The research presented in this paper is basically focused on two objectives. Firstly, the selection of parameters affecting the water supply network damage. The causes of failures were selected from a population of tens of breakdown cases and then classified in view of their importance. Secondly, attention was paid to the selection of the most suitable linguistic model which could be commonly used for selecting factors which generate failures. Finally a Mamdani-based model could be worked out as a system possessing best generalization qualities. This model can create bases for an adaptative decision system which can show the type of water supply-sewage network, depending on continuous surface strains due to the mining activity.

**Keywords:** mining damage, mining terrain, water supply, subsidence, fuzzy cluster analysis


**Słowa kluczowe:** szkoda górnicza, teren górniczy, sieć wodociągowa, ciągłe deformacje powierzchni terenu, rozmia klasteryzacja

* AGH UNIVERSITY SCIENCE AND TECHNOLOGY, AL. A. MICKIEWICZA 30, 30-059 KRAKOW, POLAND
1. Introduction

The water supply networks undergo failures, which frequently occur in areas subjected to movements and strains caused by underground construction, mining activity, rock mass tremors, earthquakes, etc. One of the applied prevention measures is usually replacement of old fragments of the network. Such prophylaxis, however, is not optimal for a number of reasons. Such operations are costly. It would be a better solution to identify damage-prone sections as in this way the number of failures and related social costs of such breakdowns could be minimized. It should be born in mind that most of water supply network failures end up in cutting off water for a large population of customers causing hazard for health and lowering the hygiene level. The analysis of water network damage allows for the identification of most frequent factors. However their evaluation and selection creates problems due to the highly diversified structure of water networks. Particular sections of the networks usually vary in age, pipe diameter and material from which they have been made. Accordingly, they have also different strength to stresses and to corrosion.

The stresses can be also analyzed using numerical calculations performed for a given network section (Cao et al., 2010; Gao, 2013; Wang et al., 2007). This approach is applied when the evaluated endangered water network section is especially important (e.g. master supply network), and which will be subject to the tunnel construction or downhole deposit extraction activities (Olajossy & Zajda, 2002; Stryczek & Wiśniowski, 2004). Most frequently the evaluation of technical state of the entire water supply network (if well documented) can be treated as damage risk analysis. In USA, when tremor areas are involved, such analyses are made with the use of, e.g. GIS (O’Rourke & Torpak, 1997). The analysis of sufficiently numerous samples does not prove strict correlation for this solution. There are methods lying in point evaluation of factors generating various breakdowns (Kliszczewicz et al., 1997), though they require highly professional experts who with their experience would select and analyze the factors. Accordingly, at present we do not have any method of evaluating damage risk for the entire water network system which would be sufficiently efficient.

The objectives of this paper are twofold. Firstly, attention has been paid to the selection of factors affecting network breakdowns. The causes of failures were selected from a population of tens of breakdown cases and then classified in view of their importance. Secondly, attention was paid to the selection of the most suitable linguistic model which could be commonly used for selecting factors which generate failures. Finally a Mamdani-based model could be worked out as a system possessing best generalization qualities. This model should create bases for an adaptative decision system which can show the type of water supply-sewage network, depending on continuous surface deformations due to the mining activity.

Following the realization of these two goals the authors worked out a model for evaluating the damage risk of water networks. This model will be based on a linguistic model selected for this purpose. The future research works will lie in determining boundary conditions for the selected parameters and testing of the obtained solution. In this sense the paper is a very important theoretical introduction and basis for the created model.

2. Engineering application of fuzzy models

Recently the fuzzy logic models started to be commonly used for solving problems where measurable quantitative factors are evaluated jointly with descriptive factors, the latter usually
assessed with expert methods. The Zadeh fuzzy set theory became popular in many scientific disciplines (Zadeh, 1965). The engineering disciplines, where reasoning is based on quantitative-characteristic variables, artificial intelligence tools are employed. The number of research and studies based on artificial intelligence has definitely increased over the last decade (Lv Yaqiong et al., 2010). Attempts at applying fuzzy modelling were undertaken in nearly all engineering disciplines. Large-scale applications of fuzzy modelling are commonly used for, e.g. solving problems related to the estimation of such environmental hazards as soil degradation (Riedler & Jandl, 2002) or environmental degradation (Feoli et al., 2002). Moreover, these tools are used for controlling environmental systems, e.g. sewage treatment plants (Jeng-Chung Chen & Ni-Bin Chang, 2007) or operation of fossil power plants (Arroyo-Figueroa et al., 1998, 2000). The fuzzy logic principles are also applicable while evaluating estate houses and engineering constructions for potential hazard (Hao-Tien Liu & Yieh-lin Tsai, 2012; Malinowska, 2011; Rusek, 2009). In water network problems they are mainly used for evaluating damage hazard (Bonvicini et al., 1998; Dong Yuhua & Yu Datao, 2005; Esayed, 2009; Han & Weng, 2011; Markowski & Mannan, 2009; Shahriar et al., 2012; Xingquan Liu et al., 2011). The results of over 30 year research have proved that in many engineering situations application of fuzzy models enables integrate and inference form ambiguity information.

3. Principles of fuzzy reasoning

The basic concept in fuzzy reasoning lies in defining a fragment of elements from the domain of variables over which it has been described (Dubois and Prade, 1980; Łęski, 2008; Piegat, 2003; Zadeh, 1965, 1973) (1).

\[ A = \left\{ (x, \mu_A(x)) \right\}, \forall x \in X \]  

where: \(\mu_A(x): X \rightarrow [0,1]\) — a membership function of a fuzzy set \(A\).

In the system approach, variables which describe a given effect are grouped, i.e. fuzzy sets are superimposed over their total extent. By linking fuzzy sets theory and incorporating it to generalized fuzzy logic principles one can build complex fuzzy reasoning systems. Generally the fuzzy reasoning systems consist of a number of concurrently defined rules of IF...THEN... type. Each rule is a result of operation of some piece of information enclosed in a fuzzy implication.

These systems act in one-dimensional space of input variables, therefore each construction rule is based on a number of premises. This signifies that the presentation of a given rule in the case of \(n\)-dimensional space of input data closes the \(n+1\)-dimensional cluster in the system reasoning space.

The schematic concept of MISO-type fuzzy reasoning system is presented in Fig. 1.

Thus presented schematic concept of fuzzy reasoning system consists of four major parts. The first one is the fuzzification block, where each input variable is fuzzified. This process lies in giving particular variables fuzzy (blunt) categories over their entire extent and expressed with fuzzy sets.

Another part of the system is the rule base. The rule base is the most important part of expert and adaptive fuzzy reasoning systems. It is just there where the information nucleus about the relation between categorized input variables and system exit is stored. As already noted, each rule
is basically formulated by a fuzzy implication operator, therefore the rule base can be presented in the form of locally defined clusters, which store information about the system operation. Each cluster, depending on the degree to which the premise has been activated, (‘ignition’ level expressed by the value of membership function of a given fuzzy set (Dubois and Prade, 1980; Łęski, 2008; Piegat, 2003; Zadeh, 1965) activates the corresponding conclusions (exit from the system) – Fig. 2.

Fig. 1. Schematic of structure and operation of MISO system (Piegat, 2003)

Fig. 2. Reasoning for a given system rule
Therefore, assuming arbitrarily $N$ input variables and one output variable, explained at the stage of reasoning, we obtain a certain rule in the following form (Łęski, 2008):

$$ R^{(i)} = \left\{ \bigwedge_{n=1}^{N} x_n \text{ is } A_n^{(i)} \Rightarrow y \text{ is } B^{(i)} \right\} $$

(2)

On the other hand, the generalized reasoning scheme based on information in the rule, assumes the following form:

$$ \left\{ \bigwedge_{n=1}^{N} x_n \text{ is } A_n^{(i)} \right\} \land \left\{ \bigwedge_{n=1}^{N} x_n \text{ is } A_n^{(i)} \Rightarrow y \text{ is } B^{(i)} \right\} \Rightarrow y \text{ is } B^{(i)} $$

(3)

Having assumed Mandani operator for the implication and Zadeh $t$-standard for conjunction, the operation of thus formulated reasoning scheme for the $i$-th system rule is presented in Fig. 3.

Fig. 3. Schematic of fuzzy reasoning operation for a given system rule

All system rules are involved in the reasoning process. As a consequence the number of the obtained activated input sets, being a conclusion of each of the rules, corresponds to the number of rules activated in the reasoning process. In this way one has to establish the ultimate fuzzy set in the input variable space, which is formed after all activated conclusions in the rule base are cumulated. This operation is performed in the aggregation block of the reasoning system.

$$ \left[ x \text{ is } A' \land (x \text{ is } A \Rightarrow y \text{ is } B) \right] \Rightarrow y \text{ is } B' \Leftrightarrow \mu_B'(y) = \sup_{x \in X} \left[ \mu_{A'}(x) \circ T_I \left[ \mu_A(x), \mu_B(y) \right] \right] $$

(4)

where:

- $\sup(\cdot)$ — upper limit of the set,
- $\circ T_I$ — fuzzy conjunction operator expressed by appropriate $t$-standard,
- $I(\cdot, \cdot)$ — fuzzy implication operator,

A graphical interpretation of this reasoning scheme is given in Fig. 3.

4. Data description

The research is based on a water network, 62 km long, running through areas staying under the mining impact. All the investigated area was subject to a longwall coal extraction with roof caving. The thickness of the exploited lots ranged between 2 m to 4.5 m. The coal extraction was
realized successfully in the Upper Carboniferous strata. Relatively thick coal beds were mined by layers at levels 500 m to 710 m. The production in the years 2003 to 2012 resulted in a subsidence of the surface reaching up to 3.2 m, horizontal strains locally over 9 mm/m. The analyzed water network remained under the influence of intense surface strains for over ten years (Fig. 4).

Fig. 4. Distribution of water network in areas subjected to continuous deformations of the surface (maximum-in-time horizontal strains)
The basic characteristic of water network is as follows. The water pipeline was principally made of steel (about 80% of cases), asbestos (15%), PE (3%) and cast iron (2%). Its average diameter was 40-200 mm. The pipes were laid at about 1.1 m of depth (73%) and the pressure in the network ranged from 0.38 MPa to 0.50 MPa. The ground in which the network was disposed was mostly wet (68% of cases) or hydrated (28%) and highly hydrated (4%). All the characteristic data were stored in the database. The basic variables referring to the construction of the water network were: material used in the pipelines network, pressure inside, depth of deposition and diameter. The base also contains data about failures and their localization. The set of variables characterizing the utilities was broadened by indices describing strains in the mining area.

Fig. 5. Distribution of major and directional strains at the breakdown site

The impact of the following factors on a linear object was analyzed (Fig. 5):
- directional strain $\varepsilon_\alpha [\text{mm/m}]$ and its rate $V\varepsilon_\alpha [(\text{mm/m})/\text{month}]$ acting along the linear object axis,
- major strain $\varepsilon_{g1}$ [mm/m], $\varepsilon_{g2}$ [mm/m] at the breakdown site,
- maximum strain $\varepsilon_{\text{max}}$ [mm/m], which occurred at the breakdown site ($\varepsilon_{\text{max}} = |\varepsilon_{g1}|$ or $\varepsilon_{\text{max}} = |\varepsilon_{g2}|$) and its rate $V\varepsilon_{\text{max}} [(\text{mm/m})/\text{month}]$,
- amplitudes of horizontal directional strain $A_{\varepsilon_\alpha}$ [mm/m],
- amplitudes of maximum strain $A_{\varepsilon_{\text{max}}}$ [mm/m].

The frequency with which strains were monitored in this area was low and the density and distribution of observation points insufficient to determine the spatial distribution of extreme strain indices Therefore the strain indices and rates of their growth were determined with the use of Knothe’s influence function. The results of observations allowed for scaling out the prediction model and adjusting it to the local conditions. Having accounted for local parameters one could very accurately estimate continuous strains.
These indices were used in various combinations, thus affecting particular calculation sets for the analyses. Moreover, thanks to the qualitative analysis of data performed at the very beginning, the set of input data could be reduced to the ones presented in table 1.

The categories indicating the type of damage were the input (explanatory) variable. On the whole 139 of damage cases were collected in the database and they were grouped according to eight damage-level categories, as presented in tab. 1.

<table>
<thead>
<tr>
<th>Description of damage in database</th>
<th>Generalized name of damage used for analyses</th>
<th>Category number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrosion</td>
<td>Corrosion</td>
<td>I</td>
</tr>
<tr>
<td>Broken welding</td>
<td>Broken welding</td>
<td>II</td>
</tr>
<tr>
<td>Broken asbestos pipeline</td>
<td>Broken asbestos</td>
<td>III</td>
</tr>
<tr>
<td>Breakings, tearings of tee, U-bend, etc.</td>
<td>Broken connection</td>
<td>IV</td>
</tr>
<tr>
<td>Lack of tightness</td>
<td>Untight</td>
<td>V</td>
</tr>
<tr>
<td>Petering out, tearing away</td>
<td>Tearing</td>
<td>VI</td>
</tr>
<tr>
<td>Bending, breaking, crashing</td>
<td>Bending</td>
<td>VII</td>
</tr>
<tr>
<td>Damaged asbestos collars</td>
<td>Collar damage</td>
<td>VIII</td>
</tr>
</tbody>
</table>

Apart from their main objective, i.e. finding out a decision system with best properties as far as adjusting and generalization are concerned, the analyses were also expected to help one select the most efficient set of input variables (in the context of representation, i.e. explain the variability in original data by the model) and show the direction of potential damage categorization depending on its burdensome effect.

5. Estimation of the parameters affecting the water pipelines on the mining terrains with a use of fuzzy clustering method

The analyses concentrated on the construction and consequently finding out the best fuzzy reasoning system for preselected combinations of input variables.

Two types of fuzzy reasoning systems were performed within the research, i.e. Takagi-Sugeno-Kanga system and Mamdani system.

The main difference in this type of systems concerns aggregation and the defuzzification method. As presented in fig. 1 the Takagi-Sugeno-Kanga (TSK) system is a typical regression model. The estimated value of such a model is acquired in line with the formula (Łęski, 2008; Osowski, 2006)

\[
y(x_1, x_2 \ldots x_N) = \frac{\sum_{i=1}^{K} \mu_i(p_i) \cdot p_i}{\sum_{i=1}^{K} \mu_i(p_i)}
\]
where:

- $N$ — number of input variables,
- $K$ — number of all system rules,
- $p_i$ — conclusion carrier for $i$-th rule,
- $\mu_i(p_i)$ — state of activation of $i$-th rule at the presentation of established variables $x_1, x_2...x_N$ at the system entry.

The parameters of such a system were adapted with the use of two methods (error back propagation method and hybrid method).

Additionally, the group of analyzed cases was checked out for the preliminarily initiated parameters, and which would later undergo adaptation. Three initiation methods for the optimization were assumed: Grid Partition method, Subtractive Clustering method and Fuzzy C-Means FCM method.

In the Mamdani system the output variable is described with a fuzzy set undergoing aggregation and sharpening, analogously as described in Chapter 2. In this case the distribution and geometrical characteristic of membership function of fuzzy sets both in the input an output space were conducted with the Fuzzy C-Means FCM method.

Apart from the main objective, i.e. finding out a decision system with the best adjusting and generalization properties, the analyses were also oriented to selecting the most efficient set of input variables in the context of representation, i.e. explain the variability in original data by the model) and show the direction of potential damage categorization depending on its burdensome effect. Moreover, the found model is further planned to be used for evaluating the significance of particular variables, mainly the mining ones, through the sensitivity analysis method.

Having assumed the original purposes of the research and accounting for the fact that:
- certain variables are uncertain in character,
- the burdensome character will be established linguistically,

the authors decided that at this stage of research the most efficient will be tools making use of fuzzy notation and fuzzy reasoning processes.

Making still another assumption that the present research is of preliminary character at the stage of which the evaluation and data mining problems appear, the fuzzy reasoning systems, self-developing in the adaptation process, were applied.

As a consequence, there were conducted investigations consisting of four parallel but independent stages. Each stage (Fig. 6) consisted in making a MISO-type fuzzy reasoning system for seven combinations of input variables (described in table 1). The first three stages were devoted to construing an adaptive Sugeno-type fuzzy reasoning system (learning) (Fig. 7). At the fourth stage the Mamdani system was used along with the clusterization-based approach (Fig. 8). The data on the assumed methods as well as the employed start-up and learning algorithms were given in table 4.

As a consequence of the assumed four-stage flowchart, a total of 28 models were obtained. Prior to building the models, each input variable and output variable was normalized to the interval $<0,1>$. After a few preliminary simulations the accepted number of epochs for teaching the adaptive system equalled to $e_p = 200$. Moreover, for the adaptive systems both for input and output variables, the assumed number of fuzzy sets was $l_R = 3$. At the fourth stage when the Mamdani fuzzy reasoning system was construsted the assumed number of clusters was $l_K = 4$. 
Fig. 6. Flowchart of the research algorithm

Fig. 7. Structure and operation of Sugeno fuzzy reasoning system
Fig. 8. Structure and operation of Mamdani fuzzy reasoning system

The results of simulations are presented in tables 4 to 7, where the MSE (Mean Square Error) was assumed as a measure of fitting the model to paradigm data. The obtained models were verified on a training set (96 learners) and test sets (41).

<table>
<thead>
<tr>
<th>Set of input variables</th>
<th>MSE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Combination no.</strong></td>
<td><strong>εg1</strong></td>
<td><strong>εg2</strong></td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------</td>
<td>---------------</td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 4
TABLE 5

Results obtained for a system created at the second stage

<table>
<thead>
<tr>
<th>Combination no.</th>
<th>$\varepsilon_g^1$</th>
<th>$\varepsilon_g^2$</th>
<th>$\varepsilon_{\text{aekstr}}$</th>
<th>$V_{\text{cmax}}$</th>
<th>$V_{\text{ca}}$</th>
<th>$A_{\text{cmax}}$</th>
<th>$A_{\text{ca}}$</th>
<th>Aver. Mat.</th>
<th>Training set</th>
<th>Test set</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.000316</td>
<td>0.006772</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.000414</td>
<td>0.001869</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td>0.000542</td>
<td>0.002309</td>
</tr>
<tr>
<td>4</td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.000188</td>
<td>0.003176</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td>$\sqrt{}$</td>
<td></td>
<td>0.000249</td>
<td>0.005132</td>
</tr>
<tr>
<td>6</td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td>0.0004</td>
<td>0.001837</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td>$\sqrt{}$</td>
<td>0.000522</td>
<td>0.00318</td>
</tr>
</tbody>
</table>

TABLE 6

Results obtained for a system created at the third stage

<table>
<thead>
<tr>
<th>Combination no.</th>
<th>$\varepsilon_g^1$</th>
<th>$\varepsilon_g^2$</th>
<th>$\varepsilon_{\text{aekstr}}$</th>
<th>$V_{\text{cmax}}$</th>
<th>$V_{\text{ca}}$</th>
<th>$A_{\text{cmax}}$</th>
<th>$A_{\text{ca}}$</th>
<th>Aver. Mat.</th>
<th>Training set</th>
<th>Test set</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.000542</td>
<td>0.00196</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.000557</td>
<td>0.001008</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td>0.000588</td>
<td>0.000957</td>
</tr>
<tr>
<td>4</td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td>0.000411</td>
<td>0.001215</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td>$\sqrt{}$</td>
<td></td>
<td>0.000447</td>
<td>0.008896</td>
</tr>
<tr>
<td>6</td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td>0.000564</td>
<td>0.001631</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td>$\sqrt{}$</td>
<td>0.000572</td>
<td>0.001598</td>
</tr>
</tbody>
</table>

TABLE 7

Results obtained for a system created at the fourth stage

<table>
<thead>
<tr>
<th>Combination no.</th>
<th>$\varepsilon_g^1$</th>
<th>$\varepsilon_g^2$</th>
<th>$\varepsilon_{\text{aekstr}}$</th>
<th>$V_{\text{cmax}}$</th>
<th>$V_{\text{ca}}$</th>
<th>$A_{\text{cmax}}$</th>
<th>$A_{\text{ca}}$</th>
<th>Aver. Mat.</th>
<th>Training set</th>
<th>Test set</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.001006</td>
<td>0.000901</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td>0.00086</td>
<td>0.001002</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td>0.00094</td>
<td>0.000705</td>
</tr>
<tr>
<td>4</td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td>$\sqrt{}$</td>
<td>$\sqrt{}$</td>
<td></td>
<td>0.000909</td>
<td>0.001133</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td>$\sqrt{}$</td>
<td></td>
<td>0.000991</td>
<td>0.000875</td>
</tr>
<tr>
<td>6</td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td>$\mathbf{0.000929}$</td>
<td>$\mathbf{0.001028}$</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>$\sqrt{}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\sqrt{}$</td>
<td>$\sqrt{}$</td>
<td>0.00089</td>
<td>0.001236</td>
</tr>
</tbody>
</table>

6. Results and discussion

The obtained results were analyzed in two stages. First, attention was paid to finding out a system, which is best as far as adaptation and
generalization are concerned. An auxiliary measure $Sel_{sys} = \frac{MSE_{TR}}{MSE_{TS}}$ was used here. With thus defined measure one can indicate a model whose adaptive properties in training and test sets are comparable, i.e. has good generalization qualities.

At the second stage the set of generated systems was inspected with the objective to find out a combination of input variables which is most influential as far as explaining of the variability of information in data goes. In this case the MSE error for the training set was the criterion.

The result of the first stage of analysis showed the Mamdani fuzzy reasoning system built on input variables (combination no. 6 in table 7.)

As a result of the analysis of procedure conducted at the second stage the Sugeno fuzzy reasoning system was indicated for a combination of input variables no. 5 in table 5. This system, however, does not reveal good generalization qualities therefore is treated as supplementary, to be used at the second stage. The graphical distribution of errors for the training, test and the assumed measure $Sel_{sys}$ is presented in Fig. 9.
The generalization properties of all created systems are graphically presented in Fig. 9. In the first row we have the distribution of errors obtained for the training sets, in the second line for the test sets and in the last row we have a difference between the errors in the training and test sets. The highest generalization quality was observed for the fourth stage model, i.e. Mamdani system, and the lowest generalization for the all of the first stage models for all combinations of input variables. The bigger is the difference between the errors in the training and test sets, the lower is the generalization capacity. That means the reduction of the models ability is to act correctly when providing new observations to the set of data.

In the course of the research the Mamdani model from the fourth stage was selected as possessing best generalization qualities. Among factors having a decisive influence on the damage level of a water network ultimately are: pipeline diameter and material from it has been made. The deciding risk factors are $A_{\text{max}}$, $e_{\text{aeko}}$, $V_{\text{ea}}$. Based on selected parameters novel model for pipeline hazard estimation will be created.

**Acknowledgements**

The research reporter in this paper has been supported by a grant from the National Science Centre No. 2011/01/D/ST10/06958.

**References**


