International Journal of Earth, Energy and Environmental Sciences
ISSN: 2517-942X
Vol:4, No:3, 2010

Statistical Analysis-Driven Risk Assessment of Criteria Air Pollutants: A Sulfur Dioxide Case Study

Ehsan Bashiri

Abstract**—**A 7-step method (with 25 sub-steps) to assess risk of air pollutants is introduced. These steps are: pre-considerations, sampling, statistical analysis, exposure matrix and likelihood, dose-response matrix and likelihood, total risk evaluation, and discussion of findings. All mentioned words and expressions are well-understood; however, almost all steps have been modified, improved, and coupled in such a way that a comprehensive method has been prepared. Accordingly, the SADRA (Statistical Analysis-Driven Risk Assessment) emphasizes extensive and ongoing application of analytical statistics in traditional risk assessment models. A Sulfur Dioxide case study validates the claim and provides a good illustration for this method.

Keywords**—**Criteria air pollutants, Matrix of risk, Risk assessment, Statistical analysis.

I. Introduction

ENVIRONMENTAL experts usually refer to risk assessment of environmental pollutants and the likes by distinguishing between four steps: *Hazard Identification, Dose-Response Assessment*, *Exposure Assessment*, and *Risk Characterization*. Due to the popularity and applicability of risk assessment among all environmental-relevant industries and companies, various models has been produced and developed.

One of the well-known methods is *Matrix of Risk* [12]. As its name suggests, this model consists of a few matrices to interpret and visualize the actual risk in a risky workplace. Matrix of risk aims to determine level of risk of hazards and significance of potential loss, and establish a framework for controls. One probable defect for this model is its severe dependence on qualitative characteristics of agents.

Other typical risk assessment methods are: *Failure Modes and Effects Analysis (FMEA), Fault Tree Analysis, Event Tree Analysis, Probabilistic Risk Assessment,* and *Hazard and Operability (HAZOP) Study* [4]. Although some of them are not environment-applicable, each contains rigid logic to develop in one or more domains of technology. Besides, on account of the quantitative nature of environmental agents, statistical analysis is highly recommended. This is due to the reliability and accuracy of mathematical concepts.

E. Bashiri, Petroleum University of Technology (www.put.ac.ir), Abadan Institute of Technology, Abadan, IRAN (phone: 631-4420050; e-mail: ebsh.ait@gmail.com). Home Address: Hanivan Town, Moin 3, 69319-96469 Ilam, Ilam, IRAN (phone: 841-3349519).

In the model described in this paper, some different techniques are integrated so that a trustable and concise method will be introduced.

II. Procedure

Approach developed here requires collection of variety of site and problem specific information. This approach has classified into seven steps and twenty five sub-steps as listed below:

A. *Pre-Considerations*

a. Type and Population of Industry or Factory
b. Worker’s Overall Psycho-social Status and Recreational Facilities
c. Regional Climate
d. All Existing Pollutants or Chemicals in Workplace

B. *Sampling*

e. Criteria Pollutant
f. Diffusion of Contaminant at Workplace
g. Sampling Instruments and their Location
h. Number of Samples
i. Sampling Method by Using Two-stage Approach
j. Samples Gathered

C. *Statistical Analysis*

k. Statistical Characteristics of Samples Gathered
l. Standard Value of Agent Based on EPA Reference
m. Confidence Interval for Population Based on Two-Sided Hypothesis Test

D. *Exposure Matrix and Likelihood*

n. Average Number of Hours per Day for People at Workplace
o. Number of Days per Week for Work
p. Number of Weeks per Year for Work
q. Finding Exposure Likelihood

E. *Dose-Response Matrix and Likelihood*

r. Carcinogenesis Classification of Contaminant
s. Potential Health effect of Agent
t. Likelihood of Occurrence or Loss
u. Finding Dose-Response Likelihood

F. Total Risk Evaluation

v. Establishing an Equation to Find Total Risk
w. Quantifying Impact Factors
x. Evaluating Total Risk

G. Discussion of Findings

y. Conclusions and Discussions of Results

III. Pre-considerations

This model, as the first priority, requires being wary of some preliminary heeds about human, workplace, climate, and other relevant industrial considerations. These characteristics play significant roles for dependant variables which predominate over risk assessment evaluation.

Size and type of industry, units' functioning, and input and output chemicals or substances that involve in operations are prior items of our survey. This information provides a complete overview of what researchers are confronted with and expected to handle. During assembling this information a list of all environmental agents in workplaces will be discovered spontaneously which cause to pollute the area. Whereas each agent has an individual influence on surrounding, some chemicals share impacts with each other bilaterally and advance reactivity of others, and so, heighten the loss. Two major and common chemicals that come to mind are Ozone and Nitrogen Dioxide.

In addition to the physical properties of industry, a lot of people including employees, employers, auditors, engineers, and managers are attending at workplace who must be cared. These professionals demand to provide dozens of supplies and facilities in and out of the workplace in order to both protecting them from injury or loss and preparing a suitable shelter for welfare and recreation. Nowadays *Psycho-Social Disorders* possess an unbelievable percentage of cause of physical and mental diseases that contribute to the worst industrial disasters [3]. This will also decrease people's threshold of tolerance and increase their vulnerability to pollutants. Some significant psycho-social problems are:

- Work overload and time pressure
- Lack of social support from supervisors or coworkers
- Too little or too much responsibility
- Lack of status rewards (appreciation)
- Discrimination or harassment
- Lack of support for work/family balance
- Lack of respect for employees and the work they do

Thus, an independent and authorized auditor would be required to measure a valid percentile level of psycho-social satisfaction of people by means of standard statistical methods. Different organizations might have different mechanisms to measure this index.

Final pre-consideration of this model refers to the geographical region in which the industry is located at. Depending on the operation used and pollutants produced, the weather condition may either relieve or even worsen spread of contaminants in environment. An ideal clean industry is one which produces fewer pollutants and locates at a susceptible and favorable climate.

IV. Sampling

Following previous section, all hazardous substances at a workplace and typical impact of climate on their release to the atmosphere could be identified. Here, subsequently, the most dominant agent, namely *Criteria Pollutant*, should be recognized. Essential and necessary knowledge prior to establishing a reliable set of samples is described in following paragraphs:

A. How the contaminant has been scattered at workplace?

This can be viewed to occur in various forms: *Random*, *Uniform (Homogeneous)*, *Patchy*, *Stratified* (homogeneous without sub-areas), and *Gradient* [6]. Each of these mechanisms is explained in environmental engineering and science textbooks, and purpose of this section is not to master them. Usually, uniform diffusion of pollutants is an ideal case.

B. What are sampling instruments? And, where they can be located on sampling?

There are several advanced instruments to collect samples. Considering the kind of pollutant, accuracy of data, and industry's budget restrictions, one standard direct-reading set could be utilized. However, researchers should be cautious regarding position of these instruments. The location can be chosen *Randomly*, *Systematically*, or in a *Stratified Random Manner*, depending on spread of the pollutant, size of industry, or number of samplers available [6].

C. What is permissible number of samples?

There are two fundamental concerns for appointing number of samples: maximum tolerable error, and economical limits. With an acceptable economical state, the number of samples is given by the following relationship:

$$n = (\frac{tS}{E})^2 \quad (1)$$

Where t is test value, S^2 the sum of the squares, and E maximum tolerable error.

S^2 is calculated by:

$$S^2 = \sum (x_i - x)^2 / (n-1) \quad (2)$$

These values are almost accurate and valid while applying *Central Limit Theorem* for *Standard Normal Distribution*. Of course, this condition itself demands to assign a standard value for n that is 30. By this value, standard normal distribution is accessible and tolerable error will be justifiable [11].

D. How to gather sample observations?

Here one proposed systematic sampling called *Two-Stage Method* is given to collect 30 samples. It should be noticed that this is only a proposed method, not a mandatory case. Each individual or group can apply one of the various standard methods that are widespread in industries.

1) First Stage: sixteen samples in one workday, two workshifts

Sixteen samples from sixteen consecutive hours of both morning and evening workshifts of a random operational workday will be observed and recorded (i.e. at 6am, 7am, 8am, 9am, 10am, 11am, 12am, 1pm, 2pm, 3pm, 4pm, 5pm, 6pm, 7pm, 8pm and 9pm of a day). Although it is optional to choose one day, it is preferred to select a day at which maximum numbers of personnel attend in workplace.

2) Second Stage: fourteen samples in one week at random hours of both workshifts

A same procedure will be kept on to collect fourteen samples, but during particular hours of both workshifts for the days of a week as shown in Table I.

V. Statistical Analysis

After collecting an acceptable number of observations,

TABLE I
2ND STAGE OF PROPOSED SAMPLING METHOD

Day	Morning Shift Work Time	Evening Shift Work Time
Sunday	7am	2pm
Monday	8am	3pm
Tuesday	9am	4pm
Wednesday	10am	5pm
Thursday	11am	6pm
Friday	12am	7pm
Saturday	1pm	8pm

statistical processing commences. At first, statistical characteristics for this set of sample data is derived, i.e. mean, variance, and standard deviation. Then, referring to standard, maximum tolerable concentration or size of agent in ambient air for any specified duration of exposure will be obtained.

The main focus of this section begins with performing a test, *Two-sided Hypothesis Test*, to clarify whether or not the population's mean (arithmetic mean of all the observations during a period of one year) is greater, or smaller than the standard value. This is in fact a representation of confidence likelihood for the population.

Consider following two-sided hypothesis test:

$$1) H_0 : \mu \le \mu_0$$
$$2) H_1 : \mu > \mu_0 \qquad (3)$$

Here, μ_0 is the standard value. Since the population's variance is unknown, *Student's t Distribution* is applied. Assuming first trial to be true, following relationship should be settled:

$$T = \frac{\overline{X} - \mu_0}{\frac{s}{\sqrt{n}}} < k = t_{1-\alpha}(n-1) \qquad (4)$$

Here, t is a statistic which has student's t distribution with (*n-1*) degrees of freedom, *k*, $(1-\alpha)^{th}$ quantile, α meaningful area, $\overline{X}$ sample's mean, *s* sample's standard deviation, and *n* number of samples.

Using this test, reader is able to get the confidence interval (likelihood) for this hypothesis that will be utilized in deriving total risk relationship described in section VIII. This is also used to infer how this might influence the potential health of people extended in section VII.

VI. Exposure Segment

In this section exposure duration to pollutants at workplace will be discussed. It can be categorized as: hours per day, days per week, and weeks per year. Although only one exposure measure could be mentioned, i.e. hours per year, the following *Exposure Matrix* provides precise values; since it takes into account *Acute-Intermediate-Chronic* exposure to agents in various time intervals that have different influences on human.

TABLE II
EXPOSURE MATRIX

Category \ Value	6	5	4	3	2	1
Hours/Day (hr/d)	7 to 8	6	5	4	3	1 to 2
Days/Week (d/w)	6 to 7	5	4	3	2	1
Weeks/Year (w/yr)	47 to 52	42 to 47	37 to 42	32 to 37	27 to 32	<27

With a reasonable induction, multiplication is utilized to derive *Exposure Likelihood* as below:

$$\text{Exposure Likelihood} = (\text{Multiplication of Corresponding Values of Categories}) / 6^3 \qquad (5)$$

VII. Dose-Response Segment

As its name hints, dose-response refers to the unfavorable response of human to the dose and exposure duration of toxicant. Here, also, three categories are given as: Carcinogenesis Classification, Potential Health Effect, and Likelihood of Occurrence. *The Dose-Response Matrix* below demonstrates logical values for any of situations.

TABLE III
DOSE-RESPONSE MATRIX

Category \ Value	6	5	4	3	2	1
Carcinogenesis Classification (Toxicity) [a]	A	B_1	B_2	C	D	E
Potential Health Effect [b]	death	permanent disability	serious injury	moderate injury	minor injury	no injury
Likelihood of Occurrence [c]	very likely	likely (probable)	rare (seldom, but possible)	very rare (very seldom, but possible)	unlikely (slight possibility)	practically impossible

[a]Based on accumulated evidences, classified by EPA [10]
[b]Based on derived average concentration and "acute-intermediate-chronic" exposure to agent, approved by standards
[c]Based on industry's periodic observations and annual reports

Again, trusting induction, but instead, using summation, *Dose-Response Likelihood* is measured as below:

Dose-Response Likelihood = (Summation of Corresponding Values of Categories) / 18 (6)

VIII. TOTAL RISK

Up to now, three probabilistic quantities are derived, each regarding one specific aspect of risk of environmental pollutants. The main purpose of this paper is to discover a distinct relationship between these values to evaluate overall risk. One possible outcome might be the arithmetic mean. This is a special case however, and can not be generalized to all conditions; since there may be other parameters which can reinforce exposure or dose-response likelihoods. These parameters are needed because although they are not involved in matrices directly, they make effects on measuring which can not be ignored. In general, a comprehensive formula written below may be applied:

$$K = \alpha(\text{Exposure Likelihood}) + \beta(\text{Dose - response Likelihood}) + \gamma(\text{Confidence Likelihood})$$
$$TR = [\frac{K}{(\alpha+\beta+\gamma)}] \quad (7)$$

Where, α, β, γ are called *Impact Factors*. Assuming impact factor of confidence likelihood to be 1 ($\gamma = 1$), other factors are derived by following equations:

$$\alpha = [1 + (\frac{\text{Number of People Exposed to Agent}}{\text{Number of All People}}) + (\frac{\text{Number of Days with Critical Weather Conditions in a Year}}{\text{365 day}})]$$
$$\beta = [1 + (\text{Employee Dissatisfaction Percent}) + (\text{Availability of Exacerbating Chemicals or Substances})]$$

$$\alpha, \beta, \gamma \geq 1 \quad (8)$$

These terms are chosen based upon experiments and experiences, and may be modified by other organizations, depending on their own preferences or situations. Considering both terms of α and β, it is transparent to realize that they affect real amount of exposure and dose-response likelihoods indirectly.

All portions have been clarified in section III (pre-considerations), and can be determined easily. However, the second term of β (exacerbating chemicals and compounds) might be to some extent confusing. This quantity relates to the properties like reactivity, and similar structure or effect of exacerbating agents, and will be evaluated by dividing their cumulative dose to the criteria pollutant's dose. For instance, water, natural gases, or other SOx compounds are exacerbating materials for Sulfur Dioxide.

IX. SULFUR DIOXIDE CASE STUDY

A. Samples

To support the claim, the Sulfur Dioxide case study as a criterion air pollutant in *Ilam Gas Refinery Co.* is to be scrutinized. This company located at *Ilam Province*, west of *Iran*, in a mountainous region with almost cold climate. 30 recorded samples according to the proposed method described in section IV are gathered and listed in Table IV. These observations are recorded by an isolated direct-reading instrument located near the operational units and reported directly by the *HSE* (Health, Safety, and Environmental) manager of the unit. He was asked to prepare observations exactly according to the method. It is also assumed that release of pollutants to the atmosphere is homogeneous.

These values are recorded for personnel who work at ambient area around the operational unit. In other words, it is supposed that a worker attend his workplace at the entire workshifts period for at least 1 year. However, these values are smaller in comparison with those for managers or head officers who are not directly and regularly confront to chemicals.

TABLE IV
MEASURED VALUES OF SULFUR DIOXIDE SURVEY

No. of sample	First Stage	Second Stage	Date (July 2009)	Time	Measured value (ppm)
1	*		2	6am	0.02
2	*		2	7am	0.024
3	*		2	8am	0.025
4	*		2	9am	0.028
5	*		2	10am	0.03
6	*		2	11am	0.031
7	*		2	12am	0.031
8	*		2	1pm	0.032
9	*		2	2pm	0.032
10	*		2	3pm	0.032
11	*		2	4pm	0.031
12	*		2	5pm	0.03
13	*		2	6pm	0.026
14	*		2	7pm	0.022
15	*		2	8pm	0.02
16	*		2	9pm	0.015
17		*	3	7am	0.025
18		*	3	2pm	0.033
19		*	4	8am	0.025
20		*	4	3pm	0.031
21		*	5	9am	0.026
22		*	5	4pm	0.03
23		*	6	10am	0.028
24		*	6	5pm	0.031
25		*	7	11am	0.031
26		*	7	6pm	0.025
27		*	8	12am	0.032
28		*	8	7pm	0.023
29		*	9	1pm	0.031
30		*	9	8pm	0.02

EPA Standard (Annual): 0.025 ppm during any 8-hour workshift of a 40-hour workweek

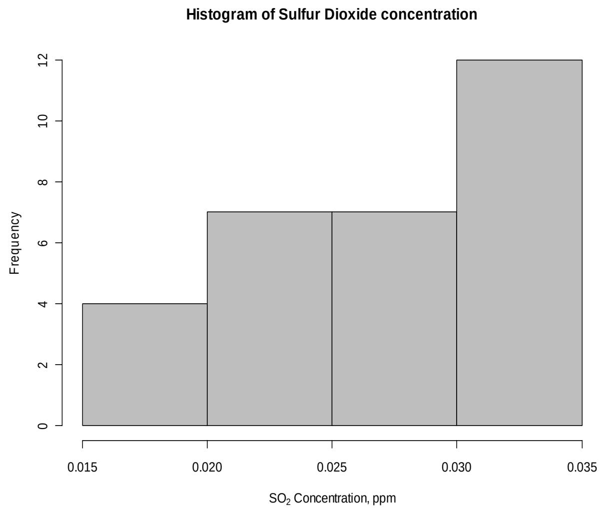

Fig. 1 Histogram of Sulfur Dioxide concentration

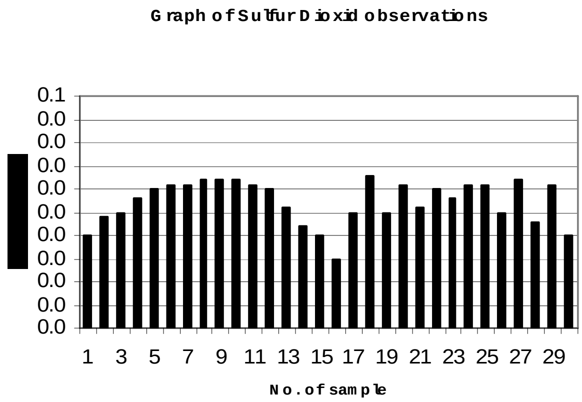

Fig. 2 Graph of Sulfur Dioxide observations

As shown in Fig. 2, Sulfur Dioxide graph resembles a normal distribution for a workday (i.e. from input no.1 until input no.16). Therefore, corresponding graph of values collected for consecutive hours of both workshifts in a day is nearly distributed normally. Readers should not be worry about remaining 14 inputs whose corresponding histogram does not have a regular form; because they are due to the way introduced for recording observations in 2nd stage of part D of section IV.

Summary of statistical characteristics of these samples is provided in Table V below.

TABLE V
STATISTICAL CHARACTERISTICS OF SULFUR DIOXIDE SAMPLES

Characteristic	Value
Mean:	0.027333
Standard Deviation:	0.004671
Max.:	0.033
Min.:	0.015
Standard Value:	0.025

B. *Confidence Interval*

As derived: s = 0.004671, μ_0 = 0.025, and $\overline{X}$ = 0.02733. These yields to:

$$T = \frac{\overline{X} - \mu_0}{\frac{s}{\sqrt{n}}} = 3.2$$

Referring to student t distribution's cumulative values for 29 degrees of freedom, it is permissible to raise the amount of confidence up to 0.995, i.e. $\alpha = 0.005$. For this value k = 2.7 and $t > k$. So, the null hypothesis is rejected and the alternative is accepted.

Therefore it is confidently accepted that $\mu > \mu_0$ with at least 99.5% confidence. This also prepares a confidence interval for the estimate of μ .

C. Matrices

Tables VI and VII contain corresponding values of each matrix for this case study.

TABLE VI
EXPOSURE MATRIX OF CASE STUDY

Category \ Value	6	5	4	3	2	1
Hours/Day (hr/d)	≥ 7					
Days/Week (d/w)		5				
Weeks/Year (w/yr)		42 to 47				

Exposure Likelihood = (150/216) = 0.694

TABLE VII
DOSE-RESPONSE MATRIX OF CASE STUDY

Category \ Value	6	5	4	3	2	1
Carcinogenesis Classification (Toxicity) [a]					D	
Potential Health Effect [b]				moderate injury		
Likelihood of Occurrence [c]			rare (seldom, but possible)			

Dose-Response Likelihood = (9/18) = 0.5

D. Total Risk

The HSE manager of organization (concerning about confidentiality of the details), reported corresponding values for impact factor terms as below:

$$\alpha = [1 + (\frac{\text{Number of People Exposed to Agent}}{\text{Number of All People}}) + (\frac{\text{Number of Days with Critical Weather Conditions in a Year}}{\text{365 day}})]$$
$$= 1 + 0.75 + 0.4$$
$$= 2.15$$
$$\beta = [1 + (\text{Employee Dissatisfaction Percent}) + (\text{Availability of Exacerbating Chemicals or Substances})]$$
$$= 1 + 0.6 + 0.5$$
$$= 2.1$$
$$\gamma = 1$$

So, based on (7), total risk is:

$$TR = [\frac{K}{(\alpha + \beta + \gamma)}]$$
$$= \frac{2.15(0.7) + 2.1(0.5) + 1}{5.25}$$
$$\approx 0.68 \text{, or } 68\%$$

X. DISCUSSIONS OF FINDINGS AND CONCLUSION

A 68% risk asserts that an illness or any health loss will occur to a person with 68% probability during one year. As a consequence, it may be expected that almost seven out of ten persons will be "injured moderately" due to unfavorable effects of Sulfur Dioxide during one year (based on dose-response matrix). This result also acknowledges the annual report of the HSE office regarding likelihood of occurrence that was ranked 4 out of 6 (67%). It must be emphasized that for this experiment a 100% probability does not signify that all people will die gradually. In fact, it shows that all people are in a danger of maximum harm effect of agent on body in specific concentration that is "moderate injury".

According to this model, in order to determine total risk of an agent, three different indices should be considered altogether. For instance, as this case study implied, the dose of a toxicant during a particular period was higher than what standard permits. This provided a confidence interval for estimate of mean concentration for population. However, it could not be inferred from this individual criterion to judge the total risk; because exposure and dose-response criteria were available and must be taken into account. Then short, average, and long duration of exposure were measured and toxicity, potential effect, and likelihood of occurrence for the pollutant were determined. It is even understood that each index itself has a different impact on the value of overall risk, thus, the indices could not be weighted equally. This leaded to the advent of impact factors. Finally the overall risk by integrating them in the form of one unified equation was evaluated. This result can be generalized to the same situations, of course with different probable interpretations.

In conclusion, SADRA leaded to prediction of valid, accurate, and convincing risk values for criteria air pollutants. This is fundamentally because of the statistical inference and analysis that has been applied to almost all elements of this model. This paper suggested a scenario to measure total risk of criteria air pollutants by integrating some useful and widespread methods. Without assistance of Statistics, this model deals only with individual subjective or objective aspects of risk alone, that of exists in traditional risk assessment models.

ACKNOWLEDGMENT

I have to appreciate Dr. David K. Stevens (Department of Civil & Environmental Engineering, University of Utah State) for his worthwhile advices and comments to revise this paper technically. I also thank Mr. Majid Sharifi for his kind attention to check grammar errors and inconsistencies.

REFERENCES

[1] Alberico, D., Bozarth, Js. (1999) Software System Safety Handbook, A Technical & Managerial Team Approach. Joint Services Computer Resources Management Group.

[2] Bartoszynski, R. (2007) Probability and Statistical Inference 2nd Ed. New Jersey: John Wiley & Sons Inc.

[3] Burton, Joan. (2006) Creating Healthy Workplaces. USA: Industrial Accident Prevention Association.

[4] Clemens, Pat L., Simmons, Rodney J. (1998) System Safety and Risk Management. Ohio: U.S Department of Health and Human Services.

[5] Cheremisinoff, Nicholas P. (2002) Handbook of Air Pollution Prevention and Control, USA: Butterworth Heinemann, Elsevier Science.

[6] Dean, John R. (2003) Methods for Environmental Trace Analysis. England: John Wiley & Sons Ltd.

[7] Freeman, H.M. (1989) Standard Handbook of Hazardous Waste Treatment and Disposal. New York: McGraw Hill.

[8] Kreyszig, E. (1988) Advanced Engineering Mathematics 6th Ed. John Wiley & Sons Ltd.

[9] Leveson, Nancy G. (2002) System Safety Engineering: Back to the Future. Massachusetts Institute of Technology: Aeronautics and Astronautics.

[10] Masters, Gilbert M. (1991) Introduction to Environmental Engineering and Science. New Jersey: Percentile-Hall International Inc.

[11] Mood, Alexander McF. (1991) Introduction to the Theory of Statistics. Mashhad, Iran: Mashhad University Press.

[12] OSSA Ins. (2008) Workplace Hazard Analysis Form, December 2000, Product No.130-001-001-D-GE, http://www.ossa.com/?Filaename=130001001DGE.pdf. accessed 19 May 2008.

[13] Ridley, J., Channing, J. (2003) Safety at Work 6th edition. London: Butterworth-Heinemann.

[14] Stephan, RA. (1997) System Safety Analysis Handbook 2nd Ed. New Mexico: System Safety Society.

[15] Stewart, Mark G, Melchers, Robert E. (1997) Probabilistic Risk Assessment of Engineering Systems. London: Chapman & Hall.

[16] U.S Department of Defense. (1993) System Safety Program Requirements. Washington DC: U.S. Department of Defense.

[17] U.S. Department of Health and Human Services. (1998) Toxicological Profile for Sulfur Dioxide. Atlanta, Georgia 30333.

[18] Weiner, Ruth F., Mathews, Robin A. (2003) Environmental Engineering Fourth Ed. USA: Elsevier Science.

[19] Whyte, W. (1991) Cleanroom Design. London: John Wiley & Sons Ltd.