

Modelling the regional vulnerability to Echinococcosis based on environmental factors using fuzzy inference system: A case study of Lorestan Province, west of IranMojtaba Ahmadijad¹, Zeinab Obeidavi², Zia Obeidavi³, Reza Alipoor⁴

¹ Associated professor, Department of General Surgery, Faculty of Medicine, Alborz University of Medical Sciences, Karaj, Iran

² M.Sc. of Remote Sensing and Geographic Information System, Faculty of Earth Sciences, Shahid Chamran University of Ahvaz, Iran

³ Medical Student, Student Research Committee, Faculty of Medicine, Lorestan University of Medical Sciences, Khorramabad, Iran

⁴ Student Research Committee, Fasa University of Medical Sciences, Fasa, Iran

Type of article: Original

Abstract

Background and aim: Echinococcosis as a zoonosis disease is one of the most important parasitic helminth that is affected by many risk factors such as the environmental factors. Thus, we predicted the regional vulnerability to Echinococcosis based on environmental factors using a fuzzy inference system (FIS) in Lorestan Province.

Methods: Our study was cross-sectional study on 200 patients from Lorestan Province (west of Iran) who underwent surgery for hydatidosis between October 2005 and November 2014. In order to model the vulnerability to Echinococcosis, first we determined the effective environmental variables. In the next step, the FIS was designed and implemented using MATLAB v.2012 software. Thus, definition and determination of linguistic variables, linguistic values, and their range were performed based on expert knowledge. Then, the membership functions of inputs (environmental variables) and output (vulnerability to Echinococcosis) were defined. A fuzzy rules base was formed. Also, the defuzzification of output was done using a centroid defuzzification function. To test the accuracy of the predictive model, we calculated the AUC (to this purpose, we used four different thresholds, 5%, 10%, 15%, and 20%) using IDRISI Selva v.17.0 software.

Results: Based on the results of this study, Aligoudarz and Koohdasht counties were identified as a highest and lowest risk area in Lorestan, respectively. The results showed that a predictive model was more efficient than a random model (AUC>0.5). Also, potential vulnerable areas cover 78.29% at threshold of 5%, 60.72% at threshold of 10%, 43.54% at threshold of 15%, and 39.82% at threshold of 20% of the study area.

Conclusion: According to the success of this research, we emphasized the necessity of attention to fuzzy approach to model vulnerability to hydatidosis. This approach can provide a practical economic basis for making informed preventive services decisions and the allocation of health resources.

Keywords: Echinococcosis; regional vulnerability; environmental variables; fuzzy inference system

1. Introduction

Cystic Echinococcosis (CE) as a zoonotic disease infects over three million persons. It is also one of the most important helminthic diseases (1), which, according to high costs and cause of death in humans and animals, is considered an important health problem (2). The cause of this disease is the larval phase of the cestode *Echinococcus granulosus sensu lato* (3). In particular, the cystic Echinococcosis is common in rural regions, especially in sheep-raising areas where the entrails of livestock are given to dogs to eat (4). This disease is native in some countries of Latin America, Asia, and southern Europe. Iran is also the most important endemic foci of this disease; studies on Iranian patients show that 1–1.12 of every 100,000 people have hydatidosis (5). Serological

Corresponding author:

Zeinab Obeidavi, Faculty of Earth Sciences, Shahid Chamran University of Ahvaz, Iran.

Tel: +98.9166186507, Email: z.obeidavi@gmail.com

Received: April 15, 2017, Accepted: July 05, 2017, Published: December 2017

iThenticate screening: June 28, 2017, English editing: October 15, 2017, Quality control: October 22, 2017

This article has been reviewed / commented by three experts

© 2017 The Authors. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

prevalence of this disease in Iran has been reported as 3% in Zanjan Province, 9.2% in Ardabil Province, 13.7% in Fars Province, 13.8% in Khuzestan Province, and 15.4% in Lorestan Province (6). In fact, Iran's presence in endemic area and lack of sanitation and inadequate planning to control and fight against this disease have resulted in it becoming a huge problem in our country (7), so that, based on the recent classification of the world health organization (WHO), Iran is one of the hyper-endemic areas in terms of hydatidosis prevalence (8). The impact of geographic location and temporal variations in the spread of infection patterns, reproductive cysts, and protoscoleces viability (9-11), along with the role of aforementioned indices in the parasite cycle stability and the survival and spread of infection in humans, has made it difficult to obtain sufficient knowledge of the distribution and severity of disease in any part of the country (Iran). Undoubtedly, coping with this problem requires access to a comprehensive understanding of the disease and its spread conditions, so that, by a holistic approach, we can take steps in line with proper management. In this regard, the use of geographic information system (GIS) techniques in combination with computational intelligence methods, particularly a fuzzy inference system (FIS), would be useful because, by providing extensive areas for studying, while taking into account the different parameters simultaneously, results in reduced time and cost of the studies and an increase in accuracy. Also, mapping the spatial distribution of the disease at regional and local scales will significantly help to analyze the relationship between the environment and the potential of pathogens spatial distribution to model the risk of disease emergence. Thus, this study, by emphasizing the spatial and ecological aspects of the disease, will clarify the geographical-epidemiological communications of the disease. The result of this research provides a good map for health centers and environment department of the Lorestan Province; on the basis, we can have many methods to harness and manage the spread of this disease. Using FIS in disease mapping was previously considered by some researchers. For example, Aroyehun et al. used fuzzy logic to predict the risk of water-borne diseases (malaria) (12). Also, Pandey et al. in their study using FIS investigated the air quality with the aim of providing risk mapping of respiratory diseases (13). In addition, Akhavan et al. in their study, which focused on risk mapping of cutaneous leishmaniasis, used the FIS in combination with neural networks (14). In this study, we predict the Echinococcosis potential distribution or the regional vulnerability to Echinococcosis based on environmental factors using FIS in Lorestan Province. In fact, the Echinococcus species transmission dynamics is associated with many risk factors, and the relationship between these risk factors and environmental parameters (such as climate, landscape characteristics or geographical location) is being increasingly recognized (15). Environmental variables affect the intermediate host density and egg survival, which are controlled by temperature, moisture, and precipitation (16). Also, in previous studies (17, 18) a significant relation amid agricultural lands and humans hydatidosis prevalence was found. Thus, we designed and implemented a FIS to predict regional vulnerability to Echinococcosis in Lorestan Province based on temperature, precipitation, and agricultural lands.

2. Material and Methods

2.1. Study area

Lorestan Province is located in the southwest of Iran and covers an area of approximately 28294 km² (Figure 1). This Province has a variety of climates (e.g., semi-arid, Mediterranean, semi-humid, humid, very humid), with average annual temperature 17.2 °C and mean annual precipitation 550 mm. The current population of Lorestan is 1,754,243; the urban population constitutes 61.34%, and rural population constitutes 38.62% of the total population (19).

2.2. Collection of Hydatidosis patients' records

Our study was a cross-sectional study on 200 cases, which were referred to the Shohada Ashayer and Shahid Rahimi hospitals of Khorramabad (as a referral center) with Echinococcosis diagnosis and underwent surgery for hydatidosis from October 2005 to November 2014. People from across the province are often referred to these hospitals for surgery. In addition to demographic data, patients' addresses were collected from medical records of CE patients. It is notable that, to prepare the spatial layer (shapefile) of occurrence of hydatidosis, we registered 200 points based on patients' addresses using GPS in study area. This spatial layer, prepared using ArcGIS software, will be used for validating the predictive model.

2.3. Ethical consideration

The protocol of present study was accepted by the Ethical Review Board of Lorestan University of Medical Sciences. Informed satisfaction forms were signed by all study participants or their parents/guardians.

2.4. Environmental variables

We selected temperature, precipitation, and agricultural lands as predictors for the regional vulnerability of Echinococcosis (Figure 2). These predictors were chosen based on literature review and collected expert knowledge (4, 16, 20).

2. 5. Fuzzy Inference System (FIS)

The FIS structure is described as follows (21, 22):

- A rule base to save fuzzy rules (conditional sentences [if-then]).
- A membership functions (MFs) interface to define the membership functions of the fuzzy sets.
- A fuzzy inference unit for decision-making based on the rules.
- A fuzzification interface to convert the crisp inputs into fuzzy inputs.
- A defuzzification interface to convert the fuzzy output into a crisp output.

In this research, the proposed FIS was developed using a Mamdani inference model, which consists of three inputs (temperature, precipitation, and agricultural lands variables) and one output (vulnerability index). The MFs of temperature and precipitation were designed using three linguistic variables, namely, low, medium, and high. The three Gaussian MFs of temperature and precipitation variables are shown in Figure 3. Also, the MFs of agricultural lands were designed using two linguistic variables, namely, “poor” and “good.” Two single-tone MFs of agricultural lands variable are shown in Figure 3. In addition, to design the MFs of output, five linguistic variables are used, namely, “very low”, “low”, “medium”, “high”, and “very high” vulnerability. The five Gaussian MFs of output are shown in Figure 4. In the following, we defined the fuzzy rules, which were introduced to the system as if-then conditional sentences (Table 1). Then, the consequence MFs were obtained based on the Mamdani model, and, using a centroid defuzzification function, the MFs quantitative values were calculated similar to the quantitative values allocated to the vulnerability function. We note that, the FIS was designed and implemented using MATLAB v.2012 software.

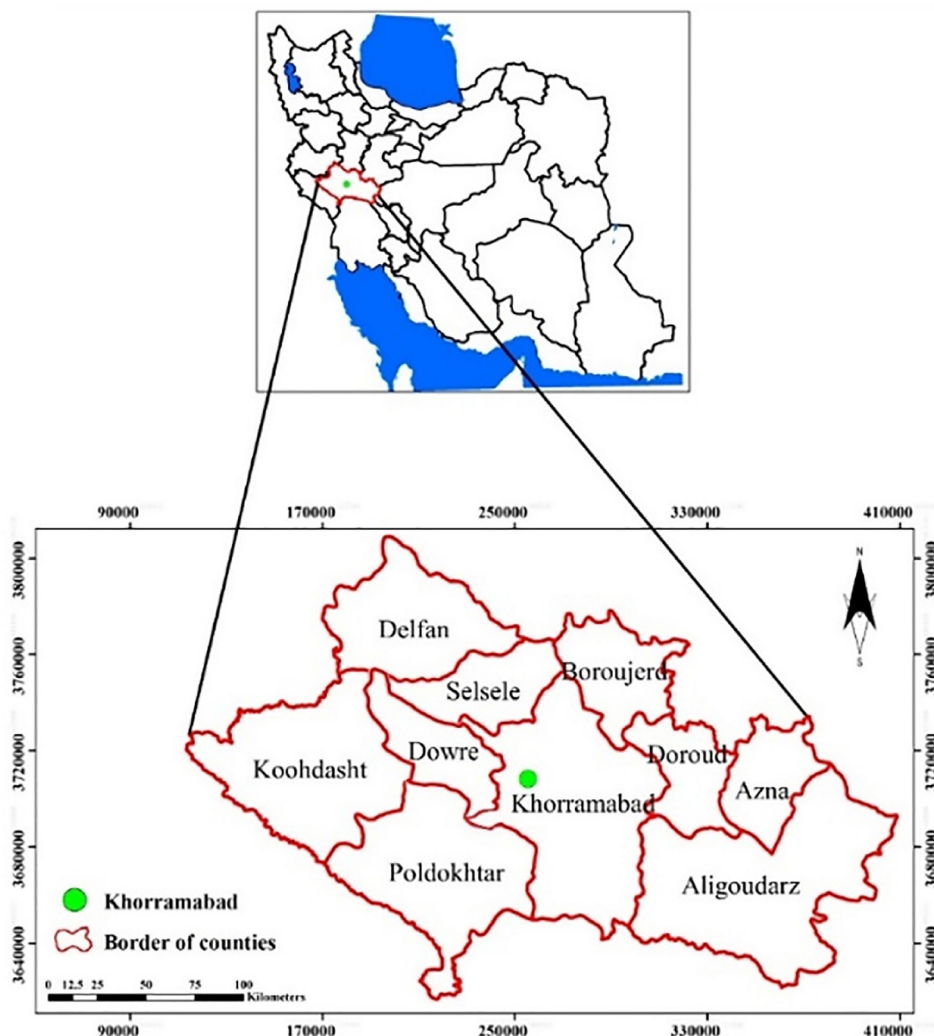


Figure 1. Study area: location of Lorestan Province in Iran with 10 counties of Lorestan.

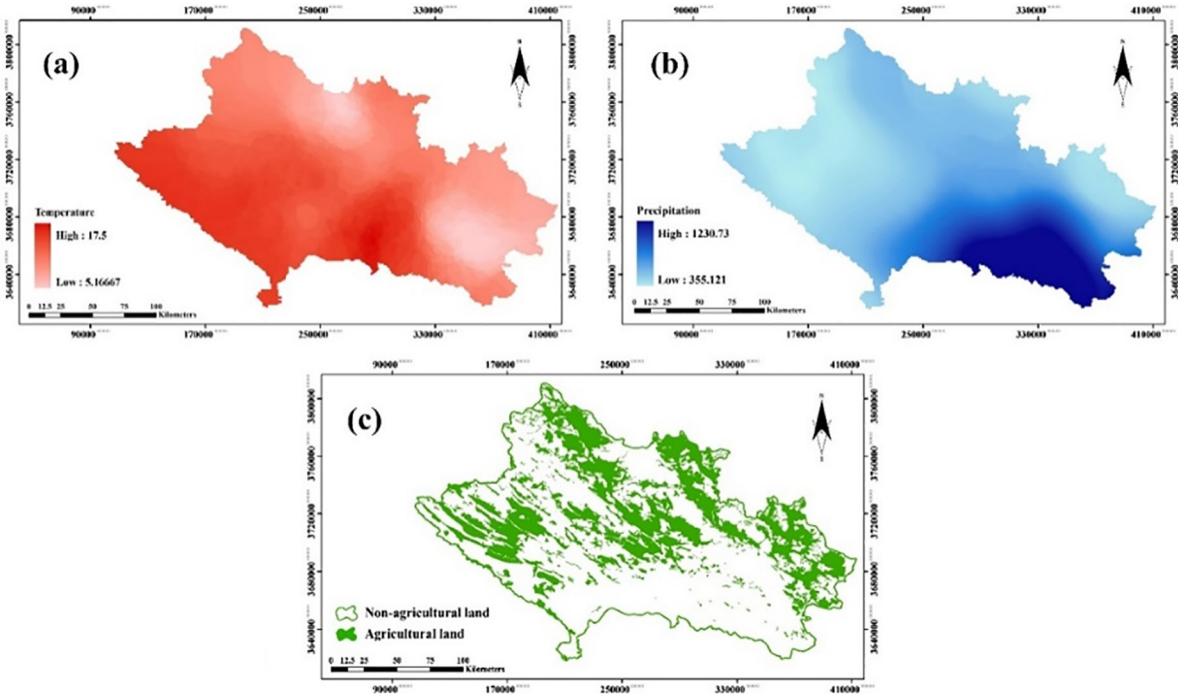


Figure 2. Environmental variables: (a) Temperature distribution map. (b) Precipitation distribution map. (c) Agricultural lands map.

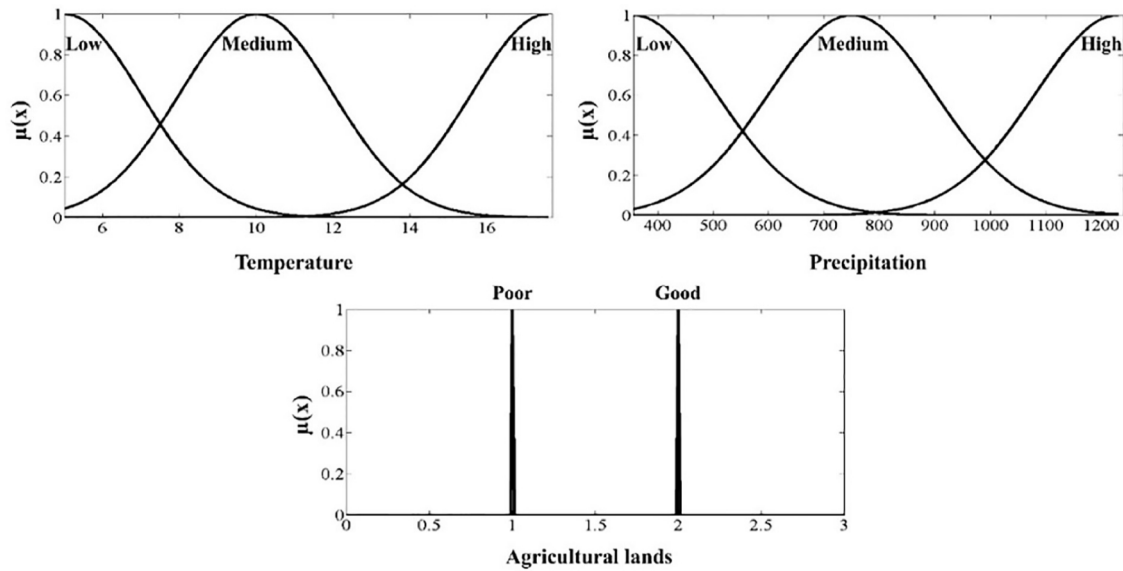


Figure 3. The membership functions of the input variables.

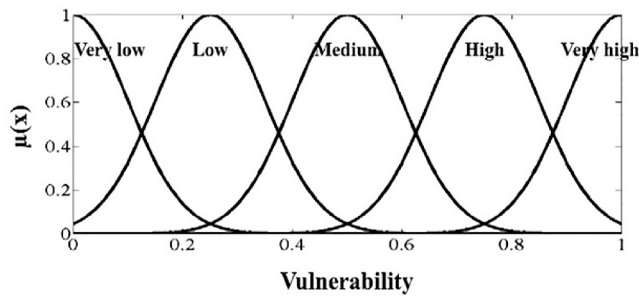


Figure 4. The membership functions of the output variable.

Table 1. Three examples of fuzzy rules to predict vulnerability.

Number	Rules
1	If temperature is “high” and precipitation is “low” and agricultural lands is “poor” then vulnerability is “very low”
2	If temperature is “low” and precipitation is “low” and agricultural lands is “poor” then vulnerability is “low”
3	If temperature is “low” and precipitation is “high” and agricultural lands is “good” then vulnerability is “very high”

2. 6. Validation of the predictive model

In present study, the area under the receiver operating characteristic (ROC) Curve or AUC was used to test the significance of the predictive model (AUC was tested using IDRISI Selva v.17.0 software). This approach requires application of a threshold to define “vulnerable” and “invulnerable” areas. So, we used four different thresholds of 5%, 10%, 15%, and 20% (23-25). Note that, for example a threshold of 5 would be expected to yield approximately 5% omission in sample of occurrence localities of disease.

3. Results

In Figure 5, the result of FIS is shown as a map to represent the regional vulnerability to Echinococcosis in Lorestan Province. Based on this result, Aligoudarz and Koohdasht counties were identified as highest and lowest risk areas in Lorestan, respectively. According to the results of applying the thresholds (Figure 6), potential vulnerable areas cover 78.29% at a threshold of 5%, 60.72% at a threshold of 10%, 43.54% at a threshold of 15%, and 39.82% at a threshold of 20% of the study area (Table 2). Also, as Table 2 shows, the threshold of 5% and threshold of 20% provide the lowest and highest AUC values, 0.752 and 0.838, respectively. It is notable that an AUC value of 1 demonstrates perfect prediction, and an AUC value of 0.5 shows that the model performance is not preferable to random. Also, AUC values between 0.7 to 0.8 show good model and values of 0.8 to 0.9, which indicate an excellent model (26); in present research the predictive model in lowest threshold (threshold of 5%) is a good model and in the highest threshold (threshold of 20%) is an excellent model. Thus, the predictive model in the present study has an acceptable accuracy.

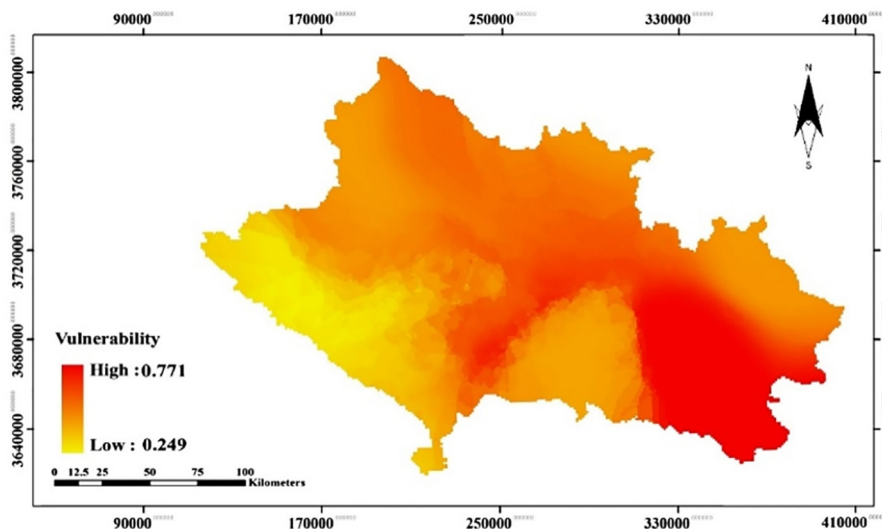


Figure 5. Regional vulnerability to Echinococcosis in Lorestan Province.

Table 2. The AUC values of different thresholds.

Thresholds	AUC	Potential vulnerable areas (%)	Potential invulnerable areas (%)
5%	0.752	78.29	21.71
10%	0.786	60.72	39.28
15%	0.815	43.54	56.46
20%	0.838	39.82	60.18

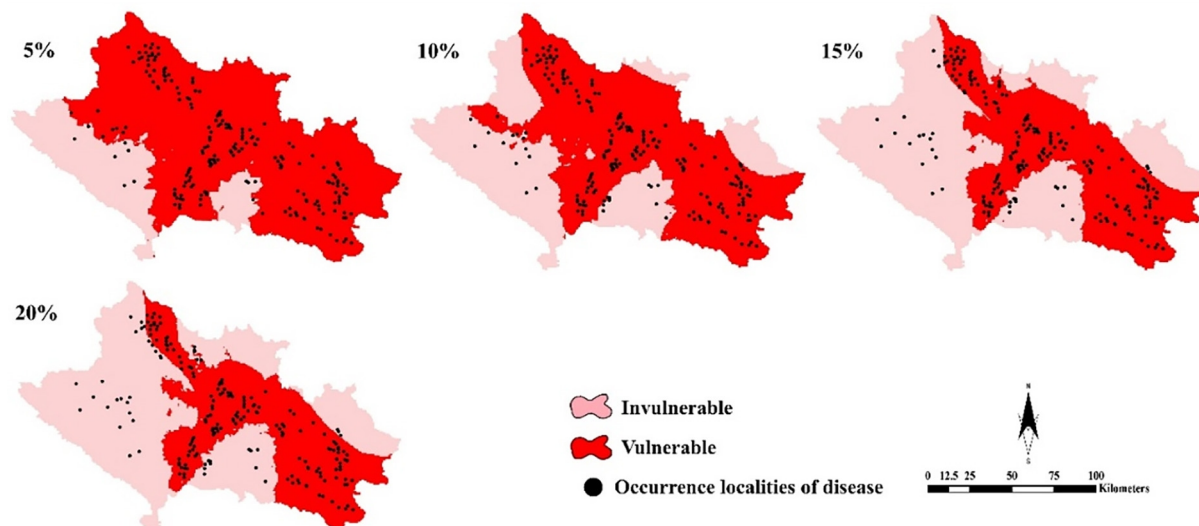


Figure 6. Vulnerable and invulnerable regions to Echinococcosis in Lorestan Province using FIS and applying four thresholds.

4. Discussion

Our findings indicate success in using FIS in modelling the risk of hydatidosis in the study area. The success of this method in modelling the risk of different diseases was emphasized by other researchers, including Akhavan et al. (14), Pandey et al. (13), and Aroyehun et al. (12). Furthermore, the results of model validation while acknowledging the potential of FIS in modelling the risk of hydatidosis confirm the significant influence of environmental variables. As mentioned above, Aligoudarz and Koohdasht counties were identified as the highest and lowest risk areas, respectively. About this result, we can emphasize mainly the effect of climatic variables as temperature and precipitation because of low temperature and high precipitation in Aligoudarz County and relatively high temperature and low precipitation in Koohdasht County. This has already been emphasized by various researchers. Yang et al. in their study emphasized that the environmental variables affect the intermediate host density and egg survival; these researchers also highlighted the temperature, moisture, and precipitation as the controller factors (16). Also, Atkinson et al. in their study found a significant relation amid agricultural lands and human hydatidosis prevalence (17); this finding was also emphasized by Otero-Abad and Torgerson (18). Also, according to results of present study, we demonstrated that setting the decision threshold changes the performance of the predictive model and the proportion of the study area that predicted as vulnerable area. In fact, a decision threshold should be chosen based on the proposed application of the model. Thus, for applications that aim to guide health specialists toward identification of unknown vulnerable areas using a low threshold (in this study: 5% threshold) is thus appropriate. In contrast, for applications that aim to manage the limited health resources, using a high threshold (in this study: 20% threshold) is thus appropriate. Finally, we note the possibility of over-covering vulnerable areas and misdirect preventive programs by setting thresholds relatively low or uncovering potentially important vulnerable areas by setting thresholds relatively high.

5. Conclusions

According to this fact that the hydatidosis is endemic in Lorestan and many other provinces of Iran and based on the outcomes of present study, we emphasized that the capabilities of GIS and fuzzy logic can facilitate the identification of regions with highest risk of disease spreading in endemic regions. Also this approach can provide a practical economic basis for taking informed preventive services decisions and the allocation of health resources. This interaction between medicine and engineering sciences promise a bright future to prevent wastage of financial resources. Finally, we note that the lack of access to some of effective variables such as distribution of livestock and domestic dogs in the province lead to model the vulnerability to Echinococcosis without these variables. Thus, we suggest using these variables in future studies to provide a more efficient model for better control of this disease.

Acknowledgments:

This study received financial support from the Vice Chancellor for Research in Lorestan University of Medical Sciences (Project No. 1843).

Conflict of Interest:

There is no conflict of interest to be declared.

Authors' contributions:

All authors contributed to this project and article equally. All authors read and approved the final manuscript.

References:

- 1) Zhang W, Ross AG, McManus DP. Mechanisms of immunity in hydatid disease: implications for vaccine development. *J Immunol.* 2008; 181: 6679–85. doi: 10.4049/jimmunol.181.10.6679. PMID: 18981082.
- 2) Zamani A, Kalikias S. Hydatid cyst of the parotid gland: a case report. *Iran J Pediatr.* 2006; 16: 95-8.
- 3) McManus DP, Zhang W, Li J, Bartley PB. Echinococcosis. *Lancet.* 2003; 362(9392): 1295–304. doi: 10.1016/S0140-6736(03)14573-4.
- 4) The Center for Food Security and Public Health (CFSPH). Echinococcosis.2011. Available from: <http://www.cfsph.iastate.edu/Factsheets/pdfs/echinococcosis>.
- 5) Iranian Doctors website. Echinococcus granulosus. 2008. Available from: <http://www.pezeshkan.org>.
- 6) Zibaei M, Azargoon A, Ataie-Khorasgani M, Ghanadi K, Sadjjadi SM. The serological study of cystic echinococcosis and assessment of surgical cases during 5 years (2007-2011) in Khorram Abad, Iran. *Niger J Clin Pract.* 2013; 16: 221-5. doi: 10.4103/1119-3077.110156. PMID: 23563466.
- 7) Tavakoli HR, Bayat M, Kousha A. Hydatidosis infection study in human and livestock populations during 2002-2007. *American-Eurasian J Agric Environ Sci.* 2008; 4(4): 473-7.
- 8) Eftekhari F. Clinical and demographic features of patients with hydatid cyst admitted in Kerman university hospitals between 1991-2000. *JKMU.* 2005; 12(4): 252-7.
- 9) Arambulo P. Public health importance of cystic echinococcosis in Latin America. *Acta Trop.* 1997; 67: 113-24.
- 10) Dueger EL, Gilman RH. Prevalence, intensity and fertility of ovine cystic echinococcosis in the central Peruvian Andes. *Trans R Soc Trop Med Hyg.* 2001; 95: 379-83. doi: 10.1016/S0035-9203(01)90188-9.
- 11) Scala A, Garippa G, Varcasia A, Tranquillo VM, Genchi C. Cystic echinococcosis in slaughtered sheep in Sardinia (Italy). *Vet Parasitol.* 2006; 135(1): 33-8. doi: 10.1016/j.vetpar.2005.08.006. PMID: 16226378.
- 12) Aroyehun AA, Sabejeje TA, Bayo-Lebi D, Olawuyi NJ, Ayinla NJ, OgunwaleYE. Fuzzy logic based predictive model for likelihood of water related disease. *Discovery.* 2017; 53(257): 321-33.
- 13) Pandey M, Singh V, Vaishya RC. Geomatics approach for assessment of respiratory disease mapping. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.* 2014; 8: 205-11. doi: 10.5194/isprsarchives-XL-8-205-2014.
- 14) Akhavan P, Karimi M, Pahlavani P. Risk Mapping of Cutaneous Leishmaniasis via a Fuzzy C Means-based Neuro-Fuzzy Inference System. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.* 2014; 2: 19-23. doi: 10.5194/isprsarchives-XL-2-W3-19-2014.
- 15) Danson FM, Giraudoux P, Craig PS. Spatial modelling and ecology of Echinococcus multilocularis transmission in China. *Parasitol Int.* 2006; 55(Suppl): S227-31. doi: 10.1016/j.parint.2005.11.047. PMID: 16371253.
- 16) Yang YR, Liu T, Bai X, Boufana B, Craig PS, Nakao M, et al. Natural infection of the ground squirrel (*Spermophilus* spp.) with Echinococcus granulosus in China. *PLoS Negl Trop Dis.* 2009; 3: e518. doi: 10.1371/journal.pntd.0000518. PMID: 19771151, PMCID: PMC2737643.
- 17) Atkinson JA, Gray DJ, Clements ACA, Barnes TS, Mcmanus DP, Yang YR. 2013. Environmental changes impacting Echinococcus transmission: research to support predictive surveillance and control. *Glob Change Biol.* 2013; 3: 677-88. doi: 10.1111/gcb.12088. PMID: 23504826.
- 18) Otero-Abad B, Torgerson PR. 2013. A systematic review of the epidemiology of echinococcosis in domestic and wild animals. *PLoS Negl Trop Dis.* 2013; 7(6): e2249. doi: 10.1371/journal.pntd.0002249. PMID: 23755310 PMCID: PMC3674998.
- 19) Craig PS, Giraudoux P, Shi D, Bartholomot B, Barnish G, Delattre P, et al. An epidemiological and ecological study of human alveolar echinococcosis transmission in south Gansu, China. *Acta Trop.* 2000; 77: 167-77. doi: 10.1016/S0001-706X(00)00134-0.
- 20) Statistical Center of Iran. 2014. Iran statistical yearbook (population). Available from: <https://www.amar.org.ir/english/Iran-Statistical-Yearbook/Statistical-Yearbook-2013-2014>.
- 21) Giraudoux P, Delattre P, Takahashi K, Raoul F, Quere JP, Craig P, et al. Transmission of ecology of Echinococcus multilocularis in wild-life: what can be learned from comparative studies and multi-scale approaches? *Conference of Cestode zoonoses: Echinococcosis and cystercercosis.* 2002: 251-66.

- 22) Mahfouf M, Abbod MF, Linkens DA. A survey of fuzzy logic monitoring and control utilization in medicine. *Artif Intell Med.* 2001; 21(1): 27-42. doi: 10.1016/S0933-3657(00)00072-5.
- 23) Shleeg AA, Ellabib IM. Comparison of Mamdani and Sugeno fuzzy inference systems for the breast cancer risk. *International Journal of Computer, Electrical, Automation, Control and Information Engineering.* 2013; 7(10): 1343-7.
- 24) Anderson RP, Gonzalez-Jr I. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecol Model.* 2011; 222: 2796-811. doi: 10.1016/j.ecolmodel.2011.04.011.
- 25) Escalante T, Rodriguez-Tapia G, Linaje M, Illoldi-Rangel P, Gonzalez-Lopez R. Identification of areas of endemism from species distribution models: Threshold selection and NEARCTIC mammals. *TIP Rev Esp Cienc Quím Biol.* 2013; 16(1): 5-17. doi: 10.1016/S1405-888X(13)72073-4.
- 26) Obeidavi Z, Rangzan K, Mirzaei R, Kabolizade M, Amini A. Wildlife Habitats Suitability Modelling using Fuzzy Inference System: A Case Study of Persian Leopard (*Panthera pardus saxicolor*) in Shimbar Protected Area. *IJAE.* 2017; 6(1): 57-66. doi: 10.18869/acadpub.ijae.6.1.57.
- 27) Giovanelli JGR, De Siqueira MF, Haddad CFB, Alexandrino J. Modeling a spatially restricted distribution in the Neotropics: how the size of calibration area affects the performance of five presence-only methods. *Ecol Model.* 2010; 221: 215–24. doi: 10.1016/j.ecolmodel.2009.10.009.