Mine Water Environ (2012) 31:138–146
DOI 10.1007/s10230-012-0178-4

TECHNICAL COMMUNICATION

Water-inrush Assessment Using a GIS-based Bayesian Network for the 12-2 Coal Seam of the Kailuan Donghuantuo Coal Mine in China

Dong Donglin · Sun Wenjie · Xi Sha

Received: 8 January 2012 / Accepted: 31 March 2012 / Published online: 17 April 2012

Abstract The Donghuantuo coal mine is geologically unusual, with 60 normal faults, 18 reverse faults, and 1 syncline. The coal seam floor is highly fractured and the fractures act as conduits for groundwater, which flows from the Ordina limestone aquifer into the no. 12 coal seam. From 2005 to 2010, there were 7 water-inrush disasters through the floor of this coal seam. The largest water-inrush event exceeded 63 m^3/min; there are five points where the water-inrush continues to exceed 1.0 m^3/min. Comprehensive modeling of the probability of water-inrush through the floor is required to reduce the likelihood and severity of such events. The water-inrush situation was assessed using a GIS-based Bayesian network (BN). In the developed BN-GIS model, the geometry of the coal mine working face was incorporated in suitable detail and resolution. The results of the modeling compared well with field water-inrush observations. Based on documented water-inrush events, the accuracy of the fit of the model data is 83.4 %, and the probability of making an incorrect prediction is less than 0.5, which means that using this method could significantly enhance coal production at the mine.

Keywords BN-GIS method · Coal mine · Probability assessment · Water inrush

Introduction

The Donghuantuo mine, which is located near the town of Hancheng in the city of Tangshan in Northern China, was established in 1975 by the Kailuan Coal Mining Bureau. This mine is the newest mine in the Kailuan coalfield, and is the highest producing mine in the coal field (Zhong and Shen 2000). The Donghuantuo coal mine is geologically unique, with 60 normal faults, 18 reverse faults, and a syncline. The coal seam floor is highly fractured and the fractures act as conduits for groundwater, which flows between the Ordina limestone aquifer and the no. 12 coal seam. From 2005 to 2010, 7 water-inrush disasters occurred (Table 1) in this seam, which is conventionally subdivided into two layers, 12-1 and 12-2.

The mine is highly profitable despite its water problems. Abundant information exists about its geology, hydrogeology, and its water-inrush events. The largest water-inrush event, in March 1995, exceeded 63 m^3/min; there are five points where the water-inrush continues to exceed 1.0 m^3/min. At present, the 12-2 coal layer is the main working seam; it accounts for 45 % of all of the water entering the mine.

In order to improve mine safety at the Donghuantuo coal mine, several studies have been undertaken by university researchers, private consultants, and China's National Academy of Sciences. Initial studies completed by the China University of Mining and Technology (Wu et al. 2000, 2004) characterized groundwater flow at Donghuantuo, but many questions still remain. Various studies have been conducted to investigate water-inrush types in

D. Donglin · S. Wenjie (✉)
College of Geoscience and Surveying Engineering,
China University of Mining and Technology,
Beijing 100083, China
e-mail: sjerr@163.com

D. Donglin
e-mail: ddl@cumtb.edu.cn

X. Sha
School of Engineering and Technology, China University
of Geosciences, Beijing 100083, China

Table 1 Water-inrush disasters from the no. 12 coal seam in the Donghuantuo coal mine

No.	Time	Types of inrush	Source of water inrush	Water-inrush channel	Maximum water yield (m^3/min)
II-87	2005.4.29	Inrush from floor	Between the no. 12 and 14 coal seams	Fissure	0.9
II-89	2005.11.16	Roof collapse	Roof water	Fissure	0.5
II-90	2006.1.1	Roof collapse	Roof sand fissure water	In-break fissure	0.55
II-94	2007.10.13	Tunnel roof leaking	Coal seam 12-1	Fissure	1.70
II-95	2009.7.10	Tunnel floor leaking	Coal seam 12-2	Fissure	0.5
II-96	2009.8.10	Tunnel floor leaking	Coal seam 12-2	Fissure	0.5
II-97	2010.01.27	Tunnel floor leaking	Coal seam 12-2	Fissure	2.0

the Donghuantuo coal seam, including water-inrush from the roof and floor of the no. 12 coal seam, and collapse caused by limestone dissolution.

To predict the probability of a water-inrush occurring during mining, it is critical to first understand what induces such an event. In particular, the distribution of groundwater, the abundance, orientation, and degree of interconnectivity of fractures, the mechanical strength and thickness of the aquifer, as well as the coal seam depth are all critical to develop a strong predictive model for water inrush. Many water-inrush events in the coal mine were investigated in order to gain a better understanding of the probability of water-inrush. Of special concern was whether a prior water-inrush incident influences the probability of future water-inrush events. There was a tremendous amount of disordered information but it was not clear how it could be used to predict water-inrush events. Bayesian networks (BN) are useful tools in such circumstances. By combining BM with a geographic information system (GIS), which has powerful spatial analysis ability, we attempted to build a model of water-inrush assessment that would provide mines in the Kailuan coalfield and perhaps, by example, elsewhere, the ability to predict and control water-inrush in the coal seam floor. The purpose of this investigation was to confirm the likelihood of water-inrush events and predict occurrence probability in the future.

The Principle of Water-inrush Assessment Based on BN + GIS

The core of BN + GIS water-inrush situation assessment in the floor of the coal seam is deducing water-inrush trends. Then, GIS spatial analysis is used to do an overlay analysis.

BN is a kind of probabilistic network, based on a graphical network of probabilistic reasoning. This network includes two parts: a directed acyclic graph (DAG), which is made up of parameters (nodal points) and a directed arc that connects nodal points, and a conditional probability table (CPT) of each parameter and all of its upper (parental) generation relationships. Bayes' formula is based on this probabilistic network.

Set S is sample space of test E and X is an event of E. Y_1, Y_2, …, Y_n is a group of events of E, and satisfy:

1. $\sum_{i=1}^{n} Y_i = S$;
2. Y_1, Y_2, …, Y_n are mutual exclusive;
3. $P(Y_i) > 0$, $i = 1, 2, \ldots n$. Then the Bayes' formula is:

$$P(Yi \ominus X) = \frac{P(X \ominus Yi)P(Y)}{\sum_{j=1}^{n} P(X \ominus Yj)P(Yj)} \quad i = 1, 2, \ldots n.$$

In this study, X stands for a specific water-inrush incident and P is the occurrence probability of Y_i under the condition of X. BN structure study uses training sample data, and finds out the best network topological structure for data and priori knowledge.

The coupled BN + GIS assessment method that we used to assess a water-inrush situation consists of:

1. Using GIS software to process basic functions like data acquisition and spatial analysis to build spatial and property data bases.
2. Using the spatial analysis function of GIS software to process multiple overlapping factors and associate spatial and property information, and to confirm the assessment unit.
3. Building a BN model, developing a design to realize the BN parametric study, inference module, and GIS spatial analysis module.
4. Using the GIS-derived property data base to process the parametric study, and then developing network inference, calculate the forward and reverse probability of water-inrush, and then create and renew the spatial data base.
5. Considering the model-calculated probability as an assessment factor, use the GIS spatial data base and GIS functions to conduct spatial queries, analysis, and display the results.

Water-inrush Hazard Assessment of the Donghuantuo 12-2 Coal Seam Floor

There are many factors that affect water inrush in the coal seam floor. These factors interact and make water-inrush incidents complex. According to the specific conditions at Donghuantuo and existing research, water-inrush incidents in the Donghuantuo 12-2 coal seam are probabilistic results of tectonic characteristics (including fault density, fault throw, fault property, the extent of fissure, fault hydraulic conductivity, and the extent of fold), aquifer conditions (e.g. aquifer pressure, water abundance, and permeability), aquiclude conditions (e.g. aquiclude thickness and strength), and mining parameters (such as mining thickness, mine pressure). The specific factors of water-inrush in Donghuantuo 12-2 coal seam floor are shown in Fig. 1.

To confirm BN nodal points, thematic maps of faults, water-inrush events, and a composite layer (which is produced by elementary overlap analysis) are considered together, and spatial associate overlapping analysis is used to produce a new composite layer. On this basis, fault density, property, throw, and water-inrush history are considered for each assessed unit, and finally, a property sheet (Fig. 2; Table 2), which includes all assessed properties, is built.

Confirming BN nodal points of coal seam floor water-inrush hazard assessment consists of building an evaluation index system. Clarification of coal seam water-inrush inducing factors involves combining GIS quantified water-inrush assessment and overlapping analysis, and finally confirming BN hazard assessment model nodal points in the coal seam water-inrush area.

In this project, 13 selected evaluation indexes were used to assess tectonic characteristics, aquifer and aquiclude conditions, and mining activities, as well as whether water inrush has occurred in the coal seam floor were used as nodal points for the BN hazard assessment model.

Discretization of Numerical Attributes

BN, like most machine learning, uncertain reasoning, and sorting algorithms, demands discretization of continuous numerical attributes (Rodrigues et al. 2000). Continuous variable discretization is the key problem to be solved because the relationship between nodal points in BN is expressed by conditional probability tables.

In Table 2, attributes of five nodal points-fault density, fault throw, the extent of fissuring, hydraulic conductivity, and aquiclude strength are continuous values. In this study, the K-means algorithm was used to discretize continuous attributes of these five nodal points. The K-means algorithm is the most well established classification algorithm, and plays an important role in the clustering algorithm (Brenda 2001; Gudrun et al. 2004; Jozef 2001; Sanguesa and Burrell 2000). Its advantages include its speed and succinctness, and the fact that it can process multiple parameters and large amounts of sample data effectively. The basic principles of the K-means algorithm are:

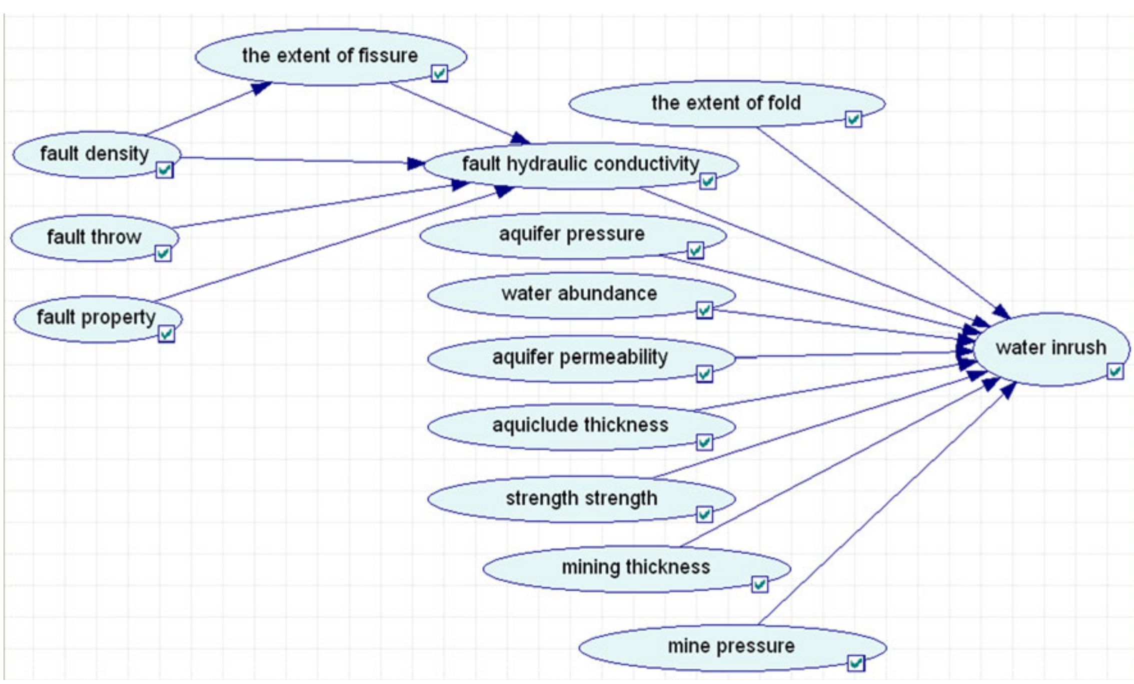


Fig. 1 BN topology structure of coal seam floor water-inrush hazard assessment

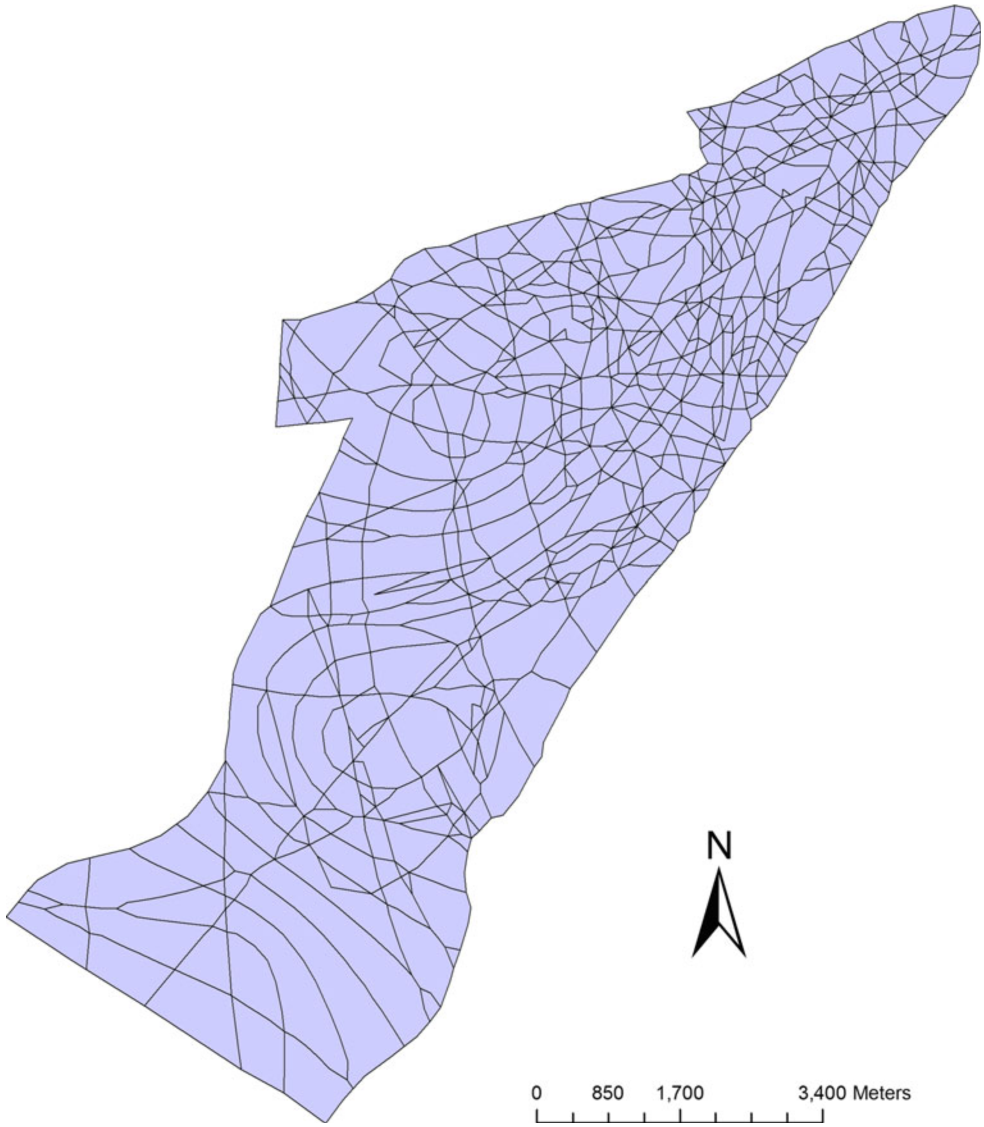


Fig. 2 The associate overlapping process of thematic maps

Viewing K as a parameter, set n objects into K clusters (classifications), and make different clusters with maximum similarity and minimum degrees of difference.

Confirming K (the number of clusters) involves selecting any K objects, where each object represents the center or mean value of a cluster. For the rest of the objects, the distance between the centers of one cluster and the nearest cluster. Then, the mean values of all of the objects in every cluster are recalculated, and the calculated value is set as the new cluster center. This process is iteratively repeated until the criterion functions converge.

$$E = \sum_{i=1}^{k} \sum_{p \in c_i} |p - m_i|^2$$

Here E represents the sum of the square deviation of all objects, p is data point, and m_i is the mean value of cluster c_i. According to this standard, the generated results tend to be compact and independent.

Thus, the process of K-means algorithm can be described as:

1. Choose any K objects as the initial center of clusters;
2. Divide all objects into clusters with minimal similarity;
3. Calculate the mean value of all objects, and distribute these values to the most similar clusters;
4. Repeat until there is no more change.

Continuous property values were set into 8 classifications in this project (K = 8) by incorporating the property data base (Table 2) and the amount of the BN conditional probability. The calculated discretization results are given in Table 3.

Table 2 Nodal points' name and property of Bayesian network

BN nodal points' name	Data source	Property
Fault density (FD)	Fault thematic map	Continuous data
Fault throw (FF)	Fault thematic map	Continuous data
Fault property (FK)	Fault thematic map	No fault, normal fault, or reverse fault
The extent of fissure (FL)	Fault thematic map	Continuous data
Drape strength (FS)	Drape strength thematic map	1,10
Fault hydraulic conductivity (FT)	Fault hydraulic conductivity thematic map	0, 1, 2, 3, 4, 5, 6, 7, 8, 10
Aquifer water stress (WB)	Aquifer water stress thematic map	30, 40, 50, 60, 70, 80, 90, 95
Aquifer water yield property (WA)	Aquifer water yield property thematic map	Weak, medium, strong
Aquifer permeability (WP)	Aquifer permeability thematic map	Continuous data
Aquiclude thickness (AT)	Aquiclude thickness thematic map	2, 3, 4, 6, 8, 10
Aquiclude strength (AS)	Aquiclude strength thematic map	Continuous data
Mining thickness (MT)	Mining thickness thematic map	0, 1, 1.2, 2
Mine pressure (MP)	Mine pressure thematic map	200, 300, 400, 500, 600, 700, 800, 900, 950
Water-inrush (I)	Water-inrush points thematic map	True (water inrush in assessment unit)
		False (no water inrush in unit)

Building a Bayesian Network

Building a BN has two aspects: construction of the topological structure and parameter distribution. This can be done three different ways:

1. The nodal points, topological structure, and the distribution of parameters are independently confirmed by experts. Obviously, this completely depends on the knowledge of the experts. Because of limited knowledge acquisition, there construction BNs and accumulated data can deviate significantly.
2. The nodal points and topological structure are confirmed by experts, while the BN parameters are learned from trained sample data using a machine-learning algorithm. If the relationship between these parameters is obvious, this method can greatly improve study efficiency.
3. The nodal points are confirmed by the field experts, while the BN structure and parameters are learned by mass trained data. This construction is driven by data, with stronger adaptability. With the development of data mining, machine learning, and artificial intelligence, this method is now possible. How to learn the BN structure and parameters from mass data is the focus of much research.

This project has used the second BN approach; that is, experts will confirm the BN nodal points and topological structure, and the BN probability distribution is being calculated. This involves:

1. Training sample data

579 assessment units were formed after overlapping thematic maps. This project used the properties of these 579 assessment units to construct the total sample. In order to verify the results, the total sample was divided into two parts. The training sample and validation sample should have a ratio of 2 to 1. In order to guarantee the validity of the results, the two parts must have separate water-inrush points.

2. Topological structure

The BN topological structure of water-inrush hazard assessment was constructed by experts (Fig. 1).

3. Parameter learning

After a model topological structure is constructed, the parameter distribution of BN coal seam hazard can be

Table 3 The results of Bayesian network continuous data discretization

Discretization results	Fault density	Fault throw (m)	Extent of fissure (m)	Hydraulic conductivity	Aquiclude strength
s1	<5.27e-005	<−15.5	<227.774	<1.5	<0.84
s2	5.27e-005 ≈ 8.28e-005	−15.5 ≈ 13.5	227.774 ≈ 410.628	1.5 ≈ 2.5	0.84 ≈ 0.93
s3	8.28e-005 ≈ 0.00011475	13.5 ≈ 27.5	410.628 ≈ 555.862	2.5 ≈ 3.5	0.93 ≈ 1.01
s4	0.00011475 ≈ 0.000144617	27.5 ≈ 46.5	555.862 ≈ 681.542	3.5 ≈ 4.5	1.01 ≈ 1.08
s5	0.000144617 ≈ 0.000164617	46.5 ≈ 62	681.542 ≈ 835.717	4.5 ≈ 5.5	1.08 ≈ 1.14
s6	0.000164617 ≈ 0.000189825	62 ≈ 87	835.717 ≈ 1164.88	5.5 ≈ 6.5	1.14 ≈ 1.19
s7	0.000189825 ≈ 0.000215603	87 ≈ 123.5	1164.88 ≈ 1436.05	6.5 ≈ 7.5	1.19 ≈ 1.23
s8	>0.000215603	>123.5	>1436.05	>7.5	>1.23

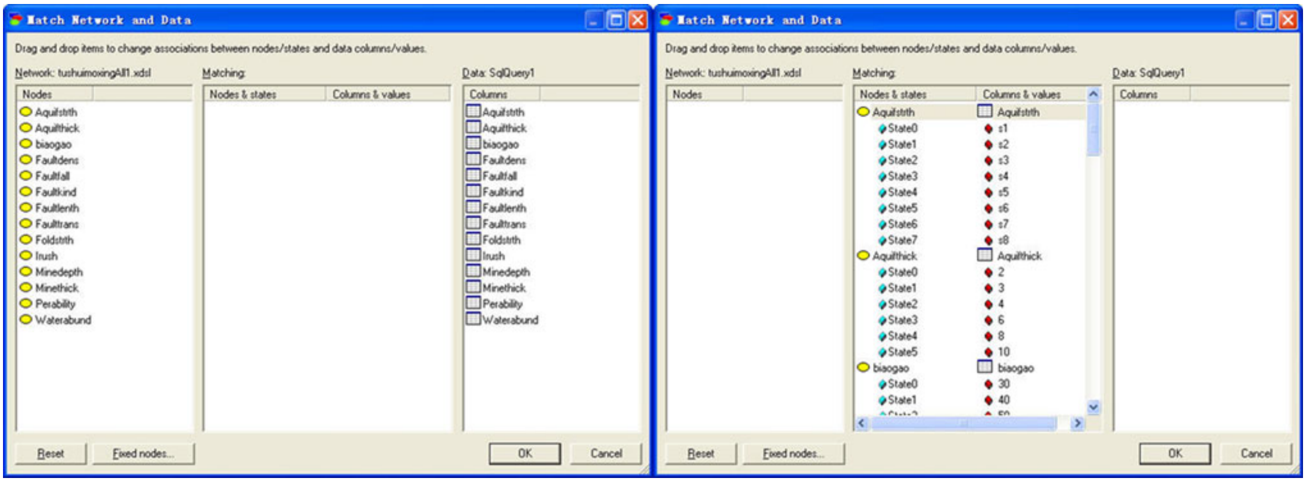


Fig. 3 Parameter learning interface of Genie 2.0

assessed by estimating the maximum likely training data. The parameter learning interface of Genie 2.0 is shown in Fig. 3. The BN nodal points and their status (their properties and their values) are separately entered in data bases and then processed using parameter learning. Table 4 shows the parameter distribution of some of the nodal points after parameter learning.

So far, we have addressed the building of the Bayesian network of the water-inrush hazard assessment for the coal seam floor. The next portion of this paper explains and analyzes this network.

Model Analysis

The purpose of BN model analysis is to further excavate inner implicit information, which is beneficial to abundantly understand problems solved by this model and better use BN to infer.

Compared to other artificial inference models, the key characteristic of BN is that any nodal point can be input nodes and others can be output nodes. According to this characteristic, we consider nodal points with “true” status as input nodes, while the others are output nodes. Observing properties of other nodes when the water-inrush point is in “true” status, we can judge whether the given probability is right. This is virtually the process of diagnostic reasoning. According to the results displayed in Table 5, when the fault density is between 0.000164617 and 0.000189825, the fault throw is larger than 123.5 m, the fault property is normal, the extent of fissure is between 1164.88 and 1436.05 m, the drape strength is 1, the fault hydraulic conductivity is 8, the aquifer’s water pressure is 7 Mpa, the water yield property is strong, the aquifer water permeability is between 5.5 and 6.5, the thickness of aquiclude is 2 m, the strength of the aquiclude is less than 0.84 Mpa, the mining thickness is 10 m, and the mine pressure is 500 Pa, so that water-inrush occurs easily. It is beneficial for us to recognize the combination of different factors that induces water-inrush. In fact, using these factors, we can predict where water-inrush is likely to occur.

Table 4 Prior probability of water-inrush hazard assessment based on BN-GIS in the 12 coal seam floor

P (FK)		P (WB)		P (WP)		…	P (AT)		P (AS)		P (MP)	
Nothing	0.7720	30	0.0483	Weak	0.5993	…	2	0.1485	s1	0.2141	200	0.0915
Normal	0.2020	40	0.0742	Medium	0.26559	…	3	0.2452	s2	0.1934	300	0.1053
Reverse	0.0259	50	0.0759	Strong	0.1347	…	4	0.3143	s3	0.1122	400	0.1934
		60	0.1278			…	6	0.2348	s4	0.1053	500	0.2227
		70	0.1951			…	8	0.0500	s5	0.0932	600	0.1226
		80	0.1537			…	10	0.0069	s6	0.1433	700	0.1036
		90	0.1882			…			s7	0.1295	800	0.0846
		95	0.1364			…			s8	0.0086	900	0.0621
						…					950	0.0138

Table 5 Maximal correlation and conditional probability of other nodal points under Bayesian network water-inrush nodal points with true status

Nodal points' name	Maximal correlation	Conditional probability (%)
The extent of fissure	s7	14
Fault hydraulic conductivity	8	16
Drape strength	1	80
Fault density	s6	14
Fault property	Normal	38
Aquifer water stress	70	20
Fault throw	s8	16
Aquifer water yield property	Strong	38
Aquiclude thickness	2	40
Mine pressure	500	28
Aquifer permeability	s6	25
Mining thickness	10	97
Aquiclude strength	s1	19

Bayesian Network Reasoning

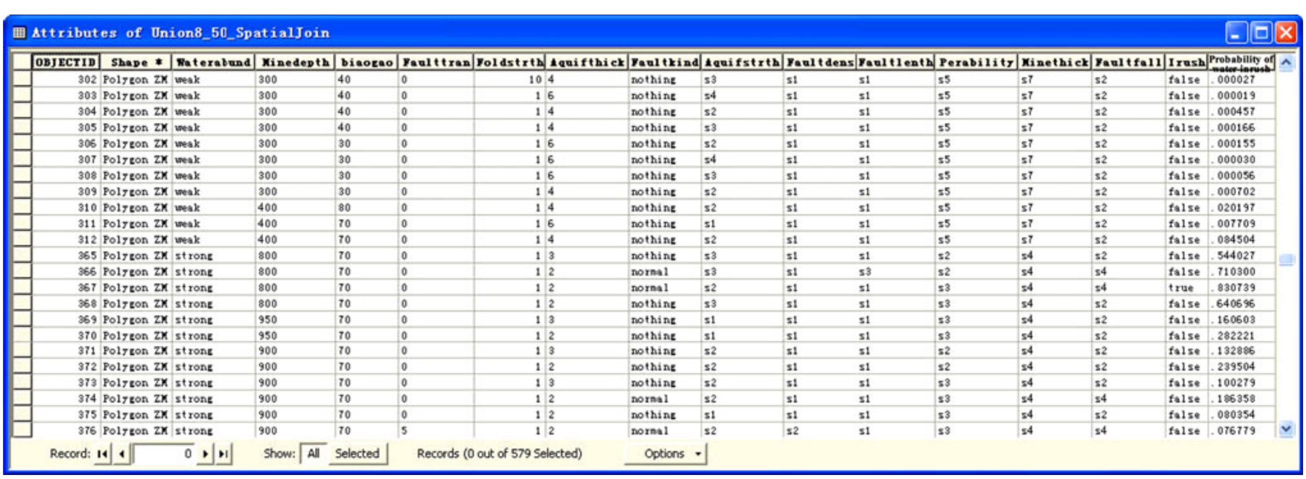
Bayesian network inference is the key context of BN applications. In this project, it is being used to establish the estimated probability of water inrush events occurring at other locations (output nodes) by comparing their values with the status of known water-inrush locations. This requires that the property data of every assessment unit be entered into the BN, and that water-inrush probabilities be calculated for all assessment units. Figure 4 presents the property data base after probability inference, showing the properties of the assessment units and water-inrush probability.

Based on their diverse water-inrush probabilities, all of the assessment units were assigned colors using GIS that represented the BN inference results (Fig. 5), where the probabilities, from high to low, are represented by gradations of three different colors, i.e. yellow, orange, and red. The region with a high water-inrush probability, between 0.61 and 0.93, is colored red, with most of this area above 0.70; the reddest region represents the highest probability (between 0.81 and 0.93) of a water-inrush event, while the lowest yellow color indicates a low probability (between 0.00 and 0.10).

All of the red regions lie in two large areas. One is located beneath the villages of Xicao, Donghuantuo, and Xihuantuo. This area, near Faults 5 and 7 (Fig. 5), has a water-rich aquifer and a relatively thin and weak aquiclude. Therefore, water-inrush events occur easily in this area. The other is located beneath the villages of Daqituo and Xiaoqituo, south of Fault 27, Faults 30–32, and Fault 34. This area has a water-rich, high-pressure aquifer. Therefore, it's also an area where water-inrush occurs easily.

Three sectors have a moderate water-inrush probability (colored orange), with a probability between 0.34 and 0.61. The first is located in the southeastern part of the mine beneath the villages of Nancao and Sanshen. The second area is large and is located beneath the villages of Darongge and Xiaorongge; it has a high aquifer pressure and a thin aquiclude. The third area is dispersed and contains many faults.

The three low water-inrush probability areas (colored yellow) have a water inrush probability between 0 and 0.34. The first two areas are relatively small: one is located in the northern portion of the mine, with low aquifer water stress or underground pressure and the second lies is in the middle of the mine. The third is located in the southwestern part of the mine; the aquifer there has a low coefficient of permeability and the aquiclude is strong.



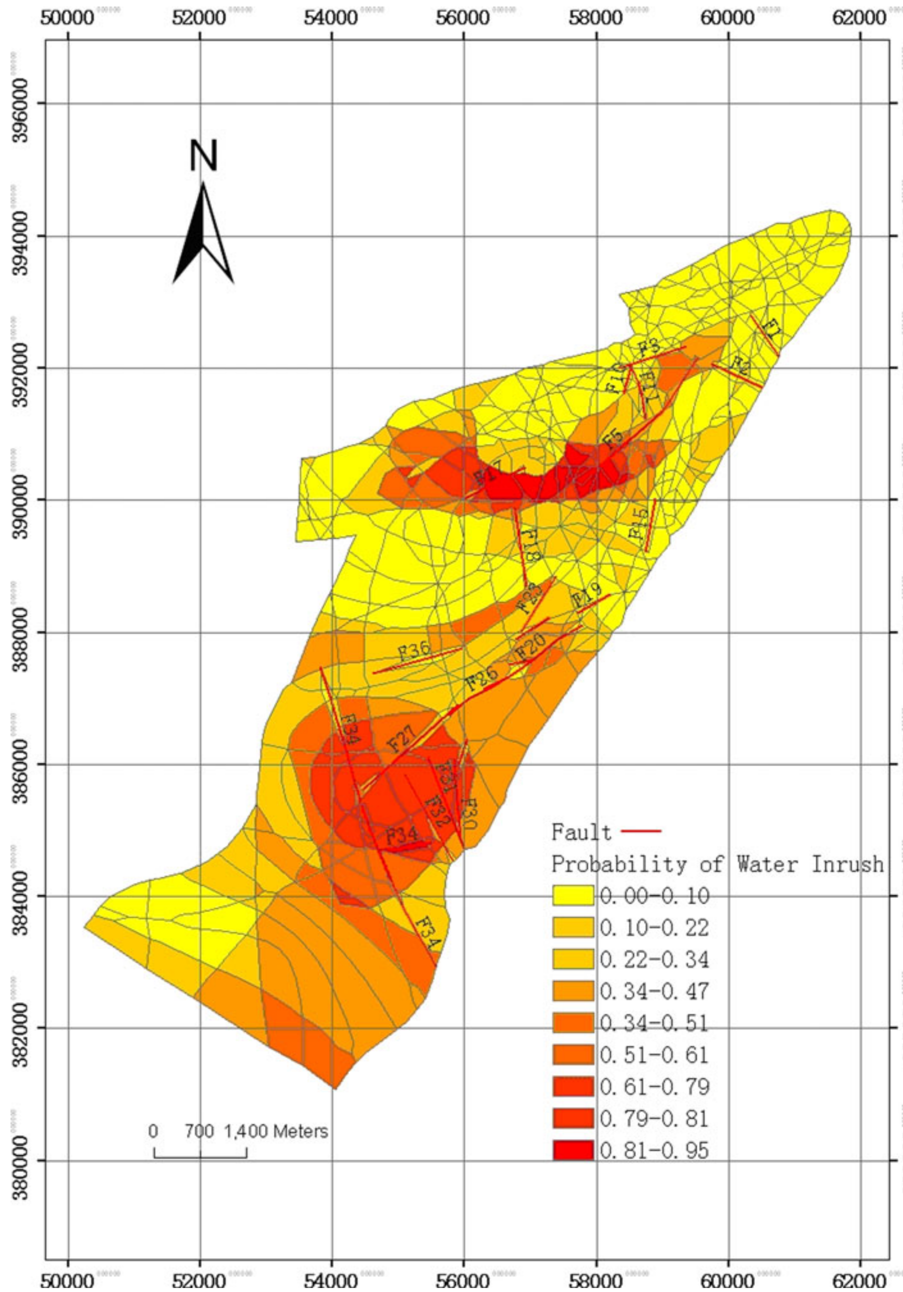
Attributes of Union8_50_SpatialJoin

OBJECTID	Shape *	Waterabund	Minedepth	biaogao	Faulttran	Foldstrth	Aquifthick	Faultkind	Aquifstrth	Faultdens	Faultlenth	Perability	Minethick	Faultfall	Irush	Probability of water inrush
302	Polygon ZM	weak	300	40	0	10	4	nothing	s3	s1	s1	s5	s7	s2	false	.000027
303	Polygon ZM	weak	300	40	0	1	6	nothing	s4	s1	s1	s5	s7	s2	false	.000019
304	Polygon ZM	weak	300	40	0	1	4	nothing	s2	s1	s1	s5	s7	s2	false	.000457
305	Polygon ZM	weak	300	40	0	1	4	nothing	s3	s1	s1	s5	s7	s2	false	.000166
306	Polygon ZM	weak	300	30	0	1	6	nothing	s2	s1	s1	s5	s7	s2	false	.000155
307	Polygon ZM	weak	300	30	0	1	6	nothing	s4	s1	s1	s5	s7	s2	false	.000030
308	Polygon ZM	weak	300	30	0	1	6	nothing	s3	s1	s1	s5	s7	s2	false	.000056
309	Polygon ZM	weak	300	30	0	1	4	nothing	s2	s1	s1	s5	s7	s2	false	.000702
310	Polygon ZM	weak	400	80	0	1	4	nothing	s2	s1	s1	s5	s7	s2	false	.020197
311	Polygon ZM	weak	400	70	0	1	6	nothing	s1	s1	s1	s5	s7	s2	false	.007709
312	Polygon ZM	weak	400	70	0	1	4	nothing	s2	s1	s1	s5	s7	s2	false	.084504
365	Polygon ZM	strong	800	70	0	1	3	nothing	s3	s1	s1	s2	s4	s2	false	.544027
366	Polygon ZM	strong	800	70	0	1	2	normal	s3	s1	s3	s2	s4	s4	false	.710300
367	Polygon ZM	strong	800	70	0	1	2	normal	s2	s1	s1	s3	s4	s4	true	.830739
368	Polygon ZM	strong	800	70	0	1	2	nothing	s3	s1	s1	s3	s4	s2	false	.640696
369	Polygon ZM	strong	950	70	0	1	3	nothing	s1	s1	s1	s3	s4	s2	false	.160603
370	Polygon ZM	strong	950	70	0	1	2	nothing	s1	s1	s1	s3	s4	s2	false	.282221
371	Polygon ZM	strong	900	70	0	1	3	nothing	s2	s1	s1	s2	s4	s2	false	.132886
372	Polygon ZM	strong	900	70	0	1	2	nothing	s2	s1	s1	s2	s4	s2	false	.239504
373	Polygon ZM	strong	900	70	0	1	3	nothing	s2	s1	s1	s3	s4	s2	false	.100279
374	Polygon ZM	strong	900	70	0	1	2	normal	s2	s1	s1	s3	s4	s4	false	.186358
375	Polygon ZM	strong	900	70	0	1	2	nothing	s1	s1	s1	s3	s4	s2	false	.080354
376	Polygon ZM	strong	900	70	5	1	2	normal	s2	s2	s1	s3	s4	s4	false	.076779

Record: 0 Show: All Selected Records (0 out of 579 Selected) Options

Fig. 4 Property data base of Bayesian network inference result

Fig. 5 The water-inrush hazard assessment chart for the Donghuantuo coal seam floor



Above all, water-inrush is more likely to occur in areas with tectonic faults (Fig. 5), which proves that geological structure has a major influence on water-inrush in the coal seam floor; water-inrush always occurred in zones with lots of faults. These areas have aquifers that contain a lot of water, high aquifer water pressure, and/or a thin and weak aquiclude. This indicates that aquifers and aquicludes also affect the likelihood of water-inrush, as one would expect.

BN Model Results Validation

Given the frequency of inrush events and the fact that the 12-2 coal seam floor is not as well studied as one might like, validation was conducted based on sample dichotomy: that is, events were sorted into two portions, with part being used to train the BN and the smaller portion used to validate the trained BN.

First, the aspects of the training samples were entered into the network and the risk and probability of water-inrush was determined (Fig. 6a). The inferred water-inrush probabilities in the assessment units were all above 0.72; the maximum was 0.91. Only one assessment unit had a probability that was close to 0.50, which correlates the facts.

Next, to test the validity of the method, the properties of the validation sample were entered into the BN; the calculated probability of water-inrush in the assessment

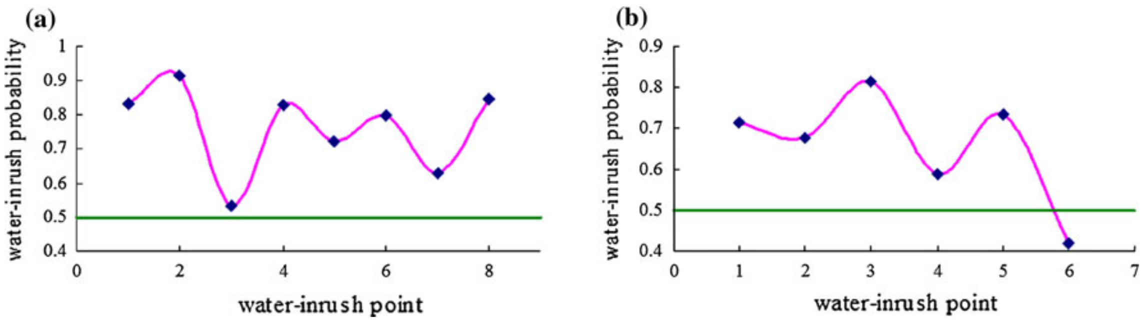


Fig. 6 Result validation probability of Bayesian network training data (**a**) and validation data (**b**)

units is illustrated in Fig. 6b. Although the validation probability was not as obvious as that of the training sample, and the calculated probability in several water-inrush assessment units was less than 0.5, most still had a validation probability between 0.58 and 0.81, which correlate the facts. Overall, the accuracy of the method was 83.4 % (validation probability of five-sixths of water inrush points is greater than 0.5) (Fig. 6b), which indicates that the BN method is both feasible and worthwhile.

Conclusions

BN + GIS proved to be a very valid method for water-inrush assessment in the Donghuatuo coal mine. It provided a better insight into past events and is now being used to predict the probability of future water-inrush events. Water-inrush hazard assessments based on GIS-BN will yield more precise predictions. In a validation test of 14 documented water inrush events, the accuracy of the fit of the model data is 83.4 %, and the probability of making an incorrect prediction is less than 0.5, which means that using this method could significantly enhance coal production at the mine. The developed BN-GIS technique for predicting water inrush events can be used in other areas.

Acknowledgments We gratefully acknowledge the financial support provided by the Fundamental Research Funds for the Central Universities (2009QD03), Beijing Natural Science Foundation (8122037), and China's National Natural Science Foundation (41172213).

References

Brenda MC (2001) Belief networks for engineering applications. Int J Technol Manage 21(3):257–270

Gudrun S, Gerhard S, Pietro P (2004) Bayesian reasoning on qualitative descriptions from images and speech. Image Vis Comput 18(2):155–172

Jozef G (2001) Financial analysis using Bayesian Networks. Appl Stoch Model Bus 17(1):57–67

Rodrigues MA, Liu Y, Bottaci L (2000) Learning and diagnosis in manufacturing processes through an executable Bayesian network. Lect Notes Compute Sc 1821:390–395

Sanguesa R, Burrell P (2000) Application of Bayesian Network learning methods to waste water treatment plants. Appl Intell 13(1):19–40

Wu Q, Jiang Z, Sun D (2000) Condition of water burst and dynamic forecast of inflow in the Donghuantuo mine. Coal Geol Explor 6:32–35 [in Chinese]

Wu Q, Wang M, Wu X (2004) Investigations of groundwater bursting into coal mine seam floors from fault zones. Int J Rock Mech Min 41(4):557–571 [in Chinese]

Zhong Y, Shen M (2000) Mining experience of the Donghuantuo coal mine. China Min Mag 9(3):1–4 [in Chinese]