Determination of the Probability of a Geap Pump Fault Using the Bayes Network and Netica Software

Jovana Perić*, Mirko Djapić[1], Tanja Miodragović[1], Stefan Pajović[1]
[1]The Faculty of Mehanical and Civil Engineering in Kraljevo, University in Kragujevać, Kraljevo (Serbia)

The literature dealing with risk analysis and finding an adequate solution in the mechanical engineering sector shows that the use of the Bayesian network as one of the methods for risk assessment in this context is well known. Therefore, the authors of this paper use the Bayesian network, while accordingly, for the purpose of simpler work with Bayesian networks, special software has been developed, one of them being Netica, as the most commonly used method for determining an adequate solution, and in this case determining the probability of malfunction. gear pumps if the probabilities of variables that directly affect the correctness of the pump are known. The procedure for determining the probability of malfunction of the gear pump in the specific example shown in this paper is obtained first by designing the Bayesian network, and then by creating the same network in Netica software, where pre-defined nodes (variables) to which assigned values are entered in program, after which the arrows show the dependences of the given variables, and then enter the conditional probabilities of these variables and form a network on the basis of which the probability of malfunction of the gear pump is determined.

Keywords: Bayesian networks, Netica software, conditional probabilities, variables, adequate solution, faulty operation of the steam pump.

1. INTRODUCTION

Bayesian networks represent directed, acyclic graphs where nodes play the role of variables and branches represent their interconnectedness. Linking variables is a basic problem, because it is necessary to clearly determine the relationship between each of the two variables that are observed. This is followed by the assignment of conditional probabilities, which also requires a lot of knowledge in the field that is the subject of the examination. From the model thus formed, it is possible to draw a conclusion about each of the variables, taking into account the influences that other variables have on it. Also, the Bayesian network provides the ability to model the interdependence of a large number of variables. When it comes to the relationship of two dependent variables, inference is simple. In this case, the answer is given by Bayes' theorem. However, a much more complicated case arises when there are several dependent variables for which the required conditional probabilities due to complexity cannot be calculated using Bayes' theorem but using Bayesian networks [1], [2].

As an adequate solution is one of the key segments in the machine system, this paper shows how to obtain a probability that shows the percentage of non-correctness of the gear pump operation if the probabilities of variables that directly affect the correct operation of the pump are known. One of the variables that directly affects the correct operation of the gear pump, which is performed using Bayesian networks and Netica software, is oil overheating.

Overheating of the oil as one of the consequences can be due to excessive friction or elevated ambient temperature or insufficient amount of oil. Furthermore, the presence or absence of excessive friction can lead to shock loads or pump overload or excessive speed. While elevated ambient temperature or insufficient oil or shock loads or pump overload or excessive speed has a direct impact on the correct operation of the gear pump. From the given example, it is easy to conclude that each of the variables affects another variable or depends on some variable. This would further mean that the variables can be chained in a certain way, which is the first step in forming a Bayesian network. For variables that are directly related, it is necessary to determine the conditional probabilities, this is done on the basis of previous experience and theoretical knowledge of the person making the conclusion. So the first step in determining the probability that shows the percentage of non-operation of the gear pump is the process of designing a Bayesian network where each variable is defined in detail, and then creating a Bayesian network in Netica software in which predefined variables are entered into the program and formed a network.

2. THE PROCEDURE FOR DESIGNING A BAYES NETWORK

In order to design a Bayesian network, it is first necessary to define the following:

• nodes in the network (variables to be solved),

• possible outcome of each node (values received by each of the variables),

• branches in the network (interdependence of variables),

• the probability distribution of each of the variables depending on the probability distribution of their parents in the network.

The next step is to assign probabilities. They represent the initial probabilities assigned to the variables,

*Corresponding author: Jovana Perić, The Faculty of Mechanical and Civil Engineering in Kraljevo of the University in Kragujevac, Dositejeva 19, 36000 Kraljevo, peric.j@mfkv.kg.ac.rs

and they are arrived at on the basis of some previous research. When it comes to the Bayesian network, probability would imply anticipating the outcome of each of the variables when no evidence is known for any node in the network.

A variable that is not affected by any other variable in the network is called a parent or root. They are placed first in the network. Then the parents connect the branches with the variables on which they have a direct influence and they are called children. This procedure continues until the final variables are reached, those that do not have children, ie they do not affect any of the variables. A similar procedure is used when assigning probabilities. It is first necessary to assign probabilities to the variables that represent the root in the network.

This is the probability in the observed example of a malfunction of the gear pump, overheating of the oil. Then, for each variable that has parents, a conditional probability table is formed. So for each variable it is necessary to take into account all combinations of probabilities of its parents. The number of parents of a variable determines the dimensionality of the table.

Figure 1 shows the analysis of the interdependence of nodes (variables) in the network. The presence of excessive friction, elevated ambient temperature and insufficient oil depends on whether the oil is overheated or not. So, this variable overheating oil has three children. Furthermore, impact loads, pump overload, excessive speed depend on whether there is too much friction or not. Well, this variable also has three children. While the variable malfunction of the gear pump, which is also the question that arises in order to reach an adequate solution to the problem, has five parents, depends on elevated ambient temperature, insufficient oil, shock loads, pump overload and excessive speed. In order for this model to be complete, it is necessary to associate each variable with the corresponding conditional probability, Figure 2.

Each of the variables that has parents is accompanied by tables of conditional probabilities that will be taken into account by each of the outcomes of her parents. As the number of parents increases, so does the number of conditional probabilities that will be associated with the mentioned variable. For example, a node that has two parents will be associated with a total of 4 probabilities, because each parent has two states in which it can be. Thus, the total number of conditional probabilities joining the node, which has n parents, is 2^n. This procedure performs the design of the Bayesian network, which is further entered into the Netica program, which leads to a percentage probability of failure of the gear pump.

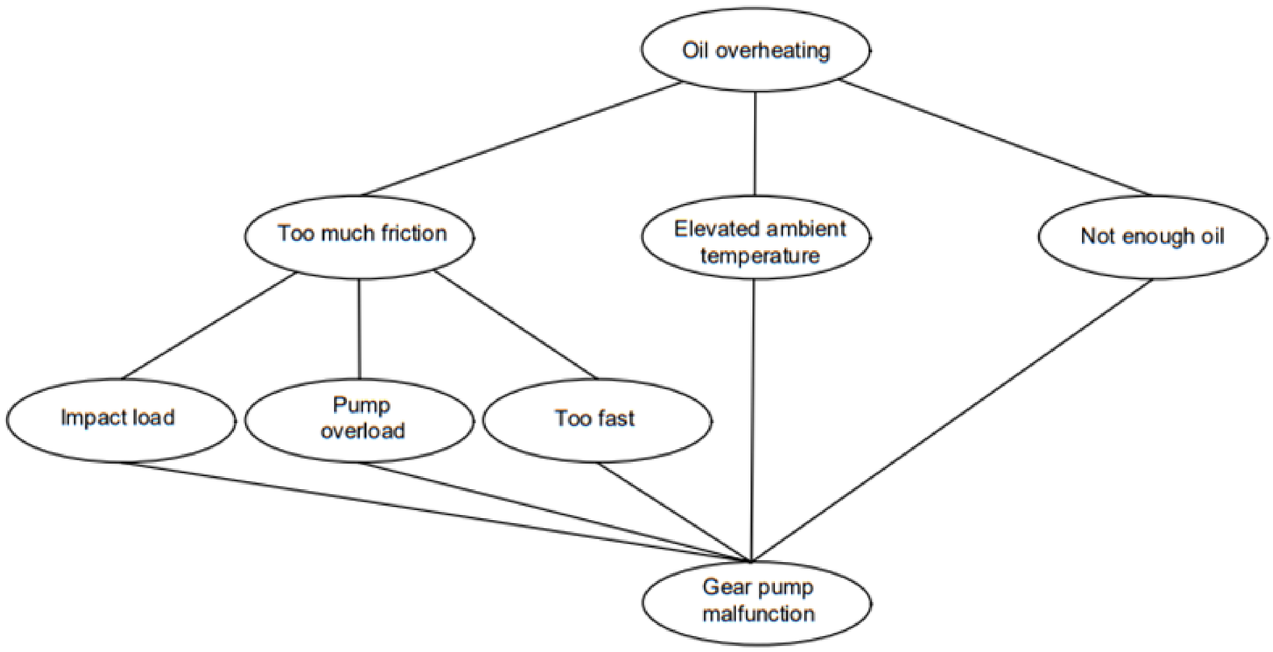

Figure 1: Bayesian network

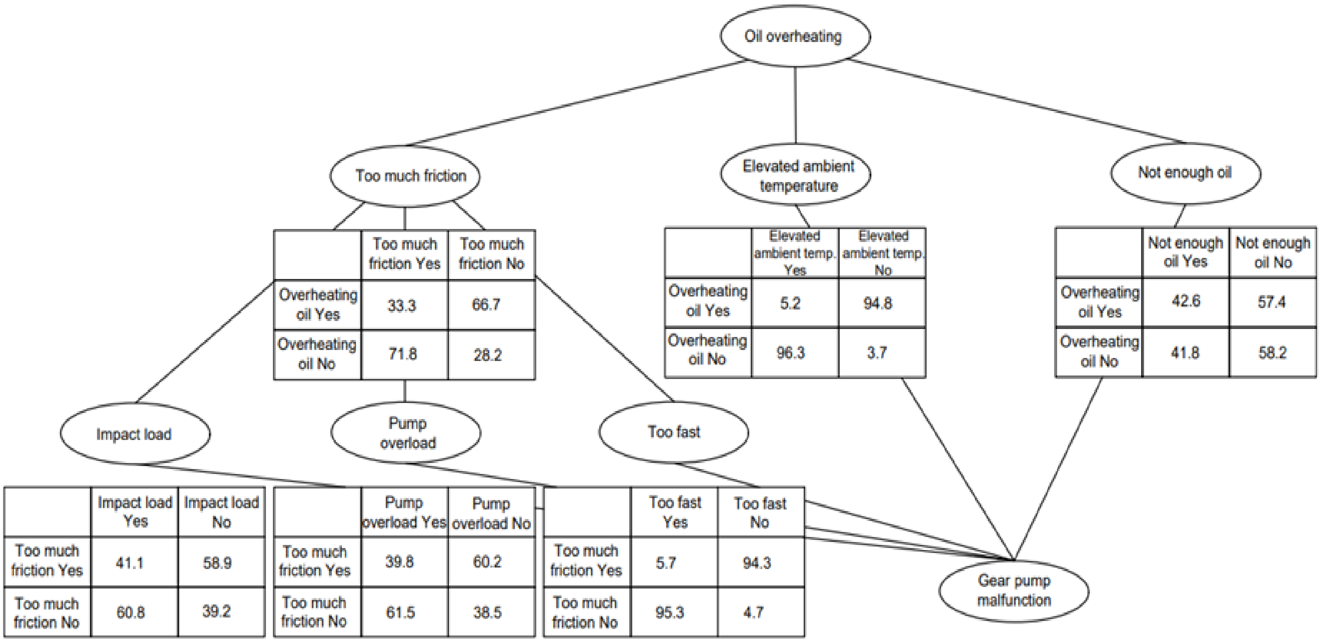

Figure 2: Bayesian network with conditional probabilities

3. CREATING A BAYES NETWORK USING NETICA SOFTWARE

Calculating probabilities in Bayesian networks that have a large number of nodes can be very complicated if the probabilities were calculated "on foot". In order to make working with Bayes' networks as simple as possible, several software tools have been developed, one of which is used in this work is Netica.

The principle of operation of this program is extremely simple. It is necessary to enter nodes (variables), values that variables can receive, conditional probabilities, more precisely everything that is defined by the formatting procedure. After that, the arrows show the dependence of the given variables and a network is formed. Figure 3 shows the appearance of the working window of the running Netica program.

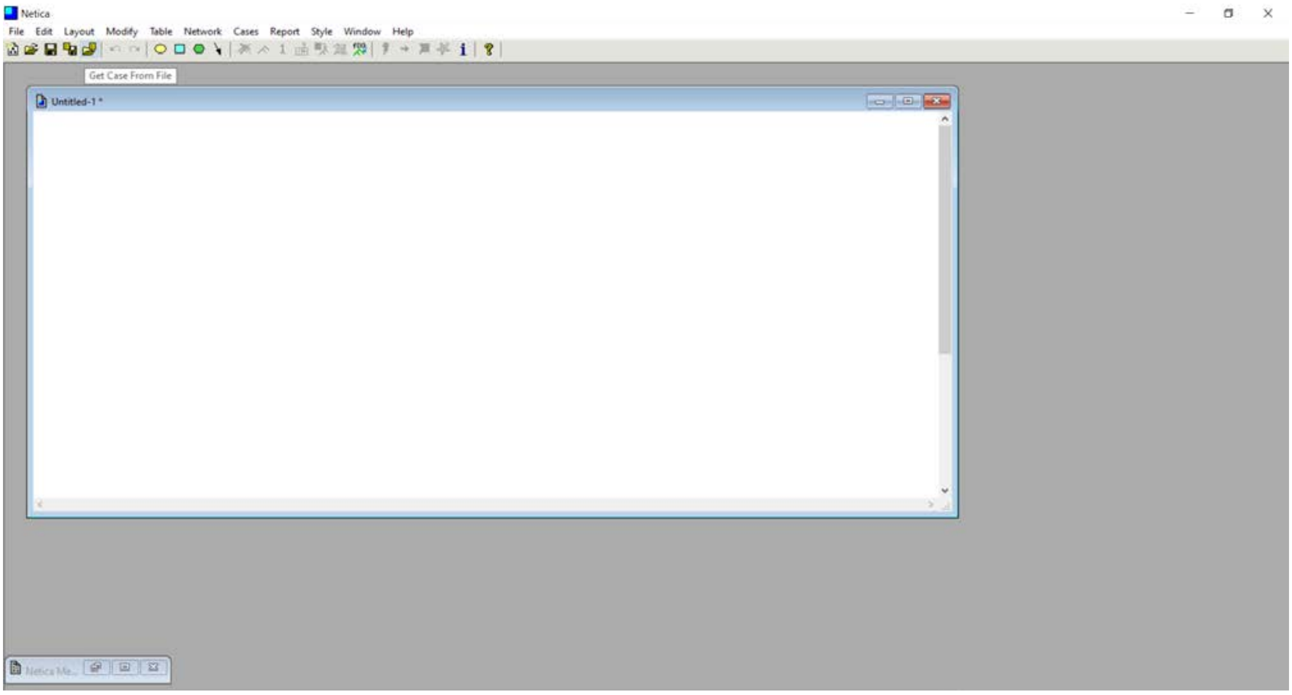

Figure 3: Appearance of the Netica program window

In the following text, it will be briefly explained how the Bayesian network is created with the help of Netika software, through the determination of the probability that shows the percentage of non-operation of the gear pump.

Creating a network in software begins with selecting defined nodes that represent variables. Netica software provides the ability to draw three types of nodes, arranged respectively in the form: circle, square and hexagon. These are, respectively: nodes of "nature", nodes of "decisions" and nodes of "values". Nature nodes are variables that can be influenced by the decision maker. They are used to represent empirical or calculated parameters and the probabilities of occurrence of different states. Decision nodes are management variables that can be directly implemented by the decision maker. These nodes represent a set of available management actions. They are closely related to value nodes, which are used to estimate the correct optimal decision on the network that will maximize the sum of the expected values of the value nodes. Next to the nodes, there is also an arrow that is used to establish the dependencies between the nodes. Relationships between nodes can be entered as individual probabilities, in the form of equations, or they can be downloaded from a data file.

In the concrete example of the network in Netica, the network will be drawn over the nodes of nature, and the relations will be entered as probabilities. When a sufficient number of nodes are displayed in the Netica software workspace, determined whether they are of discrete or continuous type, then the values that the variable can receive are entered, Figure 4, the next step is to connect the nodes to establish dependency and provide the ability of Netica to recognize who the parents of each of the variables are and thus enable easier entry of predetermined conditional probabilities, Figure 5. A further step is to enter the conditional probabilities that are known, Figure 6. Now, when all probabilities have been entered by clicking on the Compile Net option, the probabilities for all variables in the network are obtained Figure 7, after the check, based on previously defined and entered probabilities, the probability is determined, which shows the percentage of gear pump failure.

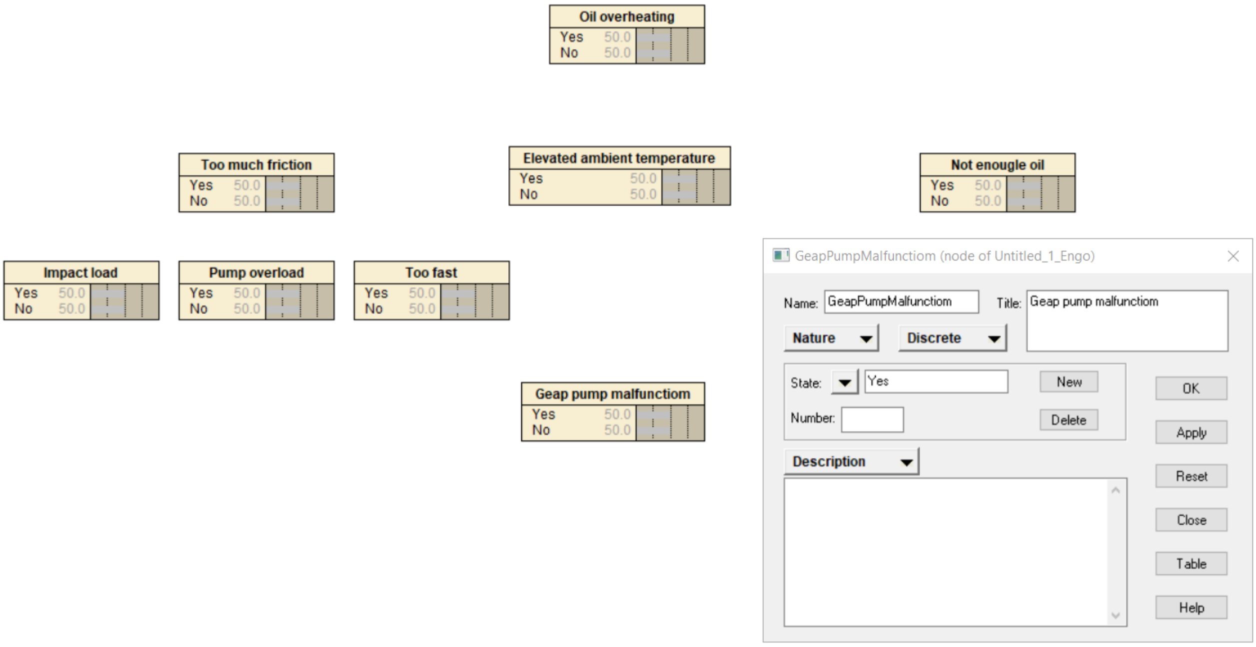

Figure 4: Node naming, input of node type and values that node receives

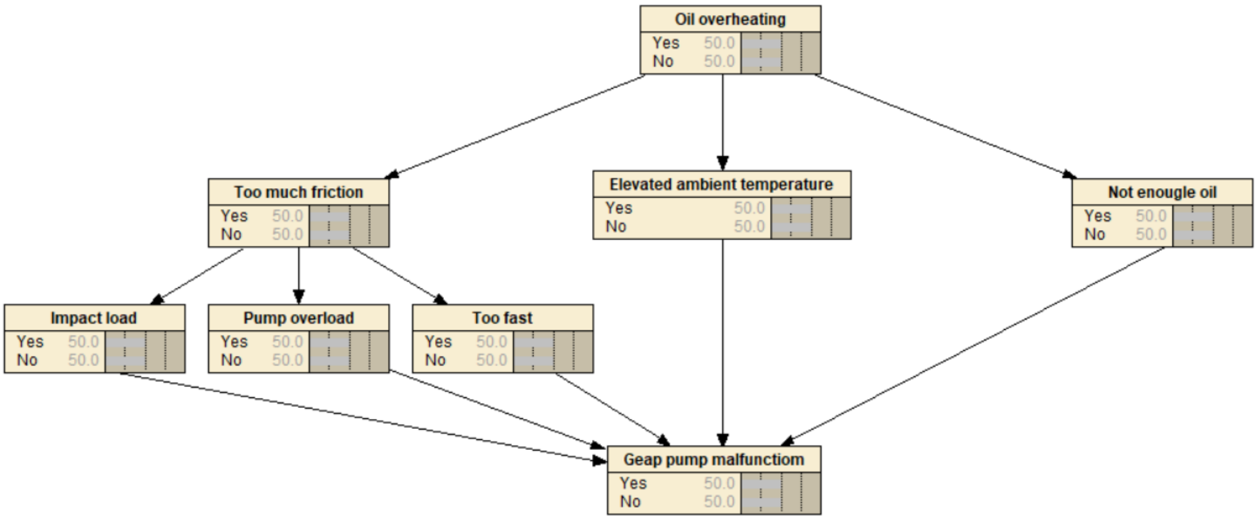

Figure 5: Connecting nodes

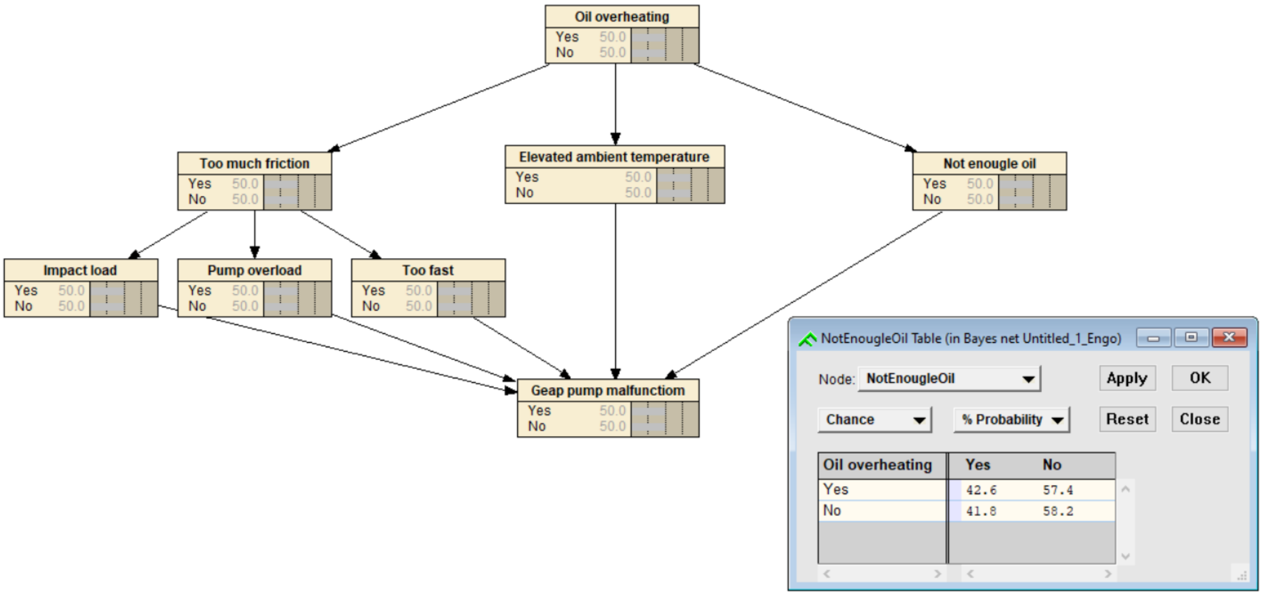

Figure 6: Entering conditional probabilities

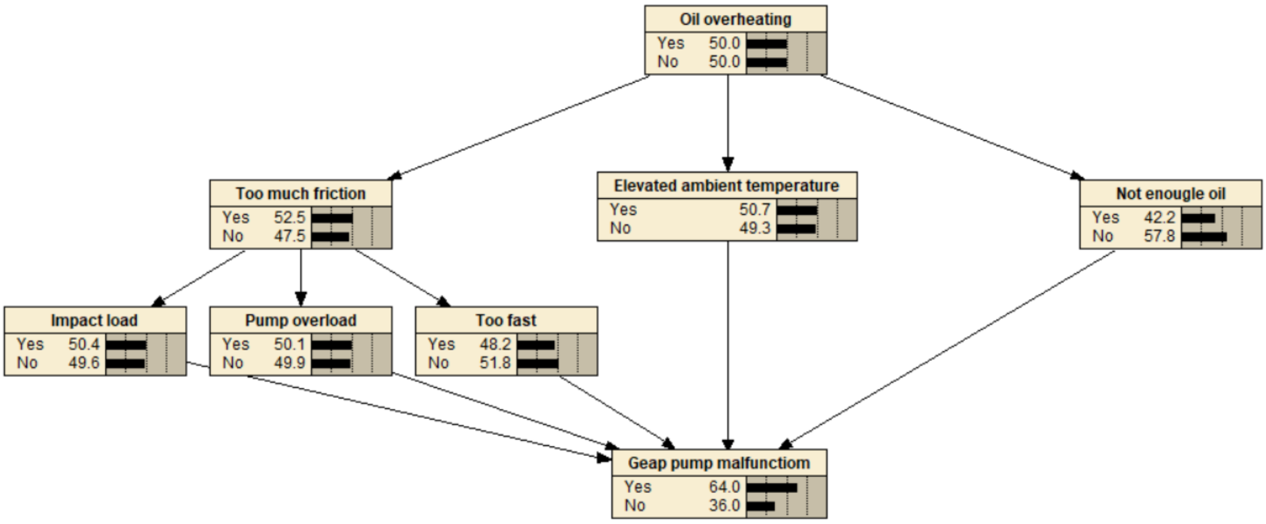

Figure 7: Assigned probabilities to all nodes

4. POSSIBILITY OF CHANGE OF PROBABILITY ON THE CORRESPONDING NODE (VARIABLE) IN THE FORMED NETWORK

In the finished Bayesian network, Netica offers the possibility of changing conditional probabilities, if there is a need for that, more precisely if there is evidence for a variable, ie it is known what value the given variable takes. In order to change the probability of the corresponding variable, it is necessary to set its probability to the appropriate one, after which the program will change the probabilities of other nodes that depend on that node (variable).

Modification within the given network for malfunction of the gear pump was performed on the node (variable) pump overload, where the problem observer set the probability for the selected node (variable) is 100, after which the program automatically adjusted the conditional probabilities of all other variables with to which it is related and in this way it is easy to calculate the percentage of malfunction of the gear pump when there is a change in the probability of the corresponding variable, Figure 8.

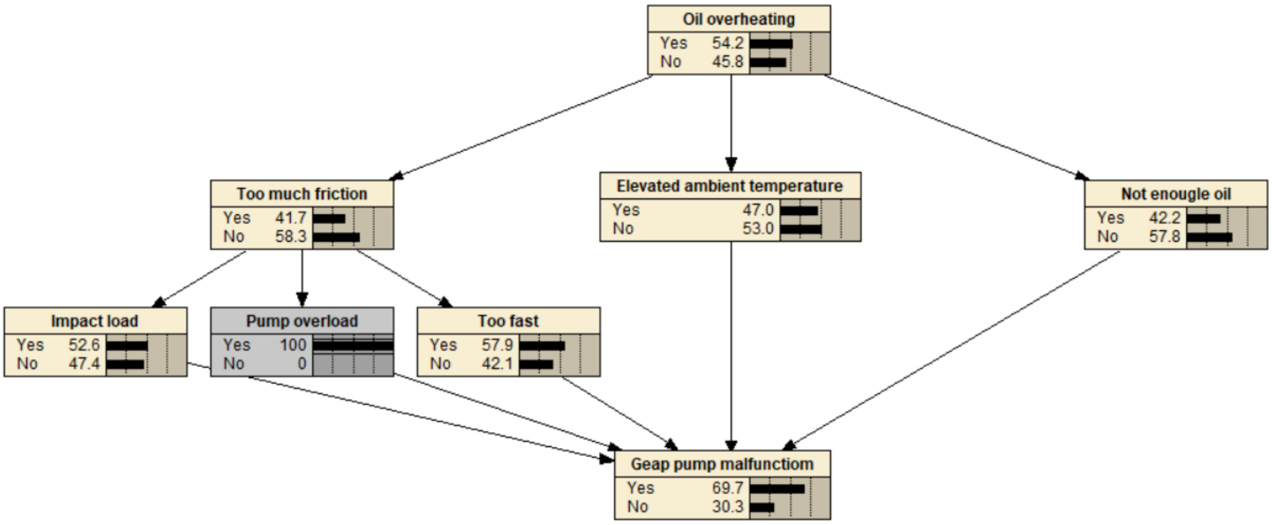

Figure 8: Probability change on the selected node

CONCLUSION

In the last few decades, new methods have been developed for mechanical engineering that enable faster and more efficient solving of problems that designers encounter when making decisions. Before solving any problem, the most important thing is to determine the exact cause of the problem. This involves a series of actions, observations and interpretations of the very cause of the problem that would lead to the presence of a suspected problem. In almost all areas of mechanical engineering, the exact cause of the problem can be solved by experience, intuition, knowledge or skill, but also by using programs that enable faster and more efficient problem solving. One such program is the Netica software, the principle of operation, design process and creation of the network is presented in this paper. The whole procedure performed to create a Bayesian network in the given example provides the possibility to determine the percentage values of gear malfunction if variables (causes) are known that directly affect the correct operation of the pump.

The paper also provides an opportunity for the problem observer to adjust the probability of a variable depending on his problem after the network is created, after which the modified Bayesian network can be easily adapted to the new data, no matter how many variables there are. This actually means that if one probability changes on one of the variables, it is possible to automatically adjust the conditional probabilities of all other variables to which it is related, which further provides the possibility of easily calculating the probability of malfunction of the gear pump.

ACKNOWLEDGEMENTS

This work is co-financed by the Ministry of Education, Science and Technological Development of the Republic of Serbia on the base of the contract whose record number is 451-03-9/2021-14/200108. The authors thank the Ministry of Education, Science and Technological Development of the Republic of Serbia for supporting this research.
.

REFERENCES

[1] R. Neapolitan, "Learning Bayesian Networks", Northeastern Illinois University, Chicago, (2000)

[2] Norman Fenton, Martin Neil, "Decision Analysis with Bayesian Networks", Second Edition, (2019)

[3] Scott Ferson, "Bayesian methods in risk assessment", (New York, USA)

[4] T. Aven, J.T. Kvaløy, "Implementing the Bayesian paradigm in risk analysis", Stavanger University College, P.O. Box 8002, 4068 Stavanger, Norway

[5] https://www.norsys.com/tutorials/netica/nt_toc_C.htm.