ACTA CLIMATOLOGICA ET CHOROLOGICA
Universitatis Szegediensis, Tomus 49-50, 2016, 41-48

MODELING ENVIRONMENTAL PROCESSES WITH BAYESIAN NETWORKS, BASED ON THE EXAMPLE OF EUTROPHICATION OF KARSTIC LAKES

M KISS[1], A SAMU[2], E TANÁCS[1] and I BÁRÁNY-KEVEI[1]

[1]*Department of Climatology and Landscape Ecology, University of Szeged, P.O.Box 653, 6701 Szeged, Hungary*
[2]*TETT Association, Holló str. 2., 6723 Szeged, Hungary*
E-mail: kiss.marton@geo.u-szeged.hu

Summary: Bayesian networks are more and more frequently used tools in environmental sciences. The advantage of the methodology is the possibility of handling conditional probabilities, thus they are able integrate expert knowledge and information from different knowledge domains. There are several examples for the use of the approach from the fields of water management, climate and land use change modeling. The evaluation of ecosystem services is an interdisciplinary research field, for which Bayesian networks have much potential. In our case study, we present a model examining the eutrophication process of karstic lakes, focusing on the share of toxic ammonia. It is important from the point of view of the health status of fish populations, which is an important indicator of ecosystem state and service providing capacity. The sensitivity analysis revealed that pH and water temperature have the most important role in the investigated process. The model is suitable for evaluating management alternatives and other types of decision support.

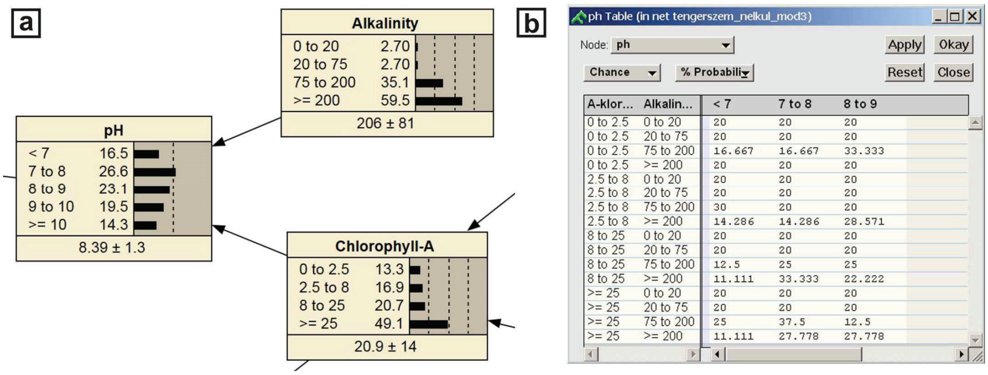
Key words: environmental modeling, Bayesian networks, eutrophication, fish populations, ammonia

1. INTRODUCTION

1.1. Bayesian networks

Bayesian networks are useful tools of environmental science and decision support applications. The use of artificial intelligence related graphic mathematical models spread at first for medical diagnostics (e.g. Kahn et al. 1997), but now they are also used in many other domains as well, like solving telecommunicational (Devitt et al. 2006) and financial problems (Gemela 2001). Bayesian networks describe relations between variables by conditional probabilities. In the graphic interpretation of the Bayesian networks, variables are marked by nodes, while causal relations are marked by edges. The relations among variables are defined in the related conditional probability tables (Fig. 1). (The whole probability table contains the combination of each relevant variable's every possible value.)

Conditional probability tables must contain discrete values, so continuous variables have to be partitioned in intervals. The networks may contain so-called utility variables as well. By assigning utility scores to different possible values of end variables in complex models, the setting of parent variable values can be optimized using the known probabilities. Using costs as benefit variables makes it possible to examine the cost efficiency of investments or environmental interventions as well. Besides direct value input, conditional probability values can be calculated by summarizing a sufficient amount of measured data as well. This is how medical diagnostics applications generally operate, just as the case study hereby presented. Besides conditional probability, there is a possibility to include equation-defined deterministic relations as well.



A-klor...	Alkalin...	< 7	7 to 8	8 to 9
0 to 2.5	0 to 20	20	20	20
0 to 2.5	20 to 75	20	20	20
0 to 2.5	75 to 200	16.667	16.667	33.333
0 to 2.5	>= 200	20	20	20
2.5 to 8	0 to 20	20	20	20
2.5 to 8	20 to 75	20	20	20
2.5 to 8	75 to 200	30	20	20
2.5 to 8	>= 200	14.286	14.286	28.571
8 to 25	0 to 20	20	20	20
8 to 25	20 to 75	20	20	20
8 to 25	75 to 200	12.5	25	25
8 to 25	>= 200	11.111	33.333	22.222
>= 25	0 to 20	20	20	20
>= 25	20 to 75	20	20	20
>= 25	75 to 200	25	37.5	12.5
>= 25	>= 200	11.111	27.778	27.778

Fig. 1 A variable and its parent variables in the graphic model (a), with the relevant conditional probability table (b)

1.2. Bayesian networks in environmental and ecosystem service modeling

There are several fields of landscape and environmental research, for which Bayesian networks may serve as a useful tool, in case of theoretical problems and practical applications as well. The possibility to include information from different domains was presented in the study of Carmona et al. (2011). It provides a decision support model in water management, which is one the most frequent domains where Bayesian networks are applied. Beyond the quantitative and qualitative aspects of surface and groundwater bodies, economic and social issues are also worth taking into account in this domain. Moreover, processes, and, this way, the results of different strategies of treatment are subject to significant uncertainty, which again justifies the usage of Bayesian networks. The possibility of including expert decisions may be useful in land use change modeling, since decision-making influence land use change but they are hard to incorporate in the models. Bayesian networks may offer a solution for this issue as well (e.g. Aalders 2008). The conditional probability values are mostly based on known quantitative interrelations, calculations or submodels, but may be provided by expert decisions as well. In the study mentioned, natural aspects influencing land use change are completed by the age distribution of inhabitants and farmers, and the presence of heirs to continue managing the farms. Besides modeling the environmental and landscape processes themselves, Bayesian networks may be suitable for creating decision support systems, through drawing up relevant abstract problems, and particularly through relations of and/or character. These are well-illustrated by the study of Newton (2010), providing a decision network facilitating the Red List classification of species.

Bayesian networks are useful tools in modeling ecosystem services as well (Landuyt et al. 2013). The focus of this approach is on the contribution of environmental attributes to human well-being. It is an interdisciplinary field, where several advantages of Bayesian networks may be utilized. The possibility to include information from different knowledge domains make it possible to evaluate cultural services (Shaw et al. 2016). As the delivery of services is mainly a complex process, all the elements and flows may not be exactly discovered and quantified. The integration of expert knowledge could have a significant added value, as it is the case with several supporting and regulating services (Grêt-Regamey et al. 2013). And the decision support systems may also be developed in an ecosystem service context, with the help of Bayesian networks (Villa et al. 2014). The aim of our paper is the presentation of a decision support oriented application of the Bayesian networks, on the example of a well-known environmental issue, the eutrophication of freshwater systems. The study was carried out in an environmentally sensitive landscape type.

1.3. Eutrophication of surface waters

Eutrophication and the related issues began emerging more vigorously in the middle of the previous century. The nutrient load, increased as a result of human activity, caused deterioration in water quality, which at the beginning manifested itself only in the loss of some functions, and finally it led to the total disappearance of their function in the ecosystem. The initial changes had worried only the experts, but later, when algal blooms, water discoloration, various odours and sometimes fish die-offs (Somlyódy and van Straten 1986), threatening public use occurred, a demand to stop and reverse the process emerged. After the exploration of the drivers different measures were carried out with more or less success in order to prevent the pollution, or in many places, to rehabilitate the water bodies. To find the best management alternatives, it is important to know the effects of different natural and anthropogenic factors in the eutrophication process. It highlights the need for targeted model development.

The lakes and their fish stock have a crucial role in providing many ecosystem services. The filling up process of the karstic lakes causes reduction in the diversity of landscapes and in the value of the scenery. Fish have a significant role in the lakes' ecosystem processes and thus provide important supporting services. The proliferation or decline of fish species on the different levels of the food chain influences the flow of elements and the population of other species, and through terrestrial fish consumers they can also affect the processes of other habitats as well (Holmlund and Hammer 1999).

In the present study we present a Bayesian network examining the eutrophication processes of karstic lakes in Hungary. One aim is to aid decision-making through seeking the influence of the different factors on one important attribute in the eutrophication process, the share of toxic ammonia. Besides that, we would like to give an example for the usage of Bayesian networks with the approach of estimating condition probabilities based on measurement datasets. It might be a usable approach in modeling supporting and other types of ecosystem services too.

2. METHODS AND STUDY AREA

Our research is carried out in the Gömör-Torna karst region. Our goal is the examination of the still waters in the area, and the quantification of the environmental conditions affecting them. These examinations are especially reasonable in karstic areas, as water is an important forming factor of the karstic system, it infiltrates fast, and thus it affects groundwater quality, underground wildlife and formations, especially caves (e.g. speleothem degradation). Contamination from human land use accumulates in the water gathered in different karstic depressions, becoming obvious firstly through the change in water quality, while later it gets transported underground. The fact that in the last few decades several still water bodies got in the state of being filled up or even disappeared in the area, also justifies the investigation. Shallow, small-scale ponds are in increased risk of eutrophication, as in their case the decrease in water quality can be followed by complete termination. In dry, karstic plateaus it means an even greater loss, firstly in terms of human use, but also in terms of decreasing biodiversity, as these ponds serve as the habitat and breeding place of several (often protected) species. The subjects of our investigation were Lake Papverme, Lake Vörös, Lake Kender and Lake Aggteleki. A three-year measurement series was carried out, from which we used data from the year 2009, from April to October.

In our study we examined the changes in the amount of ammonia, which affects the health of fish stock. Ammonia is excreted by plants and animals, and it occurs in the water due to the decomposition of living organisms, industrial emissions, and fertilizer washing-in (Randall and Tsui 2002). Ammonia is a free, non-ionic form, which can permeate through most of the biological membranes, and it affects living organisms as a neurotoxin (Szilágyi 2007). It is formed from ammonium with a higher rate in case of higher pH and higher water temperature. Algae blooms, occurring in waters overloaded with nutrients, cause assimilational alkalisation, leading to alkaline pH. Areas with limestone bedrock are especially vulnerable in this respect, as they may further increase this effect. Therefore, especially in the summer, concentrations toxic to the fish can occur.

The sampling points in the lakes were chosen in 4 or 2 directions according to the points of the compass, near the shores, because in the case of lakes, coastal water quality is decisive. In addition, samples were taken from where inflows reach the lakes. According to this, in the case of Lake Papverme 8, in the case of Lake Tengerszem 4 sampling points were marked, and 2–2 in the case of the other lakes. In the field tests we used a WTW pH/Cond 340i instrument for measuring pH and conductivity. Water temperature was measured with a Hach Lange thermoluminescent dissolved oxygen meter. The measurement of ortophosphate, nitrate and ammonium were carried out in a lab, by a Fia Star 5000 meter. The determination of alkalinity was based on the MSZ EN ISO 9963-1:1998 standard. Chlorophyll-a content was measured by the North Hungarian Environmental Protection and Water Management Inspectorate.

3. RESULTS AND DISCUSSION

3.1. Structure of the model

The probability network (Fig. 2) generated on the basis of our data illustrates one of the possible consequences of euthrophic conditions in the lakes formed and already stabilized as a result of nutrient stress, and its contribution to achieving this consequence. This means the share of toxic ammonia, converted from the potentially present ammonium, which depends on the level of alkalinity and the rise of water temperature. Ammonia, and especially its permanent presence, may be harmful to the health condition of fish, even in small quantities. In those lakes, where fish farming is going on this presents a serious problem, while in other lakes it indicates that the ecosystem is not capable of supporting higher forms of life, so species diversity decreases. The direct contribution of nutrient load to the higher rate of ammonia mainly means a higher ammonium-input, but also works indirectly: the algae proliferating due to the abundance of nutrients shift the pH towards an alkaline environment with their photosynthetic activity, and they also increase turbidity, which increases another important factor of ammonia-formation, the water temperature. The particles suspended in the water (soil, algae) absorb and scatter the sunlight, so the surface temperature of these waters increases (especially at noon) (Paaijmans et. al. 2008). Therefore, the model is built from parameters influencing the ammonia rate directly, and of other parameters affecting them indirectly. These are the anthropogenic impact, expressed by the nutrient supply; the weather conditions, which can enhance anthropogenic effects (SPI drought index); and the air temperature, that directly affects the temperature of water. As an external influence, we can mention alkalinity, which affects the quantity of ammonia through the pH. As alkalinity increases pH, in waters where the level of alkalinity is high, the ammonia is more toxic (Wurts and Durborow 1992). The value of the SPI drought index affects the quantity of the 3 types of nutrients by providing information on the appearance and length of wet and dry periods. Especially during wetter periods the rate of diffuse pollution may increase. The calculation of the index is based on the amount of precipitation of a chosen base period (McKee et al. 1993). Precipitation data are from the meteorological stations of Jósvafő and Silica, the index was calculated for the period 1958-2009. In the model, the values are valid for one-month periods; the categories used are moderately, very, or extremely wet and dry periods.

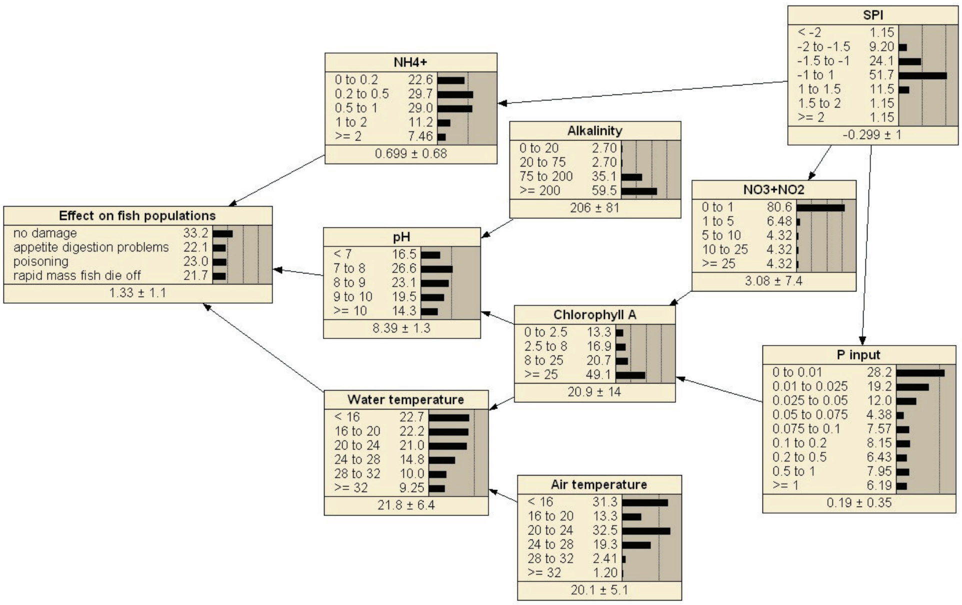


Fig. 2 Euthropication model of the lakes in Gömör-Torna karst

The next step in the model is the chlorophyll-a quantity, which is influenced by the nitrate and phosphorus content. The threshold limit values of ammonium and nitrate are given based on the no. MSZ 12749:1993 standard, in five water quality categories (from excellent to highly polluted). The limit values connected to ortophosphate are given based on the trophic state (Szilágyi and Orbán 2007). The chlorophyll-a levels were classified on the basis of OECD's (1982) trophity-grades. The limit values of alkalinity were determined based on

Wurts and Durborow (1992), according to the ideal level present in a lake ecosystem. The limit values of air and water temperature were designated arbitrarily; the lower limit is the temperature which may be relevant in terms of ammonia formation and durability of persistence. This was also the basis of the pH value categorization. The possible categories of ammonia levels were determined according to the health effects on fish, four classes were made: when there is no damage; when some irritations occur (loss of appetite, or degradation of digestion); when poisoning occurs; and the rapid development of mass fish die-off (Szakolczay 1997). For the latter there was no example amongst our measurements, but our researches led us to the conclusion that in the examined lakes at certain times (especially on summer days), there's a risk of that extreme condition occurring.

3.2. Sensitivity analysis

With the help of the sensitivity test of the program it is possible to determine in the multi-factor system of the constructed Bayesian network, how much each factor influences the final outcome, or the development of any (other) chosen variable. By this method we can test if the constructed structure is adequate or not (this has an especially high importance in the case of Bayesian networks which are based on experts' knowledge). As a result of the sensitivity analysis that was executed on the variable 'effect on fish populations' we found that in accordance with the preliminary assumptions the fish were the most sensitive to the pH and the water temperature. The next most important parameter was the ammonium rate. The chlorophyll-a amount precedes both air temperature and alkalinity. It means that this parameter plays a more important role through its effect on pH and water temperature than the air temperature or the alkalinity. Following air temperature, the drought index is the first parameter that does not have a direct impact on the two most important parameters, but its extensive impact justifies its place in the list. The health state of the fish is the least sensitive to the alkalinity, and to the nitrate and phosphate rate. Considering the initial independent variables in this respect the climate parameters – especially air temperature – are determining, in contrast with the bedrock and the atmospheric CO_2-content, which influence the alkalinity.

When measured data are used, in addition to the determination of the model structures' 'goodness' the result of the sensitivity test can also help to decide which the most important and minimally necessary parameters are to be measured if we want to find out what causes the presence of ammonia in a certain place, how persistent it will be, and where we have to intervene in the system to eliminate the unfavourable state.

We checked the accuracy of the generated model with the help of a test file from the measurement data. During the verification process we eliminated a variable from the system, and the program estimated the result for each test case on the basis of the probabilities set up for the original cases. The estimation was compared with the real value, and an error rate was calculated from it. The error rate compares the number of estimation errors to the total number of test cases. The accuracy of classification can be defined by a confusion matrix, which shows the distribution of misclassifications between the categories. Based on the 15 test cases the error rate was estimated 20%, probably caused by the low number of test data, and the disproportionate distribution of cases between the categories. Based on similar results, Marcot et al. (2001) determined that the actual occurrence of rare events is underestimated, because the forecasts of the model are based on the most probable events. It is also true in case of the data we used, because in our measurements, the state when there is no harm to the fish is dominant. The model estimated this state correctly in every case, but in opposite, instead of the state of 'decline in appetite and digestion' 2 times out of 2, and instead of the case of 'poisoning' once out of 3 'no damage' state was given as a result. For estimating the correctness of the model the software calculates even more indicators, one of them is the so-called 'spherical payoff' (Morgan and Henrion 1990, briefly introduced in. Marcot et al. 2001), its value varies between 0 and 1 (1 means the best possible model performance). In the case of our model the spherical payoff was 0.77 which, considering the low number of available data, is a relatively high value.

3. CONCLUSION

Based on our results, we consider Bayesian networks to be a useful tool in studies of water management and other environmental issues. Evaluations of ecosystem services need methods that can integrate interdisciplinary information. Bayesian networks may fulfil this need, and provide the possibility for developing decision support systems. The condition probability tables can be created based on known relationships, e.g. using measurement datasets. In this study, we gave an example for that, with an eutrophication model of karstic lakes in the Gömör-Torna Karst Region. We focused on a parameter having an effect on the health status of fish populations, which is an important element in the delivery of different ecosystem services. This approach may be used in modeling different types of ecosystem services. The model we built may need to be improved in several different aspects. Firstly, for specifying the different causal relations more precisely the model might need some changes, e.g. the use of some correlations which are not based on our own measurement data. Because of the possibility of including probability-based relations and experts' decisions in the model, Bayesian networks might be capable of handling also economical and sociological data, and might be able to examine the future tendencies as well. In the case of these karstic lakes it may mean applying the method of restoration costs for decision support: with the help of Bayesian networks different alternatives of treatment (for example dredging, protection from erosion, regional drainage) can be compared.

Acknowledgements: This research was supported by the Hungarian Scientific Research Fund in the frames of the project OTKA/048356.

REFERENCES

Aalders I (2008) Modeling land-use decision behavior with Bayesian belief networks. Ecol Soc 13: article 16

Carmona G, Varela-Ortega C, Bromley J (2011) The Use of Participatory Object-Oriented Bayesian networks and Agro-Economic Models for Groundwater Management in Spain. Water Resour Manage 25:1509-1524

Devitt A, Danev B, Matusikova K (2006) Constructing Bayesian networks Automatically using Ontologies. In: Second Workshop on Formal Ontologies Meets Industry (FOMI 2006), Trento, Italy

Gemela J (2001) Financial analysis using Bayesian networks. Appl Stochastic Models Bus Ind 17:57-67

Grêt-Regamey A, Brunner SH, Altwegg J, Christen M, Bebi P (2013) Integrating expert knowledge into mapping ecosystem services trade-offs for sustainable forest management. Ecol Soc 18: article 34

Holmlund CM, Hammer M (1999) Ecosystem services generated by fish populations. Ecol Econ 29:253-268

Kahn CE, Roberts LM, Shaffer KA, Haddawy P (1997) Construction of a Bayesian network for mammographic diagnosis of breast cancer. Comput Biol Med 27:19-29

Landuyt D, Broekx S, D'hondt R, Engelen G, Aertsens J, Goethals PLM (2013) A review of Bayesian belief networks in ecosystem service modelling. Environ Modell Softw 46:1-11

Marcot GB, Steventon JD, Sutherland GD, McCann RK (2006) Guidelines for developing and updating Bayesian belief networks applied to ecological modelling and conservation. Can J Forest Res 36:3063-3074

Morgan MG, Henrion M (1990) Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge Press, New York

Newton AC (2010) Use of a Bayesian network for Red Listing under uncertainty. Environ Modell Softw 25:15-23

OECD (Organization for Economic Cooperation and Development) (1982) Eutrophication of Waters. Monitoring assessment and control. Final Report. OECD Cooperative Programme on Monitoring of Inland Waters (Eutrophication Control). Environment Directorale, OECD, Paris

Paaijmans KP, Takken W, Githeko AK, Jacobs AFG (2008) The effect of water turbidity on the near-surface water temperature of larval habitats of the malaria mosquito Anopheles gambiae. Int J Biometeorol 8:747-753

Randall DJ, Tsui TKN (2002) Ammonia toxicity in fish. Mar Pollut Bull 45:17-23

Shaw E, Kumar V, Lange E, Lerner DN (2016) Exploring the utility of Bayesian networks for modelling cultural ecosystem services: A canoeing case study. Sci Total Environ 540:71-78

Somlyódy L, van Straten G (eds) (1986) Modelling and managing shallow lake eutrophication. With application to Lake Balaton. Springer-Verlag, Berlin

Szakolczay J (1997) Halegészségügyi alapismeretek. [The basics of fish health. (in Hungarian)] In: Tathy B (ed) Halgazdálkodás II. [Fish farming] MOHOSZ [Hungarian National Association of Anglers], Budapest. 457-487

Szilágyi F, Orbán V (eds) (2007) Alkalmazott hidrobiológia. [Applied hydrobiology. (in Hungarian)] Magyar Víziközmű Szövetség [Hungarian Water Utility Association], Budapest

Villa F, Bagstad KJ, Voigt B, Johnson GW, Portela R, Honzák M, Batker D (2014) A Methodology for Adaptable and Robust Ecosystem Services Assessment. Plos One 9:e91001

Wurts WA, Durborow RM (1992) Interactions of pH, carbon dioxide, alkalinity and hardness in fish ponds. Southern Regional Aquaculture Center Publication 464:1-4