Research Paper

Volume : 3 | Issue : 9 | September 2014 • ISSN No 2277 - 8179

Latent Profile Model: A Case Study on Water Quality

Statistics

KEYWORDS : Latent Profile, Classification, Water Quality, Iraq

Dr. Nazera Khalil Dakhil	Assistant Professor, Department of Mathematics, Faculty of Computer Science and Mathematics, University of Kufa, Iraq.
Amenah Abdullateef Hameed	Department of Mathematics, Faculty of Computer Science and Mathematics, University of Kufa, Iraq.

ABSTRACT

Water is a natural resource vital for the survival of humanity and all species on earth. However, the quality and quantity of water in Iraq is increasingly being threatened due to high demand by humans and other living organisms. It is therefore the aim of this research to attempt to study the properties of the water in Al-Najaf province in Iraq, and to evaluate the Latent Profile (LP) by means of classifying water quality. It emerged that the quality of the water in our samples varied from average to poor. Thus the water in Al-Najaf may be suitable for general domestic use but not for drinking.

1. Introduction [10], [11]

The Latent Profile model is a latent variable model with a categorical latent variable and continuous manifest indicators. It was introduced in 1968 by Lazarsfeld and Henry [14]. Many researchers have studied the Latent Profile and used it in various practical applications. Goodman (1974) [7] considered wide structure models. These models served as possible explanations of the observed relationships among a set of manifest polytomous variables. Some illustrative applications of the models to data were also considered.

Previous research confirmed the need to take serious action in monitoring the Tigris river in Baghdad for proper management [3]. Other authors confirmed that the Tigris River did not meet the WHO drinking water standard [9]. However, other research concluded that samples of water from residential districts in Baghdad may be considered of good quality [5]. Further studies which evaluated the Euphrates river at Ramadi city and Al-Dhiban indicate that all the water quality parameters were within the WHO and Iraqi drinking water standard except the total hardness [13]. Additional research showed that the water quality of the Tigris river was better than in the Euphrates river [2]. The present work uses LP analysis which is applied to Al-Najaf's water using five water quality parameters.

This purpose of this research was to classify the water in Al-Najaf. Latent Gold 5.0 [17] and MATLAB was used to analyze the data.

2. Materials and methods

2.1 Latent Profile Model [12]

The LP analysis assumes that the population consists of C unobserved subgroups that can be referred to as the latent profiles, latent classes, or mixture components. Because the indicators are continuous variables, it is most natural to assume that their conditional distribution is normal.

Let w denote a vector of r continuous manifest variables. Let µj be the meaning of the manifest variable i in class j, be the variance of the manifest variable i assumed constant across classes, (i = 1, ... , r; j = 0, ... , K — 1) and be the prior probability that a randomly chosen individual is in class j with the constraint that

$$\sum_{j=0}^{k-1} \eta_j = 1$$

The marginal distribution of the manifest continuous variables is:

$$f(W) = \sum_{j=0}^{K-1} \eta_j \prod_{i=1}^{r} g(w_i|j) \qquad (1)$$

The conditional distribution was taken in Bartholomew (1987) [4] to be the normal with mean ; and unit variance:

$$g(w_i|j) = \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left(\frac{-(w_i-\mu_i)^2}{\sigma_w^2}\right)$$

The log-likelihood for a random sample of size n is:

$$L = \sum_{h=1}^{n} \log \sum_{j=0}^{K-1} \eta_j \prod_{i=1}^{r} g(w_{ih}|j) \qquad (2)$$

Where , $g(w_{ih}|j) \equiv g(w_{ij}|\mu_{ij})$

The maximum likelihood (ML) estimates are:

$$\hat{\eta}_j = \sum_{h=1}^{n} \frac{h(j|w_h)}{n} \qquad (3)$$

where $(j = 0,1,\dots,K-1)$ and

$$\hat{\mu}_{ij} = \sum_{h=1}^{n} w_{ih}\, h(j|w_h)/(n\hat{\eta}_j) \qquad (4)$$

where $(i = 1,\dots,s;\ j = 0,1,\dots,K-1)$

and $h(j \mid w_h)$, is the posterior probability that an individual with response pattern wh will be allocated to class j given by:

$$h(j|w_h) = \eta_j\, g(w_h|j)/f(w_h) \qquad (5)$$

Equations (3) and (4) can be incorporated into an E-M algorithm to derive the ML estimates.

2.2 Water Quality

High demand for water can lead to deterioration in water quality and quantity that impacts not only the marine environment but also the availability of safe water for human consumption [16]. The assessment of water quality is very important for

Volume : 3 | Issue : 9 | September 2014 • ISSN No 2277 - 8179

knowing its suitability for various purposes. Water quality index indicates the quality of water in terms of index number for any intended use. These are usually based on scientifically assessed acceptable levels of toxicity to either humans or aquatic organisms [14], [6], and [18]. Assessing the quality and quantity of water is therefore of crucial importance to identify risks to the environment, society and human health.

2.3 Water Parameters

Water quality is determined by properties of water. These properties are characterized differently throughout the world. This section lists the common water quality parameters important in drinking water used in this research.

(1) Turbidity: the overall water clarity and the main way to identify the quality of water. It can affect the color of the water and aquatic life [1].

(2)The Value of pH: a measure of acidity in the water. It is considered an important parameter for water usage and aquatic life. It ranges from 1 to 14, where a low pH indicates high acidity and a pH of 7 refers to an equilibrium concentration which represents pure or neutral water [1].

(3) Magnesium: magnesium (Mg) salts are more soluble than calcium and less Mg is found in surface water. The concentration of this chemical element in natural water is usually low [19].

(4) Chloride: the concentration of chloride anions (Cl-) determines the water quality as the quality of water worsens after an increase in the concentration of these anions. This limits the possibilities of using natural water for different purposes [15].

(5) Total Dissolved Solids (TDS): the term used to describe the inorganic salts and small amounts of organic matter present in solution in water. The presence of dissolved solids in water may affect its taste. Water with extremely low concentrations of TDS may be unacceptable because of its flat, insipid taste [8].

2.4 Description of the data

This data consists of 110 drinking water samples that were collected during 2014 from nine water stations and water projects located along Al-Najaf city in Iraq. The data consists of some physical and chemical properties of water.

5. Results

One principle for deciding the number of classes is to look in the BIC values, and choose the number of classes which minimizes BIC. Using this criteria, the best model is the 3-class model and the number of parameters is 32.

Table 1: shows water quality index for Al-Najaf city according to the Al-Najaf and WHO index.

Figure 1 : shows a conventional triangular diagrams (Tri-Plot) that are used to represent the LP classification in which the three classes represent proportions of whole water sample.

Table 1: Water Quality Index

Parameter	WHO	Al-Najaf
Turbidity	5	5
PH	6.5-8	6.5-8.5
Magnesium Chloride T.D.S.	10-50 250 < 500	100 350 1000

Table 2: Profile of Water Classes

Variable Mean	Class I	Class II	Class III
Turbidity	11.1**	3.2	4.6
PH	7.4	7.4	7.5
Magnesium Chloride T.D.S.	49.8 144.1 904.9	47.5 163.5 93.9	60.7* 241.20 1438.9**

Table 3: Percent of Water Distribution

Area	Class1	Class II	Class III
Al-Hadaria	0	58.3	41.7
Al-Qadisyah	17.6	64.7	17.6
Al-Kufa	35.3	47.1	17.6
Al-Mashkhab	52.2	13.0	34.8
Al-Manthira	100	0	0
Al-Najaf	100	0	0
Al-Hirah	29.4	41.2	29.4
Al-Jeroah	0	66.7	33.3
Al-Horia	0	100	0

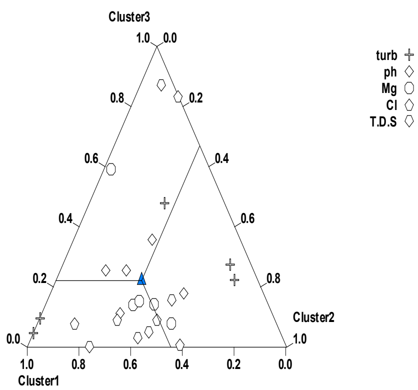

Figure 1: Distribution of Classes

Table 2: demonstrate the profile of water quality and quantity in Al-Najaf city for each classes. The most significant class is the first class which consists of around 45% of the water sample, with all variables within both limits except the turbidity which exceeds the WHO and the Al-Najaf indexes. The TDS is within the limit of Al-Najaf index but exceeds the WHO index. The next most important class is the second class, which holds around 33% of the water, with all variables within both limits except the TDS which is within the Al-Najaf index but exceeds the WHO index. The third class has around 22% of the water, with magnesium within the Al-Najaf limit but over the WHO index, and TDS exceeding both limits. All other variables are within both recommended limits.

and Al-Najaf (100%), while the second class is dominated by water from Al-Horia (100%) and the third class by water from Al-Hadaria (42%).

6. Discussion

It is important to note that turbidity offers a validation test of the entire water quality, and the LP analysis verifies that the turbidity range in the first class is more than double the recommended value by the Al-Najaf index. This may be alarming - since it all depends on the causes of soaring water turbidity in this class - because high turbidity levels are often associated with disease. On the other hand the pH values 7.4, 7.4 and 7.5 of all

classes respectively are within the WHO and Iraqi limits. Previous research suggested that any change in pH value from 7 may indicate the presence of pollutants in the water. [20]. Other research realizes chlorine becomes more toxic as the pH level of the water drops.[15].

7. Conclusion

In conclusion the quality of water of our samples varied from average to poor. Hence water in Al-Najaf may be suitable for general domestic use but not suitable for drinking.

REFERENCE

1] Al Bomola, A. H. (2011). "Temporal and Spatial Changes in Water Quality of the Euphrates River – Iraq". Division of Water | Resources Engineering Department of Building and Environmental Technology Lund University. | | | [2] Al-Janabi, Z. Z., Al-Kubaisi, A., and Al-Obaidy, A. M. J. (2012). "Assessment of Water Quality of Tigris River by using Water Quality Index (CCME WQI)". Journal of Al-Nahrain University Vol.15 (1), March, pp.119-126. | [3] Al- Shujairi, S. O. H. (2013). "Develop and Apply Water Quality Index to Evaluate Water Quality of Tigris and Euphrates Rivers in Iraq". International Journal of Modern Engineering Research (IJMER), Vol. 3, Issue. 4, Jul - Aug. 2013 pp-2119-2126. | [4] Bartholomew, D. J. (1987). "Latent Variable Models and Factor Analysis". London: Charles Griffin Si Co. Ltd. | [5] Barbooti, M. M., Bolzoni, G., Mirza, I. A., Pelosi, M., Barilli, L., Kadhum, R., and Peterlongo, G. (2010)."Evaluation of Quality of Drinking Water from Baghdad, Iraq". Science World Journal, Vol 5 (No 2.). | [6] Dodds, W. K., Oakes, R. M. (2004). "A Technique For Establishing Reference Nutrient Concentrations Across Watersheds Affected by Humans". Limnology and Oceanography Methods, 2: 333-341. | [7] Goodman, L. A. (1974). "Exploratory Latent Structure Analysis Using Both Identifiable and Unidentifiable Models". Biometrika, 61: 215-231. | [8] Holt, J. and G. Macready (1989). "A simulation study of the difference chi-square statistic for comparing latent class models under violation of regularity condi-tions". Applied Psychological Measurement 13, 221-231. | [9] Kadhem, A. J. (2013). "Assessment of Water Quality in Tigris River-Iraq by Using GIS Mapping, Natural Resources", Vol.4 No.6, 441-448. | [10] Lazarsfeld, P. F., Henry, N. W. (1968). "Latent Structure Analysis", Boston: Houghton Mifflin. | [11] McLachlan, G. J., Basford, K. E. (1988). "Mixture Models: Inference and Application to Clustering". New York: Marcel Dekker. | [12] Moustaki, I. (1996). "Latent Variable Models For Mixed Manifest Variables". Submitted in Fulfilment of the Requirement for the Degree of Doctor of Philosophy, London School of Economics and Political Science University of London. | [13] Mustafa, A. S. (2009). " Evaluation ofWater Quality Parameters in the Euphrates River within Ramadi City and Al-Dhiban Canal". J. of university of anbar for pure science : Vol.3: No.2. | [14] Robertson, D. M., Saad, D. A., and Heisey, D. M. (2006). "A regional Classification Scheme for Estimating Reference Water Quality in Streams Using Land Use Adjusted Spatial Regression-Tree Analysis". Environmental Management. 37 (2): 209-229. | [15] The Red Book. (1976). "Quality Criteria for Water". 534. | [16] United Nations Environment Programme Global Environment Monitoring System/ Water Programme (2006). "Water Quality for Eco-system and Human Health, Ontario, CANADA. | [17] Vermunt, J. K. and Magidson, J. (2005). " Latent GOLD 4.0 User's Guide ". Statistical Innovations Inc. | [18] Wickham, J. D., Riitters, K. H., Wade, T. G., Jones, K. B. (2005). "Evaluating the Relative Roles of Ecological Regions and Land-Cover Composition for Guiding Establishment of Nutrient Criteria". Landscape Ecology 20(7): 791-798. | [19] World Health Organization. (2009)."Calcium and Magnesium in Drinking-water Public health significance". WHO Press. | [20] WHO (1996). (2003) "Guidelines for drinking-water quality". 2d ed. Vol.2.UN World Water Development Report 1: Water for People, Water for life. | |