DOI: 10.26345/EGRSE-021-17-203X

MACHINE LEARNING AND STRUCTURE DETECTION USED TO SUPPORT THE NUCLEAR WASTE REPOSITORY MONITORING – INFORMATION ABOUT THE CURRENT PROJECT

STROJOVÉ UČENÍ A DETEKCE STRUKTUR JAKO PODPORA PŘI MONITOROVÁNÍ ÚLOŽIŠŤ JADERNÉHO ODPADU – INFORMACE O PROBÍHAJÍCÍM PROJEKTU

Lenka Kosková Třísková[1]

Abstract

The question how to correctly remove the spent nuclear fuel from the environment is still unsolved. In the case of disposal, it is clear that very long term monitoring of the conditions of the repository is one of the key issues. In the process of searching for the suitable monitoring process, the geophysical methods in general should be taken in focus. In general, geophysics offers non-invasive monitoring methods of the physical processes running in the repository. Regardless of the finally selected methodology and monitoring procedure, the data interpretation means to detect significant temporal changes and anomaly in the data. Machine learning methods and structure detection algorithms can be used as a useful support method for the classical geophysical data interpretation. The article presents a related research covered by project Modern2020.

Abstrakt

Otázka, jak naložit se stávajícím vyhořelým jaderným palivem, zatím nemá jasnou odpověď. Jedním z klíčových problémů trvalých úložišť je zajištění dlouhodobého sledování podmínek v úložišti. Při hledání vhodných monitorovacích metod by měly být vzaty v úvahu geofyzikální metody, které nabízejí možnost neinvazivního monitorování fyzikálních dějů v úložišti. Nezávisle na zvolené metodě a procesu monitorování lze říci, že vyhodnocení získaných dat bude vždy vyžadovat detekci dočasných a anomálních změn v datech. Metody strojového učení a metody automatické detekce struktur v datech mohou v tomto případě sloužit jako vhodná doplňková metoda ke klasické interpretaci geofyzikálních měření. Článek seznamuje se souvisejícím výzkumem řešeným v rámci projektu Modern2020.

Key words

spent nuclear fuel repository, geophysical monitoring, machine learning, structure detection

Klíčová slova

úložiště vyhořelého jaderného paliva, geofyzikální monitoring, strojové učení, detekce struktur

1 Introduction

The machine learning (ML) methods have already found its applications in geosciences and remote sensing LARY, D. (2010). The machine learning consist of a lot of techniques, in the geophysical field the Artificial neural networks (ANN), support vector machines (SVM), self-organizing maps, decision trees and other already were used SHANIN, M. A. (2005) and SHANIN, M. A. (2015). The most commonly used are the ANN and SVM methods.

The principle of the machine learning algorithm is to emulate the human learning process – the algorithm extracts the behaviour of the system from the set of the training data. Modern ML techniques do not need a prior knowledge about the relationship between the data and system parameters. The limitation of the application of the ML techniques in geophysics is the need for a large training data set.

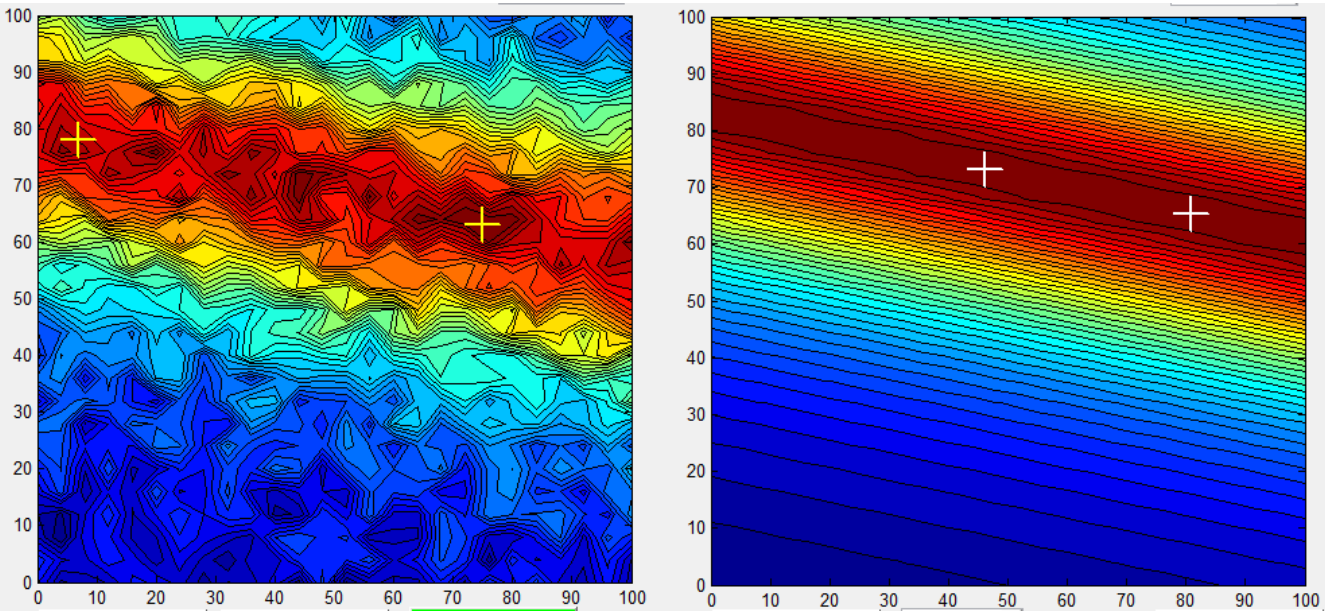

Fig.1 The illustration of the detection process, based on the synthetic data. The horizontal cylinder with noise (left picture, noise level is from 0 to 0.2 of the maximum value) and estimated value. Original depth is 22 m, estimated is 23 m

Regarding to the waste deposit monitoring, the ML algorithm can be used to detect significant anomalies in the data stream. Regardless of the monitoring technology, the physical conditions in the repository such as water saturation or temperature should either remain unchanged or change in a known manner. If any difference in monitored data is captured, it is necessary to recognize the cause of the change. Physical parameters in the repository can slightly oscillate around the equilibrium, which can be understood as a normal behaviour, or they can more dramatically increase/decrease. Such situation can be sign of a problem in the repository – for example the surrounding barrier may be corrupted and safety of the repository can be endangered.

The repository itself is strictly defined – it is a structure with defined and well known geometry, with stable homogeneous surrounding. It is possible to start pre monitoring to get the stable data stream as a reference training set. The other training set of the data can be a set of models of anomaly data which correspond to predefined problems occurring in the repository – increasing temperature over the prediction, modified water saturation, modified geometry etc. The task is to search in the data for any similarity with predefined anomaly situations.

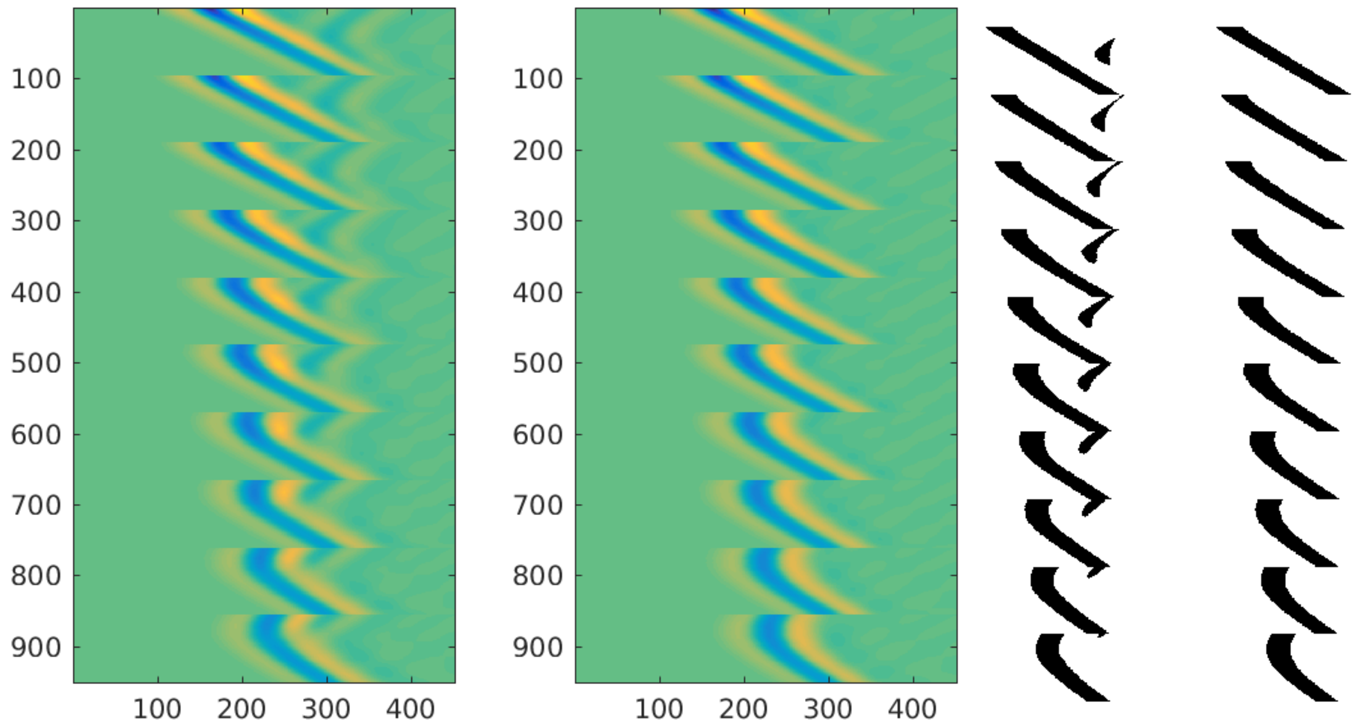

Fig.2 The difference being in the waveforms. On the left side of the picture is the original waveform for selected sources. The right side shows its conversion into the structures. The tunnel geometry is the same for both samples, the water saturation is different

The presented work is a part of the Modern2020 project which is focused to the nuclear waste repository monitoring procedures. The geophysical methods included to the project are: Electric resistivity tomography (ERT), Induced polarisation (IP) and Seismic methods (SM). For the ERT it is planned to set up a real monitoring experiment in the real operating condition of the repository. The IP is to be also run in the real operating condition with ERT as a supplementary method to distinguish the influence of changing water saturation and the temperature. For the SM, the full waveform seismic inversion is to be adjusted for the target application. The machine learning techniques are to be used as a supplementary method for the full waveform inversion. The task for the ML in SM is to classify the situation when the geometry of the repository is modified or the water saturation is dramatically different from the normal operating conditions

The structure of the input data is following – the model repository is a tunnel with circle shape. In the plane perpendicular to the tunnel are located two monitoring boreholes toward each other at an acute angle. The wave sources are located in one of the boreholes, the receivers in the other one. The model contains 114 sources and 104 receivers. The model was created for the dry tunnel, the fully water saturated tunnel, the tunnel with different geometry. For each of the configuration 114 data set are used as a training data set – one sample is the answer to one of the sources from all the receivers. The models were created and calculated by our project partner from ETH Zurich NUBER, A. (2015).

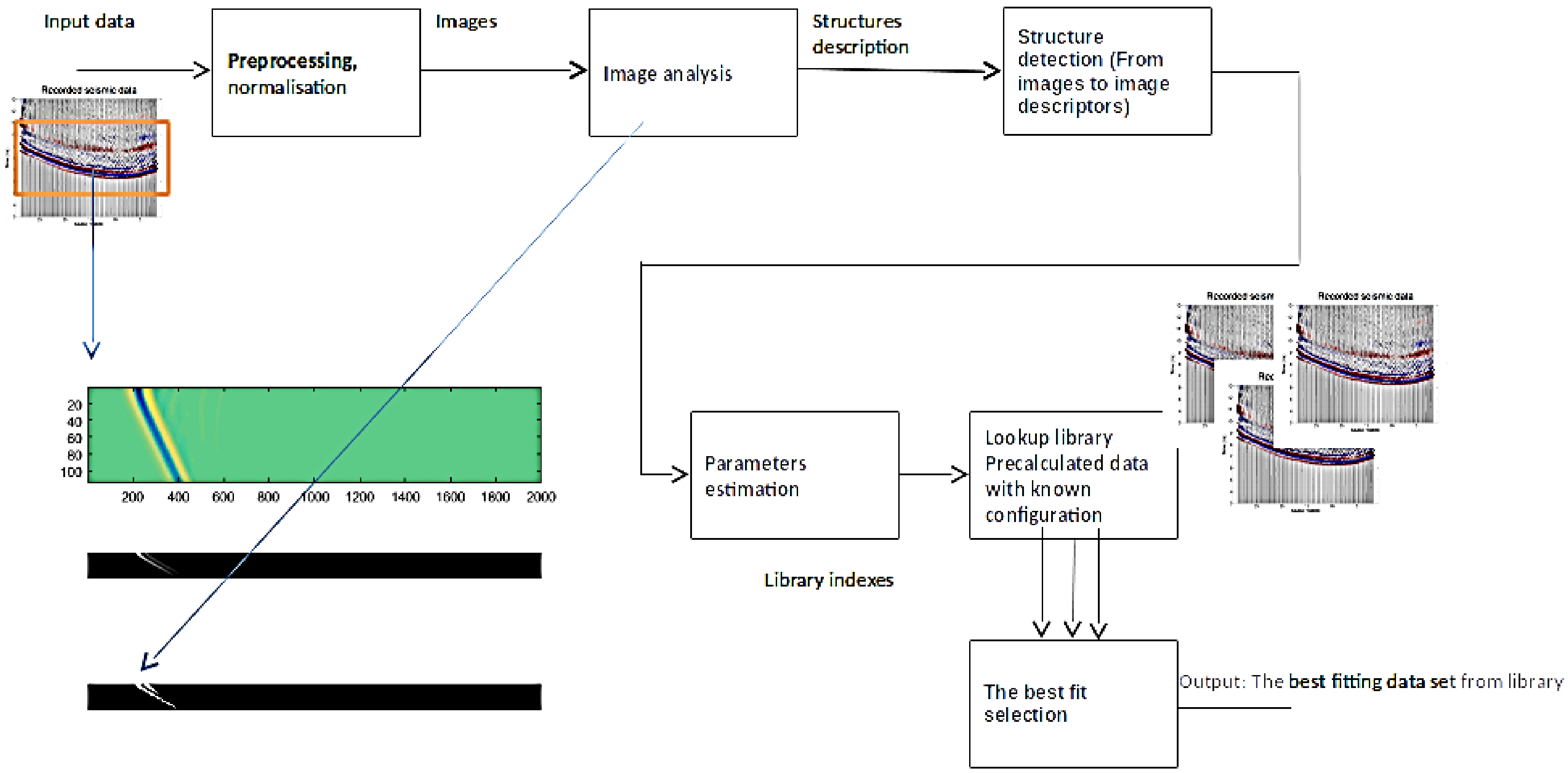

Fig.3 The structure detecting algorithm with lookup table

2 The classification based on structure detection

The task for the algorithm is to identify the anomaly dependent structures in the data and to classify the anomaly type. The structure (or feature) in the image processing domain is not clearly defined, but it can be understood as a typical image structure appearing in the data. The anomaly structure is a typical image which appears in the data. Presented algorithm, fig.1, is based on already existing implementation used to classify elementary anomaly structures in the near surface gravity data KOSKOVA, L. (2013). The initial research task for the algorithm was to classify the anomaly type (sphere, cylinder, rectangular mass) and its parameters (contrast density) in the near surface gravity data. Each type of anomaly structure has its typical appearing in the gravity data (a circle for the spherical and cylinder anomaly, sphere or rectangular shape for the rectangular mass etc.). Directly from the physical model the relationship between the size of the anomaly area and its parameters can be easily derived. A general function presenting a symmetric potential field anomaly can be expressed by simple equation SALEM, A. (2011):

$$f(r) = \frac{F}{\left(r^2 + z^2\right)^q} \tag{1}$$

Equation (1) presents symmetric potential field anomaly. For F and q symbols see tab.1, the z is the depth of the anomaly, r is the surface radius. Considering simple anomaly bodies, the *f*-function is a smooth function; maximum value is located above the anomaly centre. If the anomaly field is presented as 2D image, it gives spherical contours for sphere and vertical cylinder, for horizontal cylinder we obtain linear contours.

Tab.1 The* F *and* q *factors for simple geometrical bodies, gravity field;* γ *is gravitational constant,* $\mathbf{M_s}$ *is the contrast mass for the sphere and* $\mathbf{M_c}$ *is density contrast multiplied by cross sectional area* z *for both of cylinders

Anomaly type	F	q
Sphere	$\gamma \times M_s \times z$	1.5
Horizontal cylinder	$2\gamma \times M_c \times z$	1.0
Vertical cylinder	$\gamma \times M_c$	0.5

For all three bodies a linear dependency is between the depth of anomaly centre (*z*) and the surface location of the half-maximum value. The original algorithm takes the gravity data and converts the data into the grayscale image. As a next step, the noise reduction filter based on the Wiener filters is applied to enhance the structures in the data. Following step uses elementary transformations to detect the lines and curves in the images. If a structure is detected, it is labelled: the location, radius of the curves, and the angles of the lines are proposed to the classifier. The classifier selects the most probable synthetic anomaly model and gives the parameters (anomaly type, location, contrast density). The proposed models are calculated and compared with the original data. The best fitting proposal is finally selected as the underlying physical model. The algorithm was tested with the synthetic data, see fig.1.

For the application in the Modern 2020 project, the algorithm was modified. First problem was the last step of the algorithm: conversion from the anomaly parameters to the model. The underlying physical model for seismic is complex and therefore another approach for the anomaly estimation is being tested. Instead of direct calculation from the parameters, the possible anomaly models are pre-calculated and stored in the indexed library. Compared to the potential fields, the relation between the shape of the structure and the anomaly type is not as clear as for the potential field, see fig.2. The shape of the first arrival is detected; number of closed shapes in the image, the total area of the shapes and the distribution of the shapes is extracted. The whole modified algorithm is depicted in the fig.3. The structure parameters are used as the search keys in the pre-calculated lookup table. The lookup contains data from predefined models with different water saturation and tunnel geometry. The algorithm is now implemented and tested in the Matlab environment with limited number of models.

The possible weakness of the proposed solution is limited number of models available and complexity of the structures in the image. Unfortunately, when the water saturation reaches the expected level for the real operation in the tunnel, it is very hard to find a difference between the structures of normal and anomalous tunnel configuration. The classifier is now able to detect the dramatically modified geometry of the tunnel and can detect the situation when tunnel is dry instead of water saturated. The output of the first experiments also shows that the detection is sensitive for the different physical parameters of the rock, but not as sensitive for the different geometries of the tunnel.

The above described algorithms are based on fact, that at the beginning it is well known, what is the structure difference between the normal and anomalous data. The initial experiment shows, that the difference between the data can be unclear and the detection can misfit the anomalous data. Therefore it was decided to use the machine learning to test, if the neural network is able to extract the features itself. The task for the ANN will be the same as for the classifier: to distinguish between the normal tunnel operation (circular geometry and high water saturation) and the anomalous configuration (modified geometry, low water saturation).

Acknowledgement

The presented research is part of the Modern2020 project (Work package 3, Task 3.5). This project has received funding from the EURATOM research and training programme 2014-2018 under grant agreement No 662177. The overall objective of the Modern2020 Project is to provide the means for developing and implementing an effective and efficient repository operational monitoring programme, that will be driven by safety case needs, and that will take into account the requirements of specific national contexts (including inventory, host rocks, repository concepts and regulations, all of which differ between Member States) and public stakeholder expectations (particularly those of local public stakeholders at (potential) disposal sites).

References

LARY, D., ALAVI, J., AMIR, H., GANDOMI, W., ANNETTE, L. Machine learning in geosciences and remote sensing, *Geoscience Frontiers*, 7, 1, 2016, p. 3 – 10, http://dx.doi.org/10.1016/j.gsf.2015.10.006

KOSKOVÁ TŘÍSKOVÁ, L. Application of edge and line detection to detect near surface anomalies in potential field data, *In the ICPRAM 2013 proceedings*, 2013, p. 693 – 696.

NUBER, A., MANUKYAN, E. and MAURER, H. Enhancement of near-surface elastic full waveform inversion results in regions of low sensitivities, *Journal of Applied Geophysics*, 122, 2015, p.192–201. https://doi.org/10.1016/j.jappgeo.2015.09.020

SALEM, A. Multi-deconvolution analysis of potential field data, *Journal of Applied Geophysics*, vol. 74, 2011, p.151-156.

SHAHIN, M. A., JAKSA, M. B. Neural network prediction of pull-out capacity of marquee ground anchors. *Computers and Geotechnics*, *32*, 3, 2005, p.153–163. https://doi.org/10.1016/j.compgeo.2005.02.003

SHAHIN, M. A. A review of artificial intelligence applications in shallow foundations, *International Journal of Geotechnical Engineering*, 9,1, 2015, p. 49-60, DOI: 10.1179/1939787914Y.0000000058

Author

[1] Lenka Kosková Třísková, Ústav nových technologií a aplikované informatiky, Technická univerzita v Liberci, Hálkova 6, Liberec 461 17, lenka.koskova.triskova@tul.cz