Seminar on Case Studies in Operations Research (Mat-2.4177)

Evidential Uncertainties in Reliability Assessment - Study of Non-Destructive Testing of Final Disposal Canisters

VTT, Posiva

Backlund, Ville-Pekka
Piironen, Juho
Tuovila, Henri
Wolf, Sebastian

Mid-term Report April 12th 2012

Aalto University
School of Science

1 Introduction

The original problem for this course work comes from the final repositories of nuclear waste, and the need to study the copper canisters in which the radioactive waste is intended to bury into bedrock. A crucial part of the durability of the canister is the quality of the welding of the lid of the canister. Due to the long half-life of nuclear waste and high production costs, the copper canisters can not be tested in practice. This means, that the safety of these canisters has to be proven in other ways. The solution is to use so called non-destructive testing (NDT) methods, which are applied to the weld of actual canisters containing nuclear waste. These tests give us evidence on possible defects in the weld, which can then be used to prove what was intended. As the study aims to ensure the safety of the canisters, the underlying question is: what is the probability of accepting a defective canister? However, combining different test results is not trivial, and the dependencies between different tests is now the subject on the focus.

2 Updated scope

Our work so far has revealed that the processes and methods used in demanding industrial applications, such as welding and testing the disposal canisters, are far from trivial. They are built by teams of experts during long timespan and the reports explaining the activities are also written on from experts to experts basis. Thus it is very hard, or almost impossible, to construct a rigorous image of the processes just by reading the reports provided. This became a problem for us.

If we were to validate an application for estimating the probability of accepting a defective canister as was originally intended and planned, a very thorough understanding of the testing methods and the data they produce would be essential. Unfortunately, the reports from VTT and Posiva leave many key details unexplained. Clarifying these obscurities by discussing with our contact persons would be too time consuming for both parties and is therefore not an option. Also, if our application would end up being built with just a single erroneous assumption, it might be totally useless.

Such a complex problem does not have just one correct solution. All the answers are just estimates on some level and based on many assumptions. Nor do we have the possibility to obtain a high enough level of expertise in this field in a relative short time period in order to be able to make reasonable assumptions on our own. Also, the data we would be able to get to build and test our application is just preliminary.

As a result of the issues mentioned above, and a request from VTT, the focus of our work is moved away from building an application towards a more general discussion. The goal of our work from now on is to present new ideas and approaches for answering our main research problem: how to take account uncertainties and dependencies in evidence. Thus, our work consists of discussing the possibilities and needs the new approaches offer and require.

Next, we will present results of the work we have done so far, including our findings on the testing processes, the characteristics of the available data and suggestions for the new approaches.

3 Findings and results so far

In this part, we discuss the subjects we have encountered so far and describe briefly the main areas of interest. As the starting point the non-destructive tests are performed, and with their results probability of detection curves can be derived. The probabilities can be handled by Bayesian inference framework, and one solution to take into account the dependencies between different NDTs are copula functions.

3.1 NDT

The main tools for detecting a defect in the copper canisters used by Posiva are four different non-destructive testing (NDT) methods. Visual testing (VT) and eddy current (ET) testing are mainly surface inspection methods and ultrasonic (UT) and radiographic (RT) testing are volumetric inspection methods [1]. However the technical details of these methods are not in the focus of this project. More important for us is to derive information about the POD (probability of detection) of each NDT. The goal is then to combine the gathered information to calculate the joint probability of all four NDT failing simultaneously. One problem is that the detection by different methods cannot be assumed to be independent. At least visual and eddy current testing respectively ultrasonic and radiographic testing might be pairwise correlated. The reason is the following. Given that VT detects a defect on the surface it is more likely that the defect is big than small. Given that the defect is big the probability that ET detects it increases as well. Thus for defect x:

$$P(\text{ET detects x} \mid \text{VT detects x}) \geq P(\text{ET detects x}).$$

A analogous argument holds for dependencies between UT and RT. That is why the joint POD is not the product of the marginal PODs. To be more

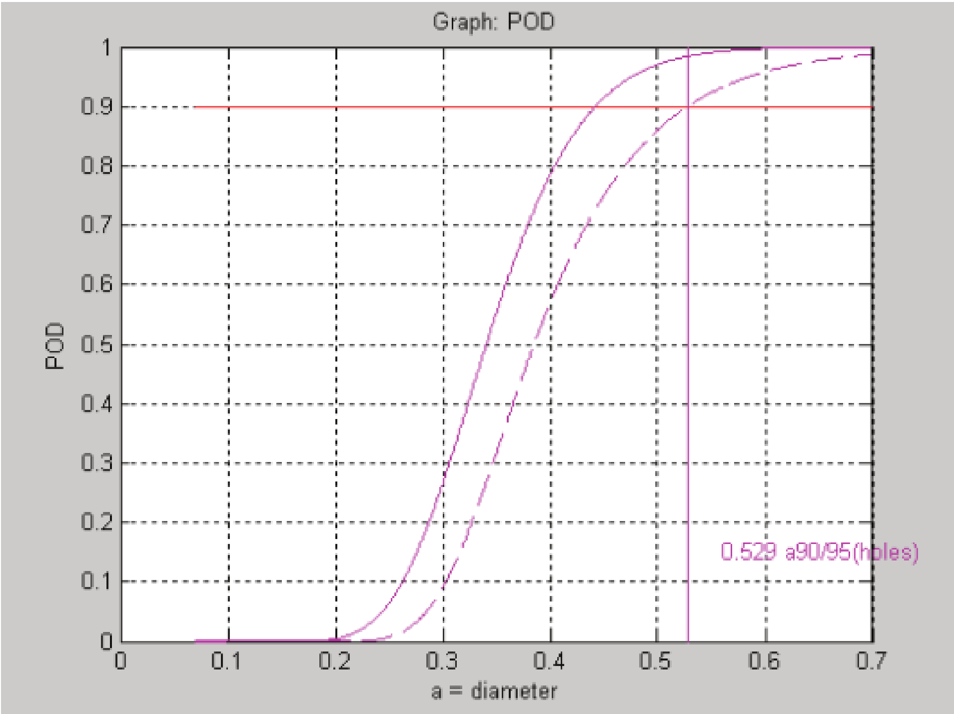

Figure 1: The POD for high frequency eddy current testing for holes. The diameter of the hole is on the x-axis and the probability of detection on the y-axis. Dashed line is the 95% confidence bound. [1]

precise the product of the POD's can be greater than the joint POD. But a too optimistic estimate is irresponsible or at least useless in a safety issue.

3.2 POD-curves

The data gathered from the NDT-tests are further processed to build up POD-curves (probability of detection). POD-curve is a sophisticated and widely used method to obtain a continuous relation and confidence intervals between the size of the defect and the probability of detecting it [2].

An example of a POD-curve is seen in figure 1, where we see the POD for high frequency eddy current testing for artificial defects, now drilled holes. The size of the hole is on the horizontal axis, whereas the probability of detecting a defect of such size is on the vertical axis. The 95% confidence interval is seen as dashed line. A point of particular interest is the 90/95-point, where we see the size of the defect which has 90% probability of detection with confidence of 95%.

Each NDT-test will have their own POD-curve or curves. For example, eddy current -test has multiple POD-curves based on the different energy settings of the test instrument. The POD-curves are the main tool we will be using when answering our research question.

3.3 Bayesian inference approach

The Bayesian approach of statistical inference is a method for updating an estimated probability function of an unknown parameter by adding more gathered information. Thus it might be a helpful tool for improving the probability of detection distribution when the number of copper canisters in the disposal increases.

The basic idea of the Bayesian inference is to improve the knowledge of the distribution of an unknown parameter θ by using the Bayesian formula of probability

$$P(\theta|X) = \frac{P(X|\theta) \cdot P(\theta)}{P(X)},$$

where the prior distribution $P(\theta)$ is the distribution of θ estimated without any further samples. $P(X|\theta)$ is the so called likelihood function which is the distribution of the observed samples given the parameter θ. Also, here X represents the samples that are used to update the prior into posterior. This function tells how likely it is to observe the observed samples, i.e. how well θ fits with the model. The posterior $P(\theta|X)$ is the distribution of θ given the information of the gathered observation. So it should be an improvement of the prior. The marginal likelihood $P(X)$ is the distribution of the samples. Since it is independent of θ it is a constant factor which is not necessarily needed if the family of $P(\theta)$ is known.

We are interested in the probability θ that a canister in the final disposal system is defective. Let p be the probability of a defect in a single canister and q the probability of failed detection of a defect in one canister. Then it holds

$$\theta = p \cdot q.$$

Let N be the number of investigated canisters and Z the unknown number of defect canisters being accepted by the testings. Z given θ is binomial distributed with parameters N and θ. That means

$$P(Z = z|N, \theta) = \binom{N}{z} \theta^z (1-\theta)^{N-z}.$$

This is the likelihood function for the Bayesian inference. The prior distribution $P(\theta)$ might be derived from the POD curves of artificially defected copper canisters. If the prior is chosen to be a Gamma distribution then the posterior will be a Gamma distribution as well with different parameters. If we consider p as a fixed number, estimated by experts or derived from the first artificially damaged canisters we have $q = \theta/p$. Thus the Bayesian approach of inference provides a applicable method for improving knowledge of the POD with increasing number of canisters.

3.4 Copula functions

The most generic way to combine dependent distributions into joint distribution is to use so called copula functions. Copula functions have been around for many years already (since 1959, first introduced by Sklar [6]), but they have not had so far much applications elsewhere than in finance. In our work we are going to study if we could use copula functions to describe the dependency structure between different NDT methods. During this project we are not going to have realistic data to make any assumptions about the dependency structure, but we might be presenting some ideas of what kind of data should be collected in order to estimate reliably the dependency structures between different testing methods.

The basic idea of copula functions is simple: to describe the dependency structure of different random variables. When dealing with copulas, we suppose we have for each different NDT method so called marginal distributions. These marginal distributions can then be combined using copula functions, and one possibility is so called Gumbel-copula, which has a property of indicating higher dependencies in the tails of distributions. This copula might be useful, as the probability of detecting a big defect with multiple methods is high, and on the other hand missing a small defect is probable for many methods. In multivariable case we can present the data as a multidimensional distribution. The multidimensional distribution contains the information about both: the marginal distributions and the dependency structure. If we want to study only the dependency structure, one solution is to separate it from the multivariate distribution and then get the actual copula. How the copulas can actually be exploited in this case is currently under investigation.

4 Task allocation and schedule

Since our goals have partly changed during the project, as was discussed earlier, it was also natural to allocate the tasks again according to the new objectives. We have already a relatively good understanding of the POD curves, how they are built and what their meaning is, so there is no need to concentrate on this anymore. We have also a simple Bayesian model built, and the concept of Bayesian inference is clear.

The most important thing from now on is to consider and provide new ideas for combining the information from different POD curves so that the dependencies between the testing methods are taken into account. As was discussed in the previous section, one idea is to apply copula functions. Thus, it is an essential task to understand the mathematical theory behind copulas and how they could be applied to our problem. Because of this, two members from the group will be concentrating on copulas, namely the members Tuovila and Piironen, since Tuovila has already some knowledge about the topic. The other two persons, members Backlund and Wolf, shall think of other ideas, like multimethod POD and so on. They can also start working with the report to avoid rush later on if there is not much to do with the other ideas. The task allocation shall be updated again instantly after we get better idea of how much work must be devoted for copulas. In short, during the next two weeks (weeks 16–17) focus is on the new ideas, copulas especially, and working on the final report shall begin at the latest on week 18.

5 Updated risk

In table 1 is presented the updated risks for this project. The risk of ambiguous starting point has realized at some level, and has lead us updating the scope of the project. As a new risk we have identified problems with the schedule, which might not hold. As a solution we have a possibility to simplify the problem to cover only two NDT methods. Efficient task allocation should help us with schedule related problems as well.

Also as a new risk we have identified the possibility that the findings or methods to cope with the dependencies are not necessarily usable for the client. At least the methods we will discuss and suggest will be presented in very general level, but not in detail using real data. This is because real data that could be applied to study the case properly is available, and we will be suggesting types of data that could be collected to make accurate assumptions about the dependency structures. This is, however, not critical.

Table 1: The risks of the project

Issue	Probability, Importance	Action	Status
Ambiguous starting point	High, High	Regular negotiations with VTT	Realized: scope has been updated
Personal workload	Medium, Medium	Working in smaller subgroups of two people	Not realized and not in sight
BBN is not applicable	Low, High	Valuable finding: case cannot be modeled as intended	Not realized
Team member quits	Low, Medium	Motivated group members	Not realized and not in sight
Difficulties in schedule	Medium, Medium	Task allocation, simplifying the study to cover only two NDT methods	New risk - not realized
Suggested methods are not usable	Medium, Low	Presenting the possibilities and suggesting what kind of data should be collected	New risk - not realized

For example copula methods for dependencies have not been widely used in industrial applications, but will probably be applied more in the future.

References

[1] Jan-Erik Holmberg, Pirkko Kuusela:
Analysis of Probability of Defects in the Disposal Canisters, Posiva Working Report 2011-36, `http://www.posiva.fi/tietopankki/` (12.4.2012)

[2] Luca Gandossi, Charles Annis:
Probability of Detection Curves: Statistical Best-Practices, ENIQ report No 41

[3] The European Commission's Joint Research Centre `http://ec.europa.eu/dgs/jrcindex.cfm` (29.2.2012)

[4] Radiation and Nuclear Safety Authority of Finland `http://www.stuk.fi/en_GB/` (29.2.2012)

[5] Judea Pearl: *Causality: Models, Reasoning, and Inference*, Cambridge University Press, 2000

[6] Alexander J. McNeil, Rüdiger Frey, Paul Embrechts: *Quantitative Risk Management*, Princeton University Press, 2005