts significantly better, or more accurately, than the null model (Table 22).Both the Pearson and Deviance statistics are chi-square based methods and here, we interpret lack of significance as indicating good fit(Table 23).Higher values of Pseudo R-square indicate better fit(Table 24).The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.We can see from the table that the predictors such as CO2, pm10 and forest area displays a significant ch-square which indicates that odp can be dropped from the model(Table 25).The Wald test (and associated p-value) is used to evaluate whether or not the logistic coefficient is different than zero. We can see that a one unit change in odp or forest area do not significantly change the odds of being classified in the third category of the outcome variable relative to the first and second categories of the outcome variable while controlling for the influence of the other predictors(Table 26). Logistic regression also satisfy main assumptions of the model such as linearity, independence of errors and absence of multicollinearity.

Table 22 Model Fitting Information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	1149.40			
Final	687.44	461.96	8.00	.00

Table 23:Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	699.23	1046.00	1.00
Deviance	687.44	1046.00	1.00

Table 24:Pseudo R-Square

Cox and Snell	.58
Nagelkerke	.66
McFadden	.40

Table 25:Likelihood Ratio Tests

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	914.08	226.65	2	.00
lco2	786.21	98.77	2	.00
lodp	691.17	3.73	2	.15
lpm10	759.28	71.85	2	.00
lforest_km2	710.49	23.06	2	.00

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

Table 26:Parameter Estimates

Event_rank[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp (B)	
								Lower Bound	Upper Bound
1	Intercept	-28.233	3.356	70.789	1	.000			
	lco2	.697	.097	51.667	1	.000	2.008	1.661	2.429
	lodp	.136	.093	2.129	1	.145	1.146	.954	1.376
	lpm10	2.337	.483	23.382	1	.000	10.350	4.014	26.689
	lforest_km2	.206	.160	1.661	1	.198	1.229	.898	1.683
2	Intercept	-28.415	3.373	70.964	1	.000			
	lco2	.744	.097	58.309	1	.000	2.104	1.738	2.547
	lodp	.180	.094	3.635	1	.057	1.197	.995	1.441
	lpm10	3.208	.490	42.854	1	.000	24.728	9.464	64.611
	lforest_km2	-.250	.155	2.610	1	.106	.779	.575	1.055

a. The reference category is: 3.

4. CONCLUSION

Using CCA technique and Wilk's lamda, we have seen that the first two canonical correlations are significant which shows that the two sets of variables, namely, the natural disasters variables and the environment variables are highly correlated. This has been validated by all other test statistics such as Pillai's trace, Hotelling-Lawley trace and Roy's greatest root.First two canonical correlationsare high which implies that only the first two canonical correlation are important.Correlations between the ENV measurements and the canonical variables of the ND measurements show that all correlations are large for the first canonical variable and it is most strongly correlated with odp and CO_2. Putting together, we see that the best predictor for natural disasters is CO_2 and odp. Univariate regression results show that CO_2 and PM10 appear to be significant factors in determining eventsNo, deathNo, PAffected and damageUS$. Odp was another important factor contributing to the eventsNo. The effect of forestation or deforestation shows positive effect on natural disasters. This means the current level of forestation is not sufficient to counter natural disasters. However, the LDA approach clearly identified two functions required for events number. LDA and multinomial logistic regression have clearly identified the positive influence of CO_2 and PM10 and negative influence of deforestation on natural disaster events. In order to reduce natural disasters, policies to reduce the emission of CO_2 and PM10 pollution are required as well as the efforts to foster forestation, afforestation and reforestation.

5. ACKNOWLEDGEMENT

I'm very grateful to the anonymous Referees for the evaluation of my paper and for the constructive critics.

6. REFERENCES

[1]Charles W.Freeman III, Amy Searight, "The Politics of Climate Change in Asia", in "Asia's Response to Climate Change and Natural Disasters:Implications for an Evolving Regional Architecture", A Report of the CSIS ASIAN Regionalism Initiative, 2010, pp.10

[2]Charles W.Freeman III, Michael J.Green, "Introduction" in "Asia's Response to Climate Change and Natural Disasters:Implications for an Evolving Regional Architecture", A Report of the CSIS ASIAN Regionalism Initiative, 2010, pp.8

[3]Jason,Anderson and Camilla, Bausch, "Climate Change and Natural Disasters: Scientific Evidence of a Possible relation between recent natural disasters and climate change". IP/A/ENVI/FWC/2005-35, 2006.

[4]NASA, http://earthobservatory.nasa.gov/Features/RisingCost/rising_cost5.php

[5]Teresita, Schaffer, "Asian Regional Institututions and Climate Change" in "Asia's Response to Climate Change and Natural Disasters:Implications for an Evolving Regional Architecture", A Report of the CSIS ASIAN Regionalism Initiative, 2010, pp.44

[6]UNESCAP, online statistics