The use of database and methods of artificial intelligence – neural networks for observing the changes in the subsoil waters in Beskid Śląski

Konrad Sikora[*]

University of Bielsko-Biała, Willowa 2, 43-309 Bielsko-Biała, Poland

Abstract

The article deals with the use of relational database and application of neural networks for observing the influence of the environment and human activities in shaping the chemism of subsoil waters. The possibility of using the database has been investigated with the use of chosen waters coming from the saturation area of Beskid Ślaski. The aforementioned area is of great regional importance for providing drinkable water. The relational database is a kind of integrated system of collecting, processing and interpreting the results of scientific research. Scientifically described software allows for collecting, selective search and analysis of the data refering to any physico-chemical parameters of water samples. The standard equipment of IBM PC is required. The material collected is analysed statistically and can be presented graphically in order to illustrate the accomplished activities in a better way. Owing to the use of software, it is possible to interpret the many – years data and compare them taking into consideration all accessible elements of the environment and the changes in urban planning or even the constantly changing legal norms. The database has the open character and can be completed with any amount of data from different spheres of knowledge. At present the research is carried out on application of artificial intelligence - neural networks for correlating the research results of the physico-chemical samples collected in the database with the factors of the environment.

1. Characterization of the area in question and the assumptons of system formation

Beskid Śląski and Mały is the most protruding part of Polish Flisch Carpathians (Beskidy). This area is of essential multilevel importance in the fields of nature and agriculture, as well as tourism and recreation. We deal here with the unique values of nature and landscape. The preservation order was put in the area of both massifs within the Landscape Parks of Beskid Śląski and Beskid Mały within which there are a lot of objects under the preservation of nature of both biotic and abiotic character.

---

[*]*E-mail address*: konradeus@poczta.onet.pl

Aforementioned area is also an important source for providing drinkable water on a local and regional scale. The Polish part of Beskidy is a very important source of drinkable water for the agglomeration of Górny Śląsk and Bielsko. There are water reservoirs in Wisła Czarne, Goczałkowice, Wapienica, Cascade of the Soła River.

The area in question is really in danger, not only on account of the local pollution but also because of the variety of different pollution deposited on the mountain slopes. Threat for the environment forces us to make their thorough observation. Consequently, the identical and integrated system of monitoring of the subsoil water should be worked out for the whole region.

One of the most significant tasks of this system will be the estimation of the influence of the environment and the local pollution on the quality of subsoil water in this area. The monitoring system is extremely difficult and requires making wise decisions which refer to assigning the functioning of the observation system, the proper control of the quality of obtained results, their processing, collecting, interpretation and accessibility [1,2].

2. Description of the programme

The data obtained from the physico-chemical analysis was rendered accessible and became the basis for constructing the database "Database of the water composition of Beskid Morawski and Śląski" and later on "Computer system of estimating the water composition coming from Beskid Śląski and Morawski". The software has been made in Microsoft Access XP, being the part of Microsoft Office XP, operating in Windows.

This is the relational database which consists of the four tables interrelated with one another by means of relations and a few forms, inquiries, macros, reports. The aforementioned database includes the results of physico-chemical analysis of waters. The results were introduced into the database in the following domains: pH, C, HCO_3, SO_4, Cl, NO_3, NO_2, Ca, Mg, Na, K, Al, B, Ba, Fe, Mn, Li, P, Sr, Zn, Pb, Cd, SiO_2 [3].

The discussed software has been divided into the following parts:
1. Introducing the database – with the help of panels for introducing data it is possible to introduce the data and to modify it.
2. Reports – present printed data from the database. There is a possibility of choosing the data print of the particular water sample and the description of the other elements connected with this sample.
3. Information retrieval – this panel was invented in order to enable the fast access to the information that can be of some interest [2,3].

There is an innovative solution applied in the software connected with the use of conditional macros for solving the problem of introducing the minimal and maximal values of the norm for quality of water. The picture below presents the example used in the constructed database (Fig. 1). If the partcular condition

remains unfulfilled (exceeded norm) the NewsWindow is displayed (Fig. 2), if the value of introducing data is about the norm the confirmation in the form of NewsWindow is displayed (Fig. 3).

	Warunek	Akcja
	[Forms]![Wyniki Formularz]![pH]<=6,5	OknoKomunikatu
	…	OknoKomunikatu
	[Forms]![Wyniki Formularz]![pH]>=9,5	OknoKomunikatu
	…	OknoKomunikatu
	[Forms]![Wyniki Formularz]![pH]>=6,5	
	[Forms]![Wyniki Formularz]![pH]<=9,5	

Fig. 1. The conditional macro Window of pH field in the form [3]
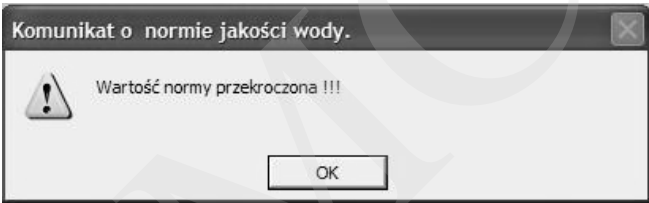

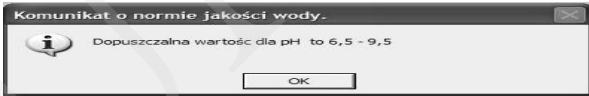Fig. 2. The example of the NewsWindow which does not fulfil the values norm [3]

Fig. 3. The example of the NewsWindow which fulfils the norm [3]

3. Application of *Statistica* neural networks for correlating
the research results

The package STATISTICA Neural Networks is used to stimulate the artifical neural networks. It has the large counting possibilities and the friendly way of exploitation for the user. Neural networks are the new informatics tool which enables the construction of a non-linear black box. It solves the complex and difficult to identify tasks both classificatory and regressive.

The application of neural networks to analyse complex data among other things the comparison of environmental parameters with the results of physico-chemical data of the water samples is really advantageous. It leads to the construction of the black box based on the property of the data but not on the arbitrary hypothesis formed by explorer black box. In the applied programme there are the best accessible algorithms. They are used to teach the networks automatic designing its structure and choosing the best set of entrance/in variables. The practical use of the (discussed) software is justified by the fact that neural networks are used in the places where there is dependence between the input and output variables. However, they are especially useful in searching

the dependence between input and output which is very complex and it is difficult to formulate this dependence with the use of common statistical methods.

During the process of building black box, the neural network is taught by means of examples, forming the black box based on the teaching data. The teaching data consists of several examples and each of them includes the values of many variables. Most variables are in/entrance data, although some are exit/out data.

The purpose of the neural networks is to learn – on the basis of gathered examples in the database – the formal dependence between input and output variables.

During the whole process of teaching the neural networks the built – in tool for data import (ODBC mechanism) or a set of data in Excel is used. This mechanism significantly increases the comfort of work with the software because of the sets of data already collected.

The taught neural network can be used in many different ways – for example to indicate the unknown values of the exit variables (out), which correspond to the known entrance variables (in). On the basis of the above mentioned comments, it can be said that the most important stage of analysis with the use of the neural network is to collect the proper set of data.

The biggest problem which has been solved recently while building the black box is the existance of many entrance data and the problem with changing them into the numeral ones [2,4].

Conclusions

Elaborated computer database allow for collecting and selective searching for data concerning subsoil waters in the area of Beskid Śląski and Beskid Mały. Owing to the application of this kind of programme we are able to carry out the interpretation of the data accumulated in the database coming from the many years' periods of their interpretation and the comparison of physico-chemical data from different places of the taken samples. In the computer programme there has been collected the information about the elements of the environment affecting the chemism of waters being the significant reservoir of drinkable water area under examination.

The software presented has been constructed on the basis of Microsoft Access Microsoft Office package, working in Windows. This tool was the most appropriate for this kind of software, for the pure practical reasons.

The elaborated programme is being constantly modified in order to complete the database with interpretative module and statistical one. The interpretative module presents the changes of the definite values of the specific element in the period of time or against a background of other indices. The statistical module is

used for the advanced processing of the gained data and their application in other programmes, or their storage in order to use them later.

The current tool that is used for this kind of tasks is the programme Statistica, the neural networks black box. Owing to the system that is being built the current research results are correlated taking into account the constantly changing surrounding and the statistical processing of the obtained data is being carried out.

The idea of the research is to obtain the completely compurerized system, whose aim will be to process the data completely together with its graphic presentation with the use of the most modern methods of information technology.

Acknowledgements

References

[1] Sikora K., Zawierucha L., Motyka J., Myrczek J., *Baza danych składu chemicznego wód Beskidów Morawsko-Śląskich*, Materiały z XII Konferencji Naukowo-Technicznej Zapobieganie Zanieczyszczeniu Środowiska, Szczyrk, (2004), in Polish.

[2] Sikora K., Zawierucha L., Motyka J., Myrczek J., *Skład wód podziemnych w wybranych punktach pomiarowych w Beskidzie Śląskim – badania i interpretacja z wykorzystaniem bazy danych*, Materiały z XIII Konferencji Naukowo-Technicznej Zapobieganie Zanieczyszczeniu Przekształceniu i Degradacji Środowiska, Szczyrk, (2005), in Polish.

[3] Sikora K., *Baza danych składu wód Beskidu Śląsko – Morawskiego*, Akademia Techniczno – Humanistyczna w Bielsku – Białej,", Bielsko-Biała, (2003), in Polish.

[4] Statistica Neural Networks – user guide, StatSoft, (2001), www.statsoft.pl