JRUCS

Journal of AL-Rafidain University College for Sciences

AL- Rafidain University College

Structural Equations Modeling in an Applied Study to Analyze the Impact of Environmental Health on the Human Health Index in Basrah Governorate

Asmaa A. Yaqoob	Lamiaa A. Dawod
asmaa.yaqoob@uobasrah.edu.iq	lamya.abd@mtu.edu.iq
Department of Statistics - College of Administration and Economics - University of Basrah, Basrah, Iraq.	College of Health and Medical Technologies, Middle Technical University, Baghdad, Iraq.

Article Information

Correspondence:
Asmaa A. Yaqoob
asmaa.yaqoob@uobasrah.edu.iq

Abstract

Pollution and disruption of the ecosystem are serious problems for human life, especially since the rate of pollution and environmental degradation is constantly increasing. Therefore, it is necessary to study the interaction between the environment and health by analyzing the impact of some environmental factors (climate change, air quality, environmental awareness, toxic drinking water and sanitation systems, noise....etc) and how they affect the health of the human. In this research, structural equation models (SEM) were used to construct causal models that study the relationships between variables at one time, especially in the presence of latent variables, where each of the environmental factors included in the study were evaluated and their contribution to the calculation of the humans health index was evaluated. The information was collected by making a questionnaire that was distributed to different groups of residents of the Basrah city in Iraq. Through the results, it was found that environmental health has a high positive effect on the health index ,therefore, it is necessary to pay attention to environmental health and reduce the sources of environmental pollution.

1. Introduction

Statistical analysis has been an essential tool for social science researchers for more than a century, and the applications of statistical approaches have expanded dramatically with the advent of computers and software and the adoption of technology as a tool for communicating knowledge. Researchers initially relied on univariate and binary analysis to understand data and relationships, but the more complex relationships made it is necessary to rely increasingly on more sophisticated methods of multivariate data analysis that analyze multiple variables simultaneously. Hence the idea of SEM, which deals with multiple variables in the case of vague variables that are difficult for us to measure directly, is called latent variables. In this research, the SEM methodology was used to analyze the relationship of environmental health to human health, which is a branch of public health

that deals with all environmental elements, whether natural or established, that affect human health. Environmental health is closely related to environmental sciences and public health, as it deals with factors affecting human health. Environmental health deals with all physical, chemical, and biological factors outside the human body, in addition to all factors that affect behavior and actions. Environmental health includes the assessment and control of those environmental factors that have the potential to affect health. Environmental health seeks to prevent disease and create an environment that supports health, in addition to analyzing the effects that are often produced indirectly and affect human health and the health of the environment in general, whether physical, psychological, social or cultural. In the modern era, new environmental problems have emerged, different from those previously known, and the most important previous studies related to the subject can be summarized as follows:

(Emmanual,et.al.,2013) They studied structural equations modeling to analyze the effects of traffic and its association with air pollution and its impact on public health, including a decrease in heart rate, and the Bayesian method was used in the estimation process.(Yi,et.al.,2016) The basic components and variables of the structural equation and the synthesis of common problems in SEM applications were presented, where it was explained how to apply SEM in environmental research, good results were reached, and studies proved that this method has enormous potential to serve environmental scientists. (Abbas,et.al.,2017) SEM was applied in the environment to study the problems of environmental sustainability, as he made a reference review of all studies related to the subject for the period from 2005 to 2016, while providing the most important instructions for the proper application of SEM. (Jerico and Venusmar,2021) In this research, the relationship of the economic cost and the effects on public health of environmental pollution was studied based on the awareness of the respondents through structural equations modeling. (Jose and Contreras,2022) They studied the relationship between means of transportation and their impact on the environment, due to the great pollution caused by means of transportation in the air.

2. Materials and methods

2.1.Observable and Latent Variables

Observed variables are those that can be measured directly, which are usually more prevalent and dealt with by researchers in various fields such as weight, height, etc., where the observed variables provide data as a primary source of information for statistical analysis [11,p.3]. However, sometimes the researcher may deal with concepts such as trends, marital status, obesity, and others. Such concepts are not directly measurable. Therefore, one of the approved approaches is to measure them indirectly by a set of indicators or through a mathematical model of observational variables that act as an alternative to that concept. The latent environment is analyzed through a variable called the latent variable that is it is meaningful but hidden and cannot be noticed [4, p.36]. And that the use of several elements to measure one concept, this procedure requires reducing the measurement error as much as possible, which is the difference between the true value of the variable and the value obtained from the measurement.

2.2. The Concept of Structural Equation Modeling

The structural equations model is a powerful multivariate tool to study the interrelationship between the observed and latent variables. This method has wide applications in psychological, educational, social, medical and biological research. SEM can be used appropriately in the analysis of complex phenomena that include relationships between one or more of the independent latent variables and one or more of the dependent latent variables, that is, it contributes to verifying causal models between the variables of the phenomenon more comprehensively [6]. One of the advantages of this method is that it deals with variables taking into account measurement errors, unlike the classical methods. Therefore, SEM estimates are more accurate, as they give unbiased results for the estimated parameters by finding unidirectional effects and mutual effects [2].

The general objective of SEM is to test the hypothesis that the observed covariance matrix for a set of measured variables is equal to the covariance matrix included in a hypothetical model and according to the following formula [6]:

$$\Sigma = \Sigma(\lambda) \tag{1}$$

Where Σ represent the covariance matrix of a set of observed variables, $\Sigma(\lambda)$ represent the covariance matrix implied by (λ), λ the vector of parameters of model, which determines the shape of x by specifying the means, variances, and regression parameters.

There are some assumptions that must be made before applying SEM, as it is assumed that the data set taken from the parent population is normally distributed, in addition to that the joint distribution of the variables that represent the data follows the multivariate normal distribution. Since the method of least squares cannot display latent variables, therefore we need to use structural equations for analysis to model the relationship between variables and analyze them at one time. If these conditions are met, the standard estimation method in SEM, which is used to estimate parameters and standard errors, is often Maximum Likelihood (ML).

2.3. Components of The Structural Equation

The basic SEM model consists of two main components, the first is the Confirmatory Factor Analysis (CFA), which is the model that combines the observed variables with the latent variables, taking into account the measurement error, as it is used to verify the structural validity of the different measures, which in turn are the latent variables [3]. After that, the latent variables are linked by arrows with the dependent variable, and this is done through the second component, which is the path analysis, the importance of this component is highlighted by the ability to study the effects of several factors on a certain phenomenon indirectly through several explanatory factors [10].

Thus, the form SEM is divided into two parts as follows [4]:

1. **Structural model:** which describes the (causal) links between the latent variables, where drawing a map of these links is the main goal of the analysis, so it works to develop the basic appearance of the SEM method.

2. **Measurement model:** which describes the links between the latent variables and the observed variables, or how to express the latent variables in light of the measured variables, taking into account the measurement error.

Thus, the SEM model is an integration between a measured model, which is a confirmatory factor analysis, and a constructive model, which is an analysis of a path between latent variables. Figure (1) below shows the structural equations model that contains latent variables, we note that the figure contains the supposed effect of A1 on A2 in addition to the links between the two non-measurable latent variables and their measurable apparent variables.
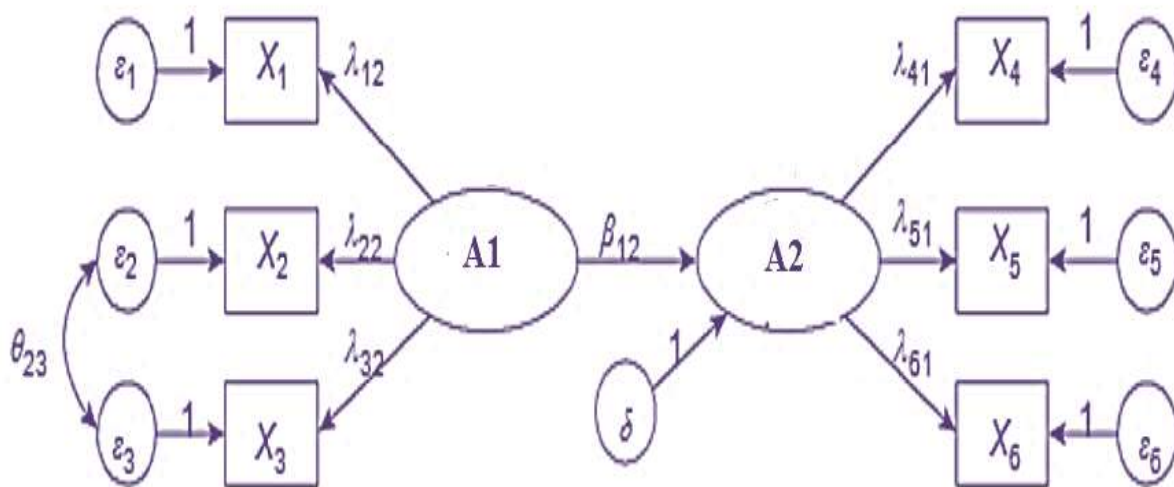
Figure(1) :The structural equations model contains latent variables

Thus, the graphic model in Figure (1) can be translated into the following equations:

$$A2 = \beta_{12}A1 + \delta$$

$$X_1 = \lambda_{12}A1 + \varepsilon_1$$
$$X_2 = \lambda_{22}A1 + \varepsilon_2$$
$$X_3 = \lambda_{32}A1 + \varepsilon_3$$
$$X_4 = \lambda_{41}A2 + \varepsilon_4 \qquad (2)$$
$$X_5 = \lambda_{51}A2 + \varepsilon_5$$
$$X_6 = \lambda_{61}A2 + \varepsilon_6$$

Where λ saturates the factor, A1 the independent latent variable, A2 the dependent latent variable, X the measured variables, δ and ε the measurement errors on the measured variables (residuals).

There are a set of parameters of the structural model and they have several types that can be summarized as follows [2,p.34]:

1. **Free parameters:** are the unknown features that are intended to be estimated, such as the variances of the residuals, factor saturations, covariances, or correlations between the independent latent variables and path coefficients, where such factors are left free during the estimation process of the model until we obtain the maximum harmonic function of the data matrix.

2. **Fixed parameters:** They are parameters that are fixed at a certain value, which is often zero or one integer, and they are fixed and do not change during the process of estimating and matching the model.

3. **Restricted features:** It is assumed that one of the features is equal to another parameter, but its value has not been determined, i.e. a restriction is placed on one of the features, and this type is between free and fixed.

3. Application
3.1. Data Description
The data was collected by designing a questionnaire distributed to a sample of the residents of the city of Basra in Iraq. The form included questions about some factors related to the health of the environment surrounding the individual, according to the paragraphs below.

Environmental Awareness:
- EA1= The economic level of the family.
- EA2= The extent of your awareness and interest in environmental matters.
- EA3= Are you ready to change your habits to save the environment.
- EA4= Are you over consuming substances.
- EA5= Do you advise people who pollute the environment.
- EA6= Do you believe in the effects of environmental pollution on health.
- EA7= Do you help report environmental pollution.

Inhaled Air Quality:
- AQ1= Home address.
- AQ2= Are you exposed to chemicals and toxic materials.
- AQ3= Do you live or work near airports.
- AQ4= Do you live or work near oil or gas companies.
- AQ5= Do you use renewable energy sources.
- AQ6= Do you incinerate household waste.
- AQ7= Do you smoke cigarettes or are exposed to frequent inhalation.

Water Quality:
- WQ1= Type of water tanks used.
- WQ2= The level of drinking water.
- WQ3= Water used for domestic consumption.
- WQ4=The quality of the water network in the region.

Waste and Sanitation Management:
- WAS1= Are there waste removal services in your area.
- WAS2= Are you managing waste properly.
- WAS3= Are you working to reduce the amount of waste.
- WAS4= Is there a sewage network in your area.
- WAS5= Are there bad smells from the sewage network in your place of residence.

There were some directly measured variables, as follows:
- DEP= Have you been exposed to a disease due to environmental pollution.
- DBP= Do you suffer from high blood pressure.
- DD= Do you have diabetes.

3.2. Analysis of The Results

In this section, the data described in the previous section (3-1) were analyzed. For the purpose of measuring the reliability of the questionnaire, Cronbach's alpha scale was used, as shown in Table (1), where we note that the value of the scale is equal to (0.83), which is greater than (0.7), so the scale is considered good

Table (1): Reliability Statistics

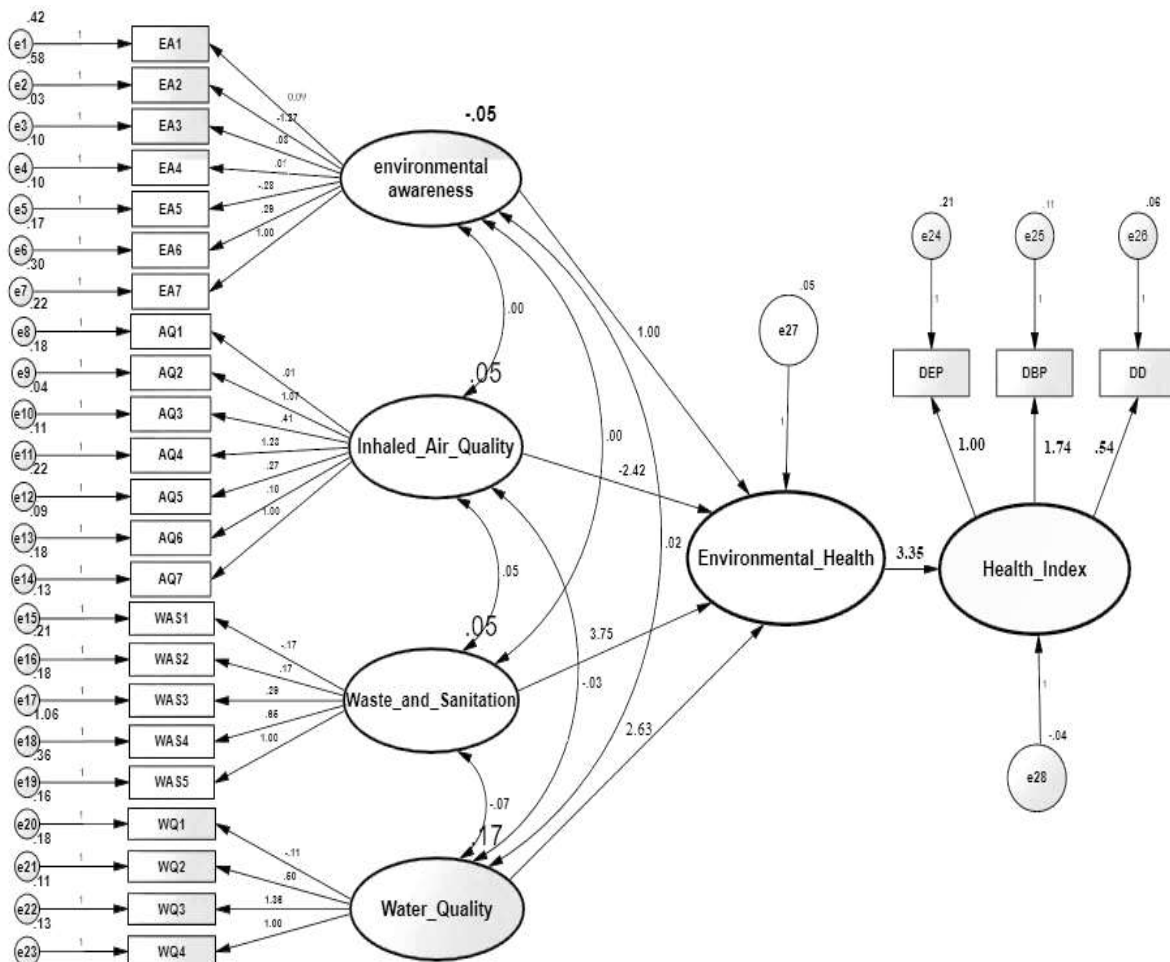| Cronbach's Alpha	N of Items
0.830	26

Figure (2):Structural equation modeling of the effect of environmental factors on the health index

The main objective of the analysis is to find out the extent of the relationship of the environmental factors surrounding the person and the extent of their impact on each of the environmental health index and the person's health index. The data was analyzed based on structural

equations modeling using the Amos statistical program, due to the inability to measure the variables directly, but rather they were in the form of latent variables. The chart was drawn to describe the relationships between the independent and dependent latent variables, as in Figure (2) below, where we note that each axis of the questionnaire is a latent variable, and each measured variable related to the axis was entered in addition to the measurement errors related to it, in addition to that the axes were linked With each other to see if there are relationships between them and then linking all the axes to find the latent variable that represents the environmental health index, then the environmental health index is entered by merging with some direct variables to estimate the main goal, which is the person's health index. We note from the results of estimating the parameters in Figure (2) that the waste and sanitation management axis had more environmental health impacts (3.75) then the water quality by (2.63), in addition, we note that the effect of environmental health on the individual's health index is a direct effect with a value of (3.35). For all the remaining variables, the relationships ranged between positive relationships for some and inverse relationships for others, as shown in the chart.

With regard to the measures through which the goodness of fit of the model is calculated, it was as shown in Table(2), where we note that the value of Chi-square/df equal to (1.187) which is less than (3) this indicates that the model is good, and it was valuable of CFI is (0.947) this value is also good, additionally we note that the value of the Root Mean Square Error Absolute (RMSEA) equal to (0.040 < 0.090) and the value of PCLOSE equal to (0.834 >0.05) we note that the values of these measures indicate that the goodness of fit of the model was good.

Table (2): Measures of goodness of fit for the structural equation model.

Scales	Default model
CMIN/DF	1.187
CFI	0.947
RMSEA	0.040
PCLOSE	0.834

4. Conclusions

The structural equation model was built to analyze the relationship between some factors affecting environmental health and their relationship to the human health index. The results indicated that the structural equations modeling method is suitable for analysis and gave good results in estimation. Through the results, we find that there are some environmental factors surrounding the individual that had a statistically significant effect that the waste and sanitation management axis had more environmental health impacts (3.75) then the water quality by (2.63), in addition, we note that the effect of environmental health on the individual's health index is a direct effect with a value of (3.35). In addition, we find from the results that the increase in the concentrations of harmful gases and radiations emitted into the air, especially in industrial areas, has a significant impact on the quality of inhaled air.

From the aforementioned, and due to the extreme importance of environmental health and its effects on human health, therefore, it is necessary to work hard and strive to preserve the environment, such as increasing the work of green spaces, which helps in purifying the air surrounding the population, and taking into account the causes leading to the deterioration of sanitation and other factors that were mentioned in studying. Must be trying to intensify efforts and conduct more research and statistical studies to solve the most important problems related to environmental health.

References

[1] M. Abbas, S. Dalia, K.Z. Edmundas, C. Fausto, N. Mehrbakhsh, J. Ahmad & Z. Habib, "Application of Structural Equation Modeling (SEM) to Solve Environmental Sustainability Problems: A Comprehensive Review and Meta-Analysis", Sustainability (2017), 9, 1814; doi:10.3390/su9101814.

[2] A.A. Abdel Nasser, Modeling the Structural Equation for Psychological and Social Sciences, Naif University Publishing, Saudi Arabia, AL-Riyadh, (2018).

[3] R.A. AL-Qatrawi, "Structural Equation Model among the Attitude towards Risk, the Big Five Factors of Personality, the Creative Thinking for Students in Palestinian Universities", PhD Dissertation, Suez Canal University, College of Education-Ismailia, (2015).

[4] N.G. Blentsh, Introduction in Structural Equation Modeling by Using IBM SPSS Statistics and AMOS, Research and Studies Center, Institute of Public Administration, Kingdom of Saudi Arabia, (2017).

[5] S.B. Emmanual, D.S. Joel, A.C. Brent, A.W. Gregory and S.V. Panter, "Structural equation modeling of parasympathetic and sympathetic response to traffic air pollution in a repeated measures study", Baja et al. Environmental Health (2013), 12:81, http://www.ehjournal.net/content/12/1/81.

[6] Y. Ferra, "The Estimation Process in Bayesian Structural Equation Modeling Approach", Journal of Physics: Conference Series 495 (2014) 012047, doi:10.1088/1742-6596/495/1/012047.

[7] D. Jerico and Q. Venusmar, "Relationships of Environmental Pollution, Economic Cost, and Public Health: A Structural Equation Modeling Study", IEOM Society International, p.923-931,(2021).

[8] A.O. Jose, M.R. Jose and I. Contreras, "An integrated model of structural equations with cognitive and environmental factors for the study of active commuting", Journal of Transport & Health, Vol. (24), (2022) 101319.

[9] F.H. Joseph, G.M. Tomas, M.R. Christain and S. Marko, "Structural Equations Modeling by Partial Least Squares", Jordan Academic Book Center, 1st, (2020).

[10] K.M. Subrat, D. Seemarekha, K.M. Subata and D. Abhiram, "Path analysis and its application in agricultural research", International Journal of Agriculture and Plant Science, Vol. 2; Issue 2; (2020); P. 01-03.

[11] S. Xin and L. Sik-Yum, Basic and Advanced Bayesian Structural Equation Modeling, John Wiley & Sons, Ltd, 1st edition, (2012).

[12] F. Yi, C. Jiquan, S. Gabrietla, J. Ranjeet, R.W.Susie, P. Hogeun and S. Changliang, "Applications of structural equation modeling (SEM) in ecological studies: an updated review", Fan et al. Ecological Processes (2016) 5:19, DOI 10.1186/s13717-016-0063-3.

AL- Rafidain
University College

JRUCS

Journal of AL-Rafidain
University College for
Sciences

النمذجة بالمعادلات البنائية في دراسة تطبيقية لتحليل اثر الصحة البيئية على مؤشر صحة الفرد في محافظة البصرة

د. لمياء عبدالجبار داود	د. أسماء أيوب يعقوب
lamya.abd@mtu.edu.iq	asmaa.yaqoob@uobasrah.edu.iq
كلية التقنيات الصحية والطبية، الجامعة التقنية الوسطى، بغداد، العراق	قسم الإحصاء ـ كلية الإدارة والاقتصاد ـ جامعة البصرة، البصرة، العراق

المستخلص

يعد التلوث والاخلال بالنظام البيئي من المشكلات الخطيرة على حياة الانسان، خصوصا ان معدل التلوث والتدهور البيئي في تزايد مستمر لذلك لابد من دراسة التفاعل بين البيئية والصحة من خلال تحليل أثر بعض العوامل البيئية (تغير المناخ، جودة الهواء، الوعي البيئي، مياه الشرب السامة وانظمة الصرف الصحي، الضوضاء....الخ) وكيفية تأثيرها على صحة الفرد. تم في هذا البحث استعمال نماذج المعادلات البنائية (SEM) في بناء نماذج سببية تعمل على دراسة العلاقات بين المتغيرات في وقت واحد ولاسيما في ظل وجود متغيرات كامنة ، حيث تم تقييم كل من العوامل البيئية الداخلة في الدراسة وتقدير مساهمتها في حساب مؤشر صحة الفرد. وقد تم جمع المعلومات من خلال عمل استمارة استبيان تم توزيعها على فئات مختلفة من سكان مدينة البصرة في العراق. من خلال النتائج تبين أن صحة البيئة لها تأثير إيجابي كبير على مؤشر الصحة، لذلك من الضروري الاهتمام بصحة البيئة وتقليل مصادر التلوث البيئي.