Iran J Public Health, Vol. 45, No.9, Sep 2016, pp.1188-1198

Original Article

Systemic Approach for Health Risk Assessment of Ambient Air Concentrations of Benzene in Petrochemical Environments: Integration of Fuzzy Logic, Artificial Neural Network, and IRIS Toxicity Method

*Vahid NOVIN [1], *Saeed GIVEHCHI [2], Hassan HOVEIDI [2]*

1. *Safety and Environment, University of Tehran, Graduate Faculty of Environment, Tehran, Iran*
2. *Faculty of Environment, University of Tehran, Tehran, Iran*

***Corresponding Author:** Email: Givehchi@ut.ac.ir

(Received 16 Dec 2015; accepted 20 Mar 2016)

Abstract

Background: Reliable methods are crucial to cope with uncertainties in the risk analysis process. The aim of this study is to develop an integrated approach to assessing risks of benzene in the petrochemical plant that produces benzene. We offer an integrated system to contribute imprecise variables into the health risk calculation.

Methods: The project was conducted in Asaluyeh, southern Iran during the years from 2013 to 2014. Integrated method includes fuzzy logic and artificial neural networks. Each technique had specific computational properties. Fuzzy logic was used for estimation of absorption rate. Artificial neural networks can decrease the noise of the data so applied for prediction of benzene concentration. First, the actual exposure was calculated then it combined with Integrated Risk Information System (IRIS) toxicity factors to assess real health risks.

Results: High correlation between the measured and predicted benzene concentration was achieved (R^2= 0.941). As for variable distribution, the best estimation of risk in a population implied 33% of workers exposed less than 1×10^{-5} and 67% inserted between 1.0×10^{-5} to 9.8×10^{-5} risk levels. The average estimated risk of exposure to benzene for entire work zones is equal to 2.4×10^{-5}, ranging from 1.5×10^{-6} to 6.9×10^{-5}.

Conclusion: The integrated model is highly flexible as well as the rules possibly will be changed according to the necessities of the user in a different circumstance. The measured exposures can be duplicated well through proposed model and realistic risk assessment data will be produced.

Keywords: Risk assessment, Exposure estimation, Benzene, Cancer risk, Fuzzy logic, Neural network

Introduction

Petrochemical production is categorized as a high-risk industry. One of the important risks is the existence of high levels of gasoline vapors including benzene that during work shifts employees are exposed to. Employees in functional units are more vulnerable population because of their continuous contact with harmful substances such as benzene. Based on USEPA (United State Environmental Protection Agency) and IARC (International Agency for Research Cancer) reports, benzene has been classified to be a group A and Class 1 human (1).

Adverse health effects of benzene appear in two different type; short-term and long-term. Short-term effects are associated with high concentration and may involve headaches, dizziness, distraction and defects temporary memory and tremors. Whereas, exposure to benzene in a long-term is connected to intricate adverse health effects such as immunological, hemato-toxicity,

Geno-toxicity, adverse effects on reproductive organs, and as well as various cancers (1, 2).
Therefore, a powerful method is required to predict existent exposure to harmful substances besides to evaluate probable adverse effects (3). Although several mathematical models have been carried out so far, there is high complexity related to chemical exposure in terms of human health. Due to the lack of information about chemical impacts on human particularly in long-term impacts, some unpredictable factors that called uncertainties can affect health risk assessment (4, 5).
Health risk assessment through inhalation concerning benzene emission from equipment at the unit of benzene production was accomplished. This project is an example of using a hybrid approach that can incorporate uncertainties to assessment's process. Major variables affecting the absorption of chemicals moreover key parameters in the dispersion of chemicals can be reflected in a hybrid system which combining fuzzy logic and neural networks (5, 6).
In this research" average inhalation and absorption rates were used for assessment. However, breathing rates are affected by many individual characteristics, including age, sex, weight, health, and level of physical activity (running, jogging, etc.)" (5, 9).
Perhaps the first and distinguished use of fuzzy sets in health risk assessment was on the application of fuzzy logic in the environmental risk assessment (10). One more example which connected to human health risk assessment is the application of fuzzy sets in human health risk assessment. Fuzzy sets were employed to estimate carcinogenic risk caused by air pollution in Ten Russian cities (11).

Methods

Study area

The project was carried out since 2013 to 2014. The study area was in the northern part of the Persian Gulf with a distance of 10 miles from the coast. It was categorized as an intermediate petrochemical throughout the country. The plant mainly produced about one million tons annually of benzene, paraxylene, orto-xylene used to produce ethylbenzene, styrene, cyclohexane, and nitrobenzene. Benzene from refinery streams was typically produced from catalytic reformats pyrolysis, gasoline, and toluene de-alkylation. This study focused on tasks and activities those are exposed to potential contact with benzene during refinery.
Main workplaces related to workers exposed to benzene such as loading of tankers were listed in Table 1. Each workplace maybe had a different group of job and activities. In order to increase the information about jobs, activities, tasks, and places a questionnaire was prepared then accomplished through an interview with employees.

Table 1: Main work places for workers

Work areas	Situation	Number of workers	Sex
Loading terminal	Out door	18	M
Process site	Out door	42	M
Tank units	Out door	17	M
Control room	In door	8	M
Laboratory	In door	6	M

Sample collection

The sampling procedure was based on absorption of benzene in an active charcoal tube (active sampling) (12). SKC model 222 pumps have been utilized for gas sampling. The Glass Tubes with a 6mm external diameter, 4 mm internal diameter and 70 mm height, containing activated charcoal holder with a restrictive orifice (separated by a 2-mm part of urethane foam) were installed. The pump was adjusted to work for 30 min at a flow rate of 100 ml/min (12).
Active sampling was accomplished in workdays. For 8 h the measurements were recorded on a daily source (08:00-16:00 Local Time). Four measuring procedures were performed lasting one week in the middle of each season (12, 13). Consequently, by the end of the week 70 samples were taken and during 15/2/2012 to 21/09/2013 totally 280 samples were collected. for detailed

information stickers attached on each tube to tag sampling number, the time of end duration of sampling, the pump number, the humidity, the wet and dry temperature, and the date of sampling (13).

The human health risk Model

The human Lifetime Average Daily Dose (LADD) equation for a single chemical exposure proposed by U.S.EPA (2011) that related to cancer risk model (presuming that the inhalation is the only route of intake) is, as shown in Eq. [1] (14):

LADD = CA.IR.EF.ED/BW.AT [1]

The Cancer Risk (CR) is calculated for exposure to benzene using Eq. [2] (15):

Cancer Risk = LADD (μg/kg/day) × SF $(\mu g/kg/day)^{-1}$ [2]

For equation 1, CA is the concentration of a chemical in an exposure medium ($\mu g/m^3$), IR represents the inhalation rate (m^3/h), EF is the exposure frequency (number of working days per year), ED is the exposure duration (working years), BW is the body weight (kg), AT is averaging time (AT=70 yr 365 d/year for carcinogens), LADD shows lifetime average daily dose (μg/kg/day) [14] and SF in equation 2, is the cancer slope factor of benzene (linear low-dose cancer potency factor) (15).

The quality of data in equation 1 is uncertain even if definite exposure-related measurements are available for variables (16). To explain more, some unpredictable parameters can change the results. As an illustration, movement of contaminants among environmental media changes remarkably the quality of data, therefore, uncertainty plays a crucial role involved in each variable as well as affected by physical and chemical factors (11, 16). The concentration of the air pollutant is derived as a function background concentration. These variables; wind speed, ambient temperature, humidity and rainfall cause fluctuation in concentration of air pollutants (17). Furthermore, the inhalation rate is also influenced by multiple individual aspects including age, body weight, and amounts of physical activity (5, 11).

Owing to these numerous variables, the trustworthy approach is required to make a better risk evaluation. Therefore, to overcome the problems of uncertainties the hybrid method was prepared (5, 18). The construction of the system is displayed in Fig. 1 (5). In this figure, each tablet represents a sub-system combined; they determine the chemical absorption rate through inhalation. The receptor description (Inhalation rate) and dose estimation sections use a set of fuzzy rules and obtain the average daily inhalation dose, based on fuzzy inference. The exposure prediction section consists of a designed neural network using a new back-propagation algorithm. This section created in order to calculate of real ambient concentration. In the following sections, details about the subdivisions will be explained (5,18).

Using fuzzy logic for estimation of inhalation rate

Standard quantities technique in health issues has been often vague and unclear adjectives such as a little, too much…; so at the start it is necessary to quantify these adjectives and change them into fuzzy sets (19, 20). In this section, 36 fuzzy rules were used to determine the inhalation rate of adults. Age, body weight, and activity tend to be the significant reasons of daily energy expenditure in healthy people for as long as is also in energy balance (20, 21).

The triangular membership function is selected for age. Variables must be delineated with linguistic values rather than numerical values (5, 21). Fig. 2 shows one of the prepared fuzzy sets (fuzzy numbers) that related to age with three linguistic values. The horizontal vector indicates the size of the parameter and the vertical vector represents the degree of membership (degree of dependence) of each value (5).

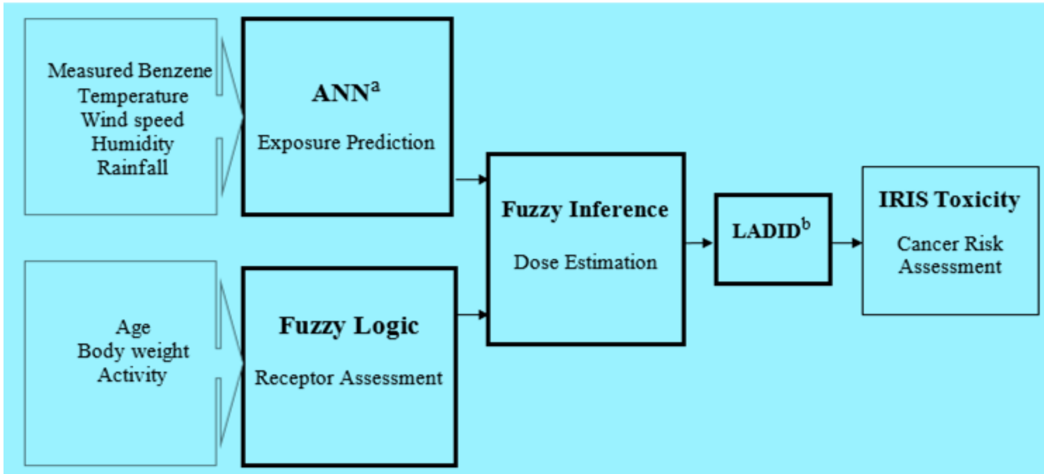

Fig. 1: Integrated system of risk assessment (5)
a. Artificial Neural Network
b. Life Lifetime Average Daily Inhalation Dose

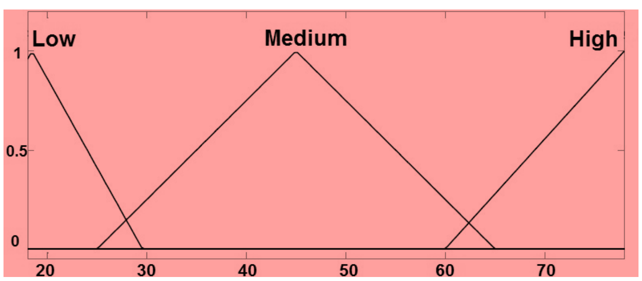

Fig. 2: Membership functions of the variable "Age" (5)

The membership function for age was set to range from 18 to 80 yr (22). Using census and demographic details, age (in years) was divided into three fuzzy linguistic sets: low, ranging from 18 to 30; medium, ranging from 25 to 64; and high, ranging from 59 to 80 (5, 22).
According to Brainard and Burmaster (23), body weight (45–110 kg) follows a normal distribution. It has been divided into three linguistic sets: low (from 45 to 68 kg), medium (from 56 to102 kg), and high (from 90 to110 kg). Fuzzy set was made for physical activity by means of four linguistic values and follows skewed (normal) distribution (5).
As for long-term physical activity levels (PALs), it is ranging from 1.2 to 2.5 times the BMR (Basal Metabolic Rate), where 1.2 represents the minimal intensity of activity and 2.5 shows a very physically active lifestyle. Aging is a parameter that can remarkably decline high-intensity activity (21). Moreover, seven linguistic values ranging from 5-25 m^3/day with normal distribution have been provided to show inhalation rate, which is compatible with the human capability (14).

Artificial neural network for estimation of exposure to benzene

ANNs are working based on specified transfer function and made by artificial neurons (17). There are neuron networks so-called weight factors or simply weights. These networks are composed of neurons that interact with each other. Networks or connections transmit signals from other neurons. Signals contain either positive or negative weights. Using particular function ANNs can adjust the value of different weights properly.
In order to neural network learning, the data should be split into two subsets; training and testing sets. There is no specified rule to define the size of each subset. Hence, the data set for this project were randomly divided into a ratio of 3:1 between training and testing sets, respectively (7, 10, 17). As a result, available data (280 samples) arbitrarily fragmented into two subsets. One of

those was training groups and included 75% of data. By developing the model this group predicted the concentration of benzene for the second group that contained the 25% of data (10, 17).

Parameters; wind speed, temperature, relative humidity, rainfall and measured concentration was chosen as model input data. The predicted benzene concentration was the output parameter. Details of measurement devices for each meteorological parameter are shown in Table 2.

Proposed ANN designed in three layers. Input parameters located in the first layer in the form of five neurons. To gain best results experimentally founded to put 10 neurons in the second layer. The third layer involved just one neuron which is output parameter and shows actual benzene concentration. A typical three-layer neural network is shown in Fig. 3 (17).

Table 2: Details of the meteorological monitoring devices

Parameter	Reaction time	Procedure
Wind speed	1s	3-cup anemometer
Temperature	10 s	Temperature sensor
Humidity	15 s	Sensing element
Rainfall	3 s	Tipping bucket rain gauge

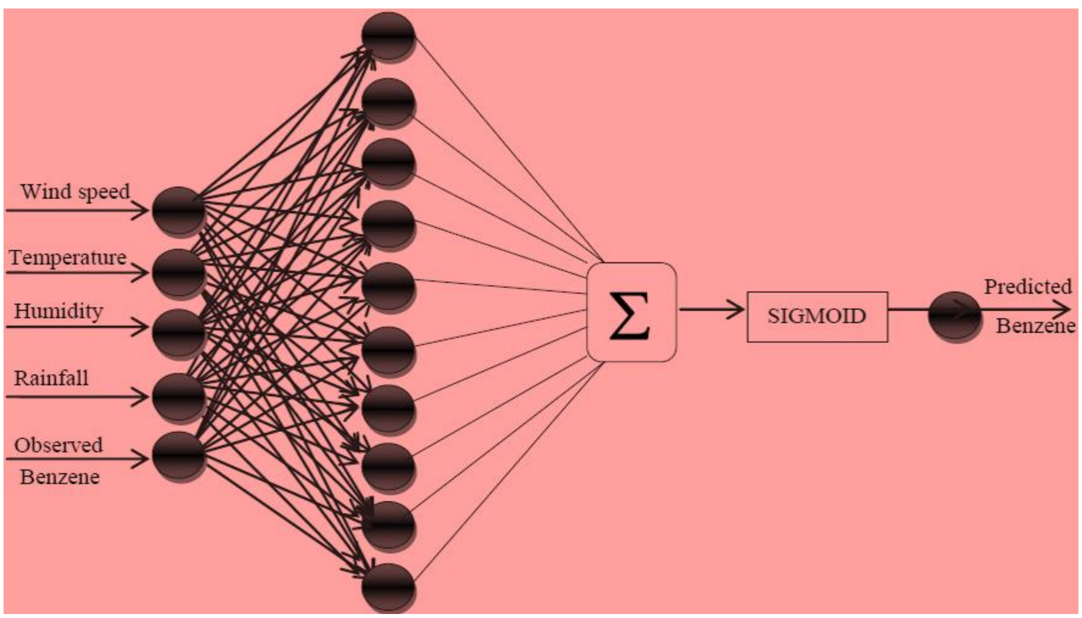

Fig. 3: Selected three layer neural network (17)

Before conducting the network preparation procedure, a training group (75% of the data) consisting of 210 cases had to be prepared from the environmental data done by the back propagation paradigm. Training group itself became divided into training set and testing set that this process conducted randomly. First one is the training set consisted of 90% of the data or 189 cases used to train the model. The second or final 10% contains 21 cases were left out as a test set (7, 17, 24).

Fuzzy inference calculation for LADD estimation

Output variables from the artificial neural network stage and fuzzy logic section contributed to the last part of the hybrid model and produced a chronic daily intake dose (18). The predicted benzene concentrations lie between 0 and 67 (μg/m^3) considered very high. Although the global average is 0 to 15μg/m^3 (25), the peak concentrations in the study area are high and the membership function for output and some fuzzy rules must be adjusted to range 0 and 67 (μg/m^3). The process of adjustment was easily done that indicates a powerful feature of the dose assessment modeling system as well as the adaptability to different scenarios. In different circumstances, the user can change fuzzy rules according to the different situations (6, 9).

The outputs from the benzene prediction divided into six linguistic variables based on the frequency distribution of the data. The membership function of variable concentration is shown in Fig. 4 (18). The output from the receptor block was divided by the body weight to derive m^3/d/kg then contributed as input to the dose estimation section (5, 18). The daily intake rate for the dose estimation is expressed by four linguistic fuzzy sets and is expressed in m^3/d/kg (18). Finally, the life average daily intake dose defined as seven fuzzy linguistic values and ranging from 0 to 6 μg/kg/d.

A sample fuzzy inference calculation for the dose estimation block is shown in Fig. 5. Where IR denotes inhalation rate normalized for body weight (m^3/day/kg); CC indicates chemical concentration (μg/m^3); LADID = Lifetime Average Daily Inhalation Dose (μg/kg/d). As an illustration, if the inhalation rate were considered 0.35 m^3/kg/d, the ambient benzene concentration is 15 μg/m^3, lifetime average daily inhalation dose would be 1.63 μg/kg/day.

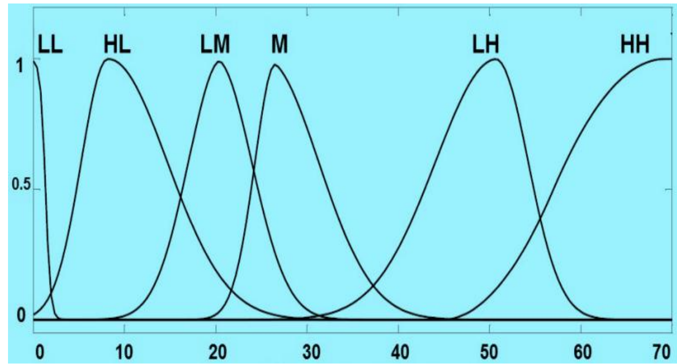

Fig. 4: Input variable "predicted concentration (μg/m^3)"

The output from this last module gives the overall dose through inhalation for each individual around the site, taking into account both the exposure factors and the receptor factors in a quantitative way, based on neural network and fuzzy inference. When the exposure is quantified, combined with Integrated Risk Information System (IRIS) toxicity factors, then as a final point the related risk can be estimated.

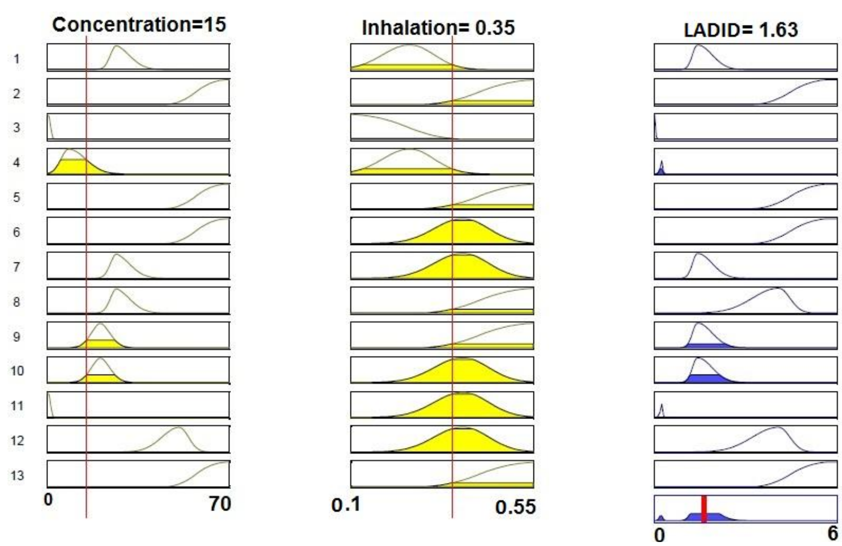

Fig. 5: Fuzzy inference for estimation of LADID (Matlab software; fuzzy toolbox)

Results

Artificial network validation and results

In this section, Leave-One-Out Cross-Validation (LOOCV) was followed with the aim of training and testing the ANN model. Frequently, one sample is kept for testing while the rest is used for training up to all samples are finally tested (26). Before the proposed model is applied to the particular application, it must be trained using all available samples (27). The difference between the observed and the predicted values are shown in Fig.6.

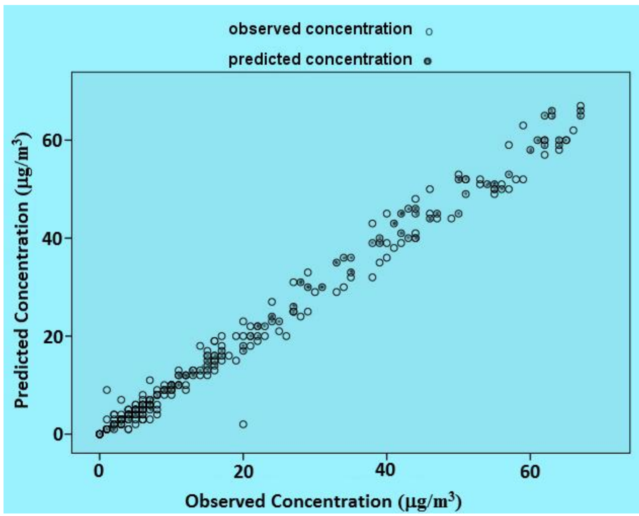

Fig. 6: Comparison of observed and predicted concentration (Source: Author's calculations)

The training of network continued until maximum correlation within the measured and predicted output was achieved (Table 3). Correlation expressed by R-squared that R^2 is coefficient of multiple determinations and relative root mean square error (RMSE) (26). Correlation results are perfect when an R-squared value of 1, a very good fit is next to 1 and a very poor fit less than 0. On the other side, how much the value of RRMSE is smaller; the performance of the model is better.

Table 3: ANN evaluation parameters (Source: author's calculations)

Pollutant	Structure of ANN	Evaluation parameters	Results
Benzene	5-10-1	Mean Absolute Error	5.9
	5-10-1	Minimum Absolute Error	0
	5-10-1	Maximum Absolute Error	13.1
	5-10-1	RMSE (g/m^3)	4.8
	5-10-1	R^2	0.941

Predicted concentration results

In Fig. 7, the annually averaged results of passive sampling are presented. Exposure values in terminals were generally higher than another area. There is seasonal variation, especially in outdoor units. Ambient air temperature affects significantly the exposure levels, as a result, the high temperatures in the summer and spring explains the increased benzene concentrations in exposure values. When temperatures are very low (wintertime), the exposure levels are less than usual for an equivalent quantity of close benzene level (Fig. 8). The presence of wind reduces exposure levels, especially to employees who are performing outdoor activities. No strong correlation found between wind speed and exposure levels of employees working in laboratory and control room.

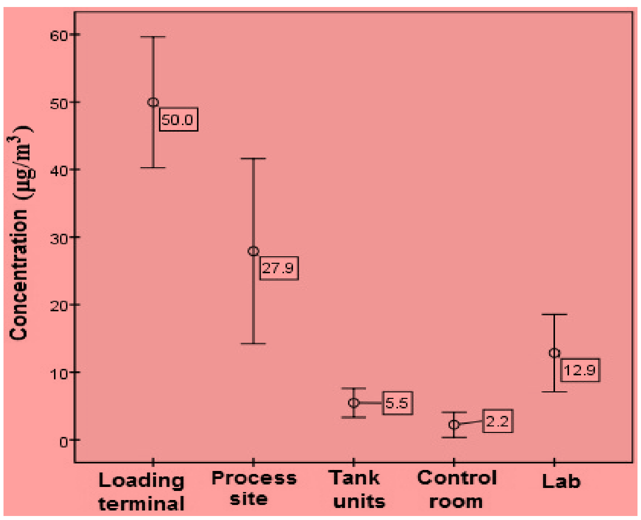

Fig. 7: Annually average concentration (Source: Author's calculation)

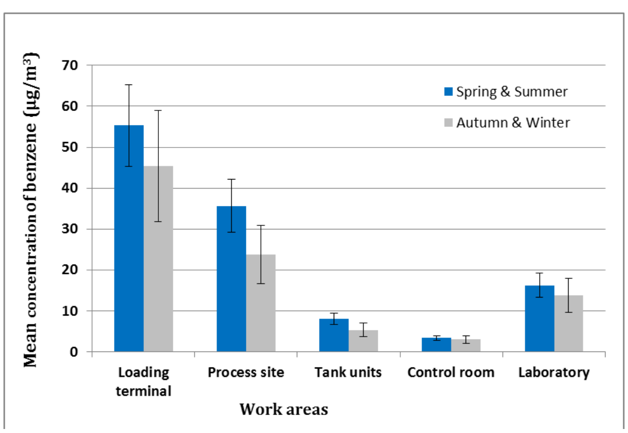

Fig. 8: Seasonal comparison of ambient concentration of benzene (Source: Author's calculations)

Health risk assessment results

The measured benzene concentrations and the related exposure levels are not considered as an acute health risk; meanwhile, this is the possibility of leukemia existence after chronic occupational contact with benzene due to measured concentration. Long-term individual risk will be calculated by combining USEPA integrated risk information system (IRIS) toxicity features with results from exposure assessment section (18). The individual excess CR was calculated for exposure to benzene through inhalation is presented in Table 4. The best estimate for the variable distribution of risk in a population implied 33% of people exposed less than 1×10^{-6}. Based on EPA Clean Air Act Risk Range, 1×10^{-6} risk range is considered as the most health protective end of the range (27, 28). Whereas 1×10^{-5} is the midpoint of risk range and 67% of the population is ranging from 1×10^{-5} to 9.8×10^{-5} cancer risk probability (27, 28).

The average estimated risk for all work areas considering exposure to benzene is equal to 2.4×10^{-5}, ranging from 1.5×10^{-6} to 6.9×10^{-5}. The results counsel potential cancer risk for period exposure to benzene within the numerous areas but at different levels (Fig. 9). These differences result from differences in the employee's type of activity, age, weight, and breathing rate that in common methods of risk assessment they did not pay attention to them.

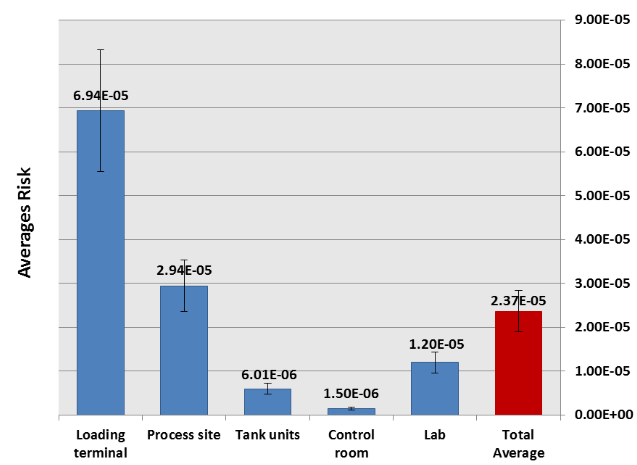

Fig. 9: Potential cancer risk

Table 4: Individual health risk assessment (based on average exposure obtained by passive sampling)

Average Life Time Cancer Risk Value	Frequency (Number of workers)	%	Average Life Time Cancer Risk Value	Frequency (Number of workers)	%
1.4×10^{-6}	3	3.3	2.7×10^{-5}	3	3.3
3.0×10^{-6}	3	3.3	2.8×10^{-5}	4	4.4
4.3×10^{-6}	2	2.2	2.9×10^{-5}	3	3.3
5.2×10^{-6}	3	3.3	3×10^{-5}	3	3.3
6.1×10^{-6}	3	3.3	3.2×10^{-5}	3	3.3
6.5×10^{-6}	2	2.2	3.4×10^{-5}	3	3.3
6.8×10^{-6}	4	4.4	3.5×10^{-5}	5	5.5
7.5×10^{-6}	2	2.2	3.6×10^{-5}	7	7.7
8.1×10^{-6}	3	3.3	4.2×10^{-5}	3	3.3
8.4×10^{-6}	2	2.2	6×10^{-5}	3	3.3
9.1×10^{-6}	3	3.3	7×10^{-5}	2	2.2
1.6×10^{-5}	3	3.3	8.2×10^{-5}	2	2.2
2.1×10^{-5}	3	3.3	9.5×10^{-5}	1	1.1
2.2×10^{-5}	5	5.5	9.8×10^{-5}	4	4.4
2.4×10^{-5}	4	4.4	**Total**	**91**	**100 %**

Discussion

The complete conception in establishing an integrated product is a dynamic model for risk estimation, able to evaluate probable interactions among the levels of benzene exposure through different task designs, those lead to different levels of benzene metabolic rate and subsequently to discount risk estimations (29, 30). In this model, risk factors that can influence inhalation rate were age, body weight, and physical activity of persons. These parameters considered as the input variable of fuzzy logic in order to estimate exposure level for the various situation of employees. The hybrid system also used Artificial Neural Network approach to predict actual atmospheric benzene concentration. ANN can effectively forecast the changes in measured atmospheric benzene concentration. Major variables affecting pollutant concentration were wind speed, ambient temperature, humidity, and rainfall.

The model is highly flexible so can easily accommodate any situation. Expert viewpoint is used as fuzzy rules and the rules can be changed according to the needs of the user. The produced composite model shows promise as a new tool for chemical exposure and health risk assessment, which it allows multiple uncertainties incorporate into health risk assessment.

Conclusion

In this project, the new hybrid dose model has been able to replicate a measured exposure prospering believed to be more representative of actual intake dose than data from previously accepted methods.

Overall, according to aforementioned reasons this system can produce risk assessment data that appear realistic. It may also be extended to other risk management applications where multiple uncertainties exist.

Ethical considerations

Ethical issues (Including plagiarism, Informed Consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the author.

Acknowledgments

The authors would like to thank the Editor and anonymous reviewers for their helpful comments and suggestions. No fund was received for conduction this study. The authors declare that there is no conflict of interests.

References

1. Sorahan T (2011). Occupational benzene exposure and lymphoma risks. *Environ Health Perspect*, 119 (11): a468.
2. Navasumrit P, Chanvaivit S, Intarasunanont P, Arayasiri M, Lauhareungpanya N, Parnlob, V, Ruchirawat M (2005). Environmental and occupational exposure to benzene in Thailand. *Chem Biol Interact*, 153-154: 75–83.
3. Badjagbo K, Loranger S, Moore S, Tardif R, Sauvé S (2010). BTEX exposures among automobile mechanics and painters and their associated health risks. *Hum Ecol Risk Assess*, 16(2):301–316.
4. Zhang X, Huang GH (2011). Assessment of BTEX-induced health risk under multiple uncertainties at a petroleum-contaminated site: An integrated fuzzy stochastic approach. *Water Resour Res*, 47(12): W12533.
5. Rajkumar T, Guesgen HW (1997). Analysis of chemical exposure through inhalation using hybrid neural network. *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*.
6. Kartalopoulos SV (1995). *Understanding Neural Networks and Fuzzy Logic*. IEEE Press, Los Alamitos, pp.: 165-283
7. Yan Chan K, Jian L (2013). Identification of significant factors for air pollution levels using a neural network based knowledge discovery system. *Neurocomputing*, 99: 564–569.
8. Jo WK, Pack KW (2000). Utilization of breath analysis for exposure estimates of benzene associated with active smoking. *Environ Res*, 83(2): 180–187.
9. Matbouli YT, Hipel KW, Kilgour DM, Karray F (2014). A fuzzy logic approach to assess, manage, and communicate carcinogenic risk. *Hum Ecol Risk Assess*, 20 (6):1687–1707.
10. Sarigiannis DA, Karakitsios SP, Gotti A, Papaloukas CL, Kassomenos PA, Pilidis GA (2009). Bayesian algorithm implementation in a real time exposure assessment model on benzene with calculation of associated cancer risks. *Sensors (Basel)*, 9(2):731–755.
11. Reshetin VP (2008). Fuzzy assessment of human-health risks due to air pollution. *Int J Risk Assess Manag*, 9 (1/2):160.
12. Karakitsios SP, Kassomenos PA, Sarigiannis DA, Pilidis GA (2009). Exposure modeling of benzene exploiting passive–active sampling data. *Environ Model Assess*, 15(4): 283–294.
13. Król S, Zabiegała B, Namieśnik J (2011). Measurement of benzene concentration in urban air using passive sampling. *Anal Bioanal Chem*, 403(4):1067–82.
14. USEPA (2011). *Exposure factors handbook 2011 edition (Final)*. U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-09/052F, chapter 4, pp.: 32-45
15. USEPA (2005). *Guidelines for carcinogen risk assessment, risk assessment forum*. Technical Report. Washington, DC, USA, chapter 3, pp.: 5-32
16. Vesely WE (2011). *Probabilistic Risk Assessment*. System Health Management: With Aerospace Applications, pp.: 253–263.
17. Cai M, Yin Y, Xie M (2009). Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach. *Transport Res D-Tr E*, 14(1): 32–41.

18. Rajkumar T, Guesgen HW, Robinson S, Fisher GW (2000). A new dose model for assessment of health risk due to contaminants in air. *J Air Waste Manag Assoc*, 50(1): 3-20.
19. Mofarrah A, Husain T (2011). Fuzzy based health risk assessment of heavy metals introduced into the marine environment. *Water Qual Expo Health*, 3(1): 25–36.
20. Kentel E, Aral MM (2004). Probabilistic-fuzzy health risk modeling. *Stoch Envir Res Risk Ass*, 18 (5): 324–338.
21. Stifelman M (2007). Using doubly-labeled water measurements of human energy expenditure to estimate inhalation rates. *Sci Total Environ,* 373(2-3): 585–590.
22. Meijer GAL, Westerterp KR, Van Hulsel AMP, Ten Hoor F (1992). Physical activity and energy expenditure in lean and obese adult human subjects. *Eur J Appl Physiol Occup Physiol,* 65(6): 525–528.
23. Brainard J, Burmaster DE (1992). Bivariate distributions for height and weight of men and women in the United States. *Risk Anal*, 12(2): 267–275.
24. Olden JD, Joy MK, Death RG (2004). An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecol Model*, 178(3-4): 389–397.
25. Wallace L (1996). Environmental exposure to benzene: an update. *Environ Health Perspect*, 104(Suppl 6): 1129–1136.
26. Witten IH, Frank E, Hall MA (2011). *Credibility. Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann Publishers is an imprint of Elsevier, pp.: 147–187.
27. "Clean Air Act | US EPA." U.S. Environmental Protection Agency. N.p., 19 Dec. 2008. Web. 9 Aug. 2016. Available from: *http://www.epa.gov/air/caa/*
28. Raun LH, Marks EM, Ensor KB (2009). Detecting improvement in ambient air toxics: An application to ambient benzene measurements in Houston, Texas. *Atmos Environ*, 43(20): 3259–3266.
29. Fontes T, Silva LM, Pereira SR, Coelho MC (2013). Application of artificial neural networks to predict the impact of traffic emissions on human health. *Progress in Artificial Intelligence, Springer-Verlag Berlin Heidelberg,* LNAI 8154, pp.: 21–29.
30. Elangasinghe MA, Singhal N, Dirks KN, Salmond JA (2014). Development of an ANN-based air pollution forecasting system with explicit knowledge through sensitivity analysis. *Atmos Pollut Res*, 5(4):696–708.