Machine learning based multi-sensor fusion for the nondestructive testing of corrosion in concrete

Timo Haller [1], Christoph VÖLKER [2], Timo HARTMANN [3]

[2] Bundesanstalt für Materialforschung und -prüfung, Berlin, Germany

[3] Technische Universität Berlin, Berlin, Germany

Contact e-mail: Christoph.Voelker@bam.de

ABSTRACT: Half-cell potential mapping (HP) is the most popular non-destructive testing method (NDT) for locating corrosion damage in concrete. It is generally accepted that HP is susceptible to environmental factors caused by salt-related deterioration, such as different moisture and chloride gradients. Additional NDT methods are able to identify distinctive areas but are not yet used to estimate more accurate test results. We present a Supervised Machine Learning (SML) based approach to data fusion of seven different signal features to obtain higher quality information. SMLs are methods that explore (or learn) relationships between different (sensor) data from predefined data labels. To obtain a representative, labelled data set we conducted a comprehensive experiment simulating the deterioration cycle of a chloride exposed device in the laboratory. Our data set consists of 18 measurement campaigns, each containing HP, Ground Penetrating- Radar, Microwave Moisture and Wenner resistivity data. We compare the performance of different ML approaches. Many outperform the best single method, HP. We describe the intrinsic challenges posed by a data-driven approach in NDT and show how future work can help overcome them.

1 INTRODUCTION

Reinforced concrete is one of the most widely used building material. The main cause for deterioration is the corrosion of reinforcing steel (Nelson, 2013). Buildings in marine environments and traffic structures are particularly affected because they are exposed to large amounts of salt from seawater and de-icing agents. The most popular NDT method for the detection of active corrosion processes is half-cell potential mapping (HP) - where the electrochemical potential of the embedded reinforcement is measured against a reference electrode at the components surface. A highly negative potential indicates an active corrosion spot. It is generally accepted that the measurements are influenced by environmental factors (such as moisture and chloride gradients in the component, oxygen availability, etc.) and structural noise (caused by varying surface properties such as inhomogeneity, varying concrete cover, etc.) (DGZFP B03E (2014)). Theoretically, many of these influencing parameters can be determined with additional measurements. Nevertheless, a thorough physical understanding of the signal-environment interaction in the sheer variety of test scenarios is not available for data interpretation.

Earlier work demonstrated that even simple supervised machine learning (SML) functions outperform individual sensors in the detection of pitting corrosion in concrete under various controlled environmental conditions due to the use of correlation patterns in multi-sensor data sets (Völker, 2018). However, the use of more advanced approaches is by no means an implementation problem - numerous methods are freely available - but depends mainly on the variability contained in the training data. This means that predictions of SMLs are essentially only as expressive as the data from which they were derived. The acquisition of variable NDT training data involves considerable effort as damaging process are slowly and the test specimens in civil engineering are large. It is therefore particularly worthwhile to take a closer look at the mathematical effects which also influence the prediction quality. This work compares several methods from SML and provides empirical advice how these methods can improve our test data analysis. The reciprocal effect of these methods on small sample size (in our case 18 experiments), high dimensionality (seven features), heterogeneity (experiment contains the whole life cycle), and unequal group sizes (0.5 million intact VS. 2000 defect pixels) on the data side are analyzed. Our analyses complements existing studies (Völker et al. 2017 and 2018) in data-driven approach in NDT focusing on highly sensitive data analysis by exploring possibilities of improving the sensitivity of NDT methods with data fusion of multiple sensing methods through machine learning. Thus, the general objective to accomplish in this research project is to exhibit the possibilities to improve the sensitivity of NDT techniques for active corrosion detection through the process of multi-sensor data fusion with supervised machine learning.

2 THEORETICAL BACKGROUND

The experimental work is based on a case study conducted by the Federal Institute for Materials Research and Testing (BAM), which simulates the accelerated life-cycle of salt-exposed concrete under various controlled environmental conditions (Völker 2017). The design of the concrete specimens (L x W x H = 1.4 x 1.0 x 0.3 m³) is based on a traffic structure according to DIN Technical Report 102 - as these structures are exposed to large amounts of de-icing salt and are particularly susceptible to chloride-induced corrosion. The reinforcement is 10 mm structural steel (BST 500) spaced crosswise at15 cm.

The core aspect of the study was to achieve a high variability of the environmental factors moisture and salt on the one hand and to obtain referenceable conditions for later data labelling on the other hand. The first was achieved by accelerated chloride migration with potentiostatic polarization - a method that has been proven to accelerate the determination of the diffusion coefficient of concrete samples (e.g. according to ASTM C1202). This uses the endeavour of ions to strive for electric potential equalization. The migration of negatively charged chloride ions is accelerated by the build-up of an electric field in which they move to the positive pole (anode). In the experiment the embedded rebar was anodically polarized against a saline solution on the specimen (see figure 1 bottom, left.). Monitoring of the chloride penetration front to prevent unwanted corrosion was conducted using self-built anode leader sensor (marked in blue in figure 1, top, left).

A referenceable defect position was ensured by the eventual corrosion initiation at a predefined location according to Figure 1 bottom, right, on an electrically insulated rod (marked red in Figure 1 top, left). The rod was anodically polarization against a counter electrode beyond the critical corrosion causing potential. After the reaction started, the bar was short-circuited with the reinforcement cage to form an active corrosion system.

This simulated life cycle of a chloride exposed structure was regularly monitored with NDT-methods. The data set contains NDT methods that are commonly used for corrosion detection, namely ground-penetrating radar, HP, Wenner-resistivity (WR) and microwave-moisture (MW). A-priori Information for data labelling is available. GPR measurements were carried out using a GSSI, Inc. SIR20 device with a 2 GHz palm antenna (GSSI 2005). The data consist of two perpendicular polarizations per measurement that were collected subsequently with an automated scanner system with a lateral measurement spacing of 5 mm and a line spacing of 2 cm. The HP, WR and MW measurements were collected manually along a predefined measuring grid with a spacing of 10 cm. The HP data were collected using a Canin+ corrosion analysis system from Proceq. The reference electrode is a copper sulfate rod electrode (Proceq 1, 2016) DGZfP Merkblatt B 03. The WR measurements were collected using the Resipod probe from Proceq (Proceq 2, 2016) and the MW measurements were performed using an ID10 probe from HF-Sensorsimpermeability and porosity, which are due to the production-related higher compaction of the bottom part. The first campaign includes 10 measurements and second one consists of 8 measurements. The size of the measurement field is 126 x 86 cm². The measuring dates and the corresponding specimen conditions are listed in table 1. , the second consists of 8 measurements. The size of the measurement field is 126 x 86 cm². The measuring dates and the corresponding specimen conditions are listed in table 1.

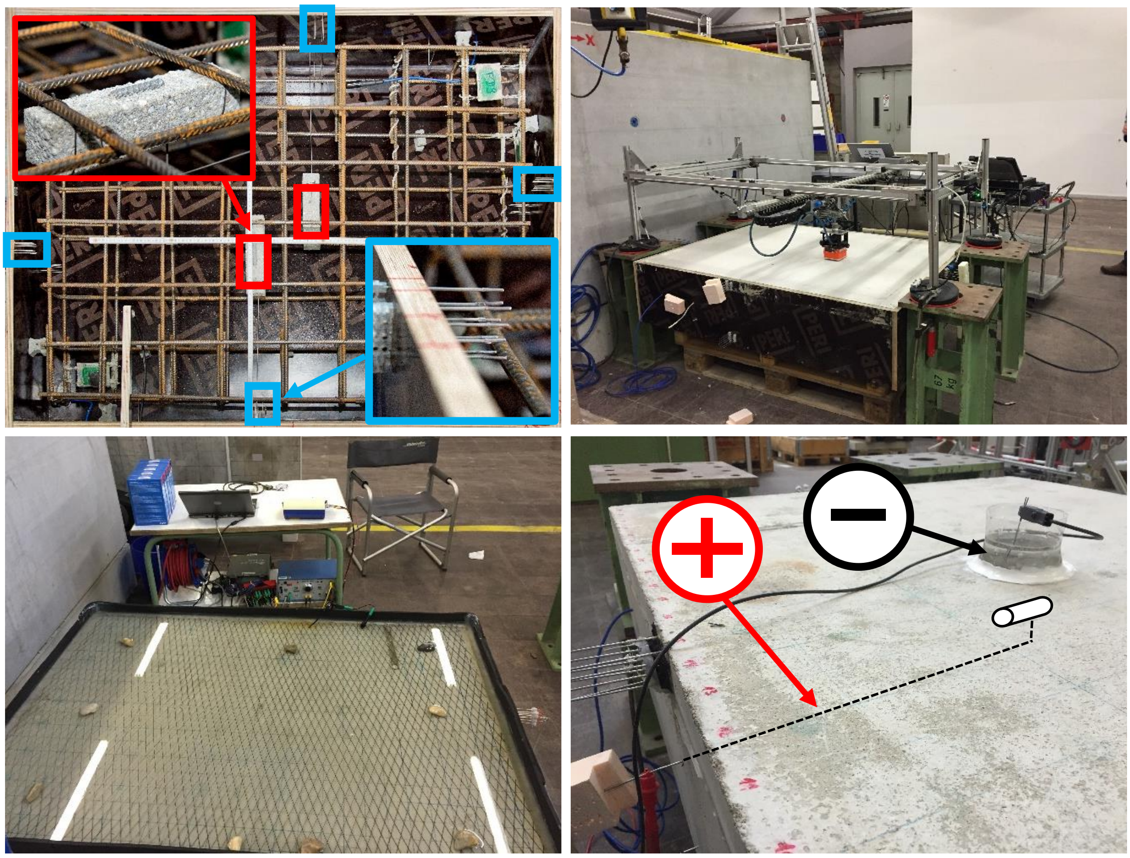

Figure 1: Different stages of corrosion experiment; top left: specimen before concreting with corrosion rod (red) and chloride migration sensors (blue); top right: automated data collection on specimen with scanner system; bottom left: salt water bath with counter electrode mesh for potentiostaticly accelerated chloride migration; bottom right: sketched corrosion initiation through anodic polarization of embedded corrosion rod (+) against counter electrode (-).

The data set contains 18 independently collected measurement campaigns with four different NDT methods on a specimen with varying conditions including moisture, chloride content, concrete quality and corrosion activity. Seven features were extracted from the NDT methods and are summarized in table 1. A representation of the features in spatial coordinates is provided in Figure 2. The characteristics grid-shape results from the fact that the features are with regard to the position of the reinforcement which is installed crosswise. The top row contains features from a pre-damaged condition of the specimen. The lower row shows the features in a scenario where environmental conditions have changed and corrosion has occurred. The dataset contains in total 50,037 observations, of which 1954 were labeled as defect/corroding pixels.

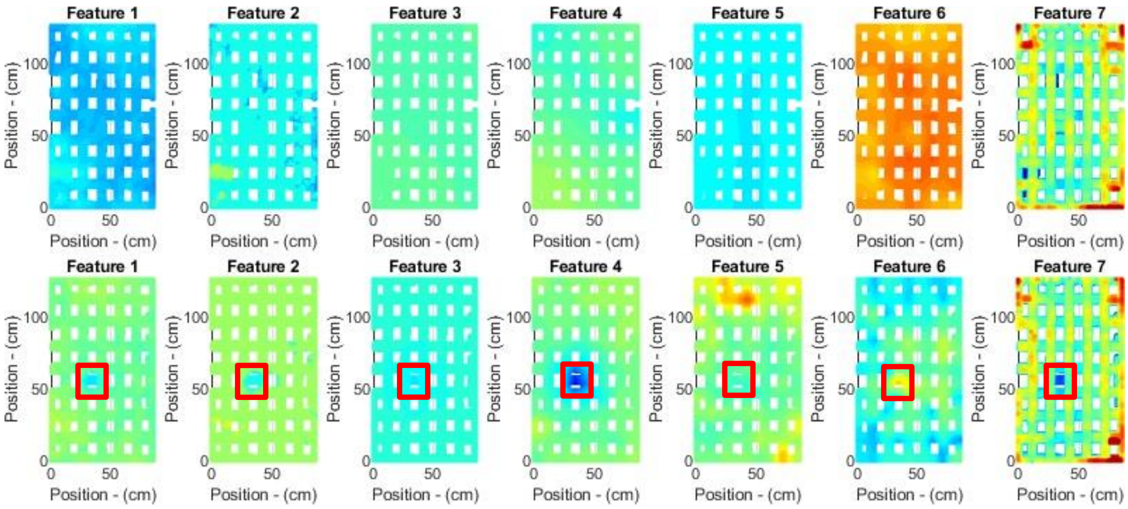

Figure 2: Example feature set in spatial coordinates; top: seven features (see table 1) for intact condition (campaign 2); bottom: seven features for corroding condition (campaign 5) with red-marked corrosion spot.

3 RESEARCH METHOD AND DATA

The basic hypothesis of this work is to evaluate the possibilities of improving the sensitivity of NDT methods with data fusion through machine learning; in this case for detection of the active corrosion in reinforced concrete. Based on the previously presented experiment for corrosion detection, the obtained results for the tested specimen in various controlled environmental conditions were used for data fusion. The realization of data fusion can be accomplished by heuristic approaches, such as supervised learning techniques.

Dependencies between features are typically represented in the feature space (FS). The FS is a mathematical vector space in which the measured features for each point are plotted over each other. A common application is regression analysis, in which a two-dimensional FS is used to describe the relationship between two features. However, the FS is a general concept that allows to describe the relationships between any number of variables. The key concept is that different classes (such as corrosion and no corrosion) populate different areas in the FS. SML algorithms estimate a so-called decision boundary (DB) to discriminate between these regions. So-called classification assigns new data to a class depending on which side of the boundary they appear on. In this study, we use support vector machines (SVM) as a machine learning tool for classification, which can recognize the relationships in data, based on the given labels (corrosion and no corrosion). SVMs provides prediction models, which can distinguish between intact and defect pixels in the dataset. The estimation of the DB with SVMs is based on the appropriate choice of kernels that transform regions in the FS into higher dimensional spaces. While it may be difficult to develop an intuitive understanding of this approach, it is a widely accepted and powerful technique for the estimation of DBs.

The performance assessment is based on the error rates sensitivity (a.k.a. true-positive-rate, TPR) and specificity (a.k.a. true-negative-rate, TNR) from test data that had not been used to calculate the DB. However, the data set is relatively small (18 experiments) and very heterogeneous (data for the entire life cycle are included). Thus, performance is determined in terms of an average performance using cross validation where all combinations of test and training campaigns are considered. More specifically, the dataset contains seven features, which are summarized in Table 1. From the overall 50,037 intact observations, 1954 observations were classified as defect pixels. Different feature constellations were tested for the application of SVM. The implementation of the classifier for supervised machine learning technique, as well as the preparation of the data basis, are done with the help of RStudio (version 1.1.456).

Table 1. The extracted Features from the underlying NDT signals

	NDT-Method	Parameter
Feature 1	Ground-penetrating radar (GPR)	energy of direct wave
Feature 2	GPR	dominant frequency of direct wave
Feature 3	Half-cell potential (HP)	rebar potential
Feature 4	HP	change of rebar potential
Feature 5	Wenner-resitivity (WR)	electrical resistivity
Feature 6	Microwave-moisture (MR)	relative moisture of concrete
Feature 7	GPR	depth-corrected top rebar reflection

4 RESULTS

The aim of this study was to improve the sensitivity of the prediction results. This corresponds to an improvement of wrong classification of defect pixels as intact. This type of misclassification is more problematic for chloride-induced pitting corrosion. Nevertheless, the balance between maximum sensitivity and appropriate false alarm rate should be treated on a case-by-case basis.

For the application of the machine learning technique, seven different feature constellations were tested. The large amount of data may contain information that does not contribute to the creation of the model. Reducing redundant information can therefore help to improve the model and to facilitate the pattern recognition. For the given database, the best result was achieved using all features.

Table 2. Prediction results using a linear SVM-Kernel for different feature selections. We report the True Positive Rate (TPR) of pixels that were rightfully detected as corroded and the True Negative Rate (TNR) of pixels that were correctly identified as non-corroded.

Used Features	TPR	not detected defect pixels	TNR	not detected intact pixels
All	0,688	609	0,987	460
3-7	0,682	622	0,992	381
3-6	0,588	1214	0,983	409
1-4,7	0,604	1018	0,995	245
3,4,7	0,544	1496	0,987	605
2-4, 6-7	0,644	1009	0,986	313
3,4,5	0,622	1198	0,985	460

For a higher accuracy in SVM based classification results may be possible by using more advanced kernels the transform FS with a non-linear function. Table 3 shows the result from various more advanced kernels which are detailed in Shalev-Schwartz et al. (2014).

Table 3. Prediction results for different kernels by using all features. We report the True Positive Rate (TPR) of pixels that were rightfully detected as corroded and the True Negative Rate (TNR) of pixels that were correctly identified as non-corroded.

kernel	TPR	not detected defect pixels	TNR	not detected intact pixels
Vanilladot	0,688	609	0,987	460
Anova	0,546	888	0,981	1385
Bessel	0,312	1345	0,972	1345
Laplace	0,643	697	0,985	171
Polydot	0,688	609	0,987	460
Rbf	0,576	828	0,983	242
tan	-	1954	-	1644

The creation of additional features based on mathematical transformation of existing features can help to improve the detection performance of the prediction models. Following, two features will be presented that have a positive effect on the prediction model.

The first feature (F1) determines the difference between each two data points of a feature and creates a new feature from it. The underlying assumption is that the slope of a feature may be a more sensitive indicator. However, it should be considered that the determination of the difference does not make a distinction in the coordinate system of the measuring field, as in some cases the compared pixels are spatially separated and not adjacent to each other. The application of F1 has shown an improvement in the detection performance. In the case of LR, the sum of the wrong classified pixels could be reduced from 1036 to 960, which corresponds to a decrease of 7.34%, compared to the results of LR without further feature engineering. To the same extent (7.28%), the detection rate of the unrecognized defect pixels could be reduced. The application of SVM showed that the number of total errors were decreased by 26.15%, while the unrecognized defect pixels decreased by 9.6%.

The second feature F2 uses a mathematical transformation of the data values of selected features. In particular, the decimal logarithm function is applied to the data values. The logarithmic transformation is used to modify the distribution of a feature when the original data values have an asymmetry in the distribution function. In case of LR, it can be seen that both the total number of misclassified pixels and the unrecognized defect pixels could be reduced by the application of the logarithmic transformation. Compared to the difference feature F1, the improvement is slightly smaller. The same effect can be observed in case of SVM. In general, using F2 coupled with SVM, the best overall results can be achieved. The summary results of the conducted experiments is summarized in Table 4.

Table 4. Prediction results by using F1 and F2

Feature selection	applied on feature	TPR	not detected defect pixels	TNR	not detected intact pixels
Prediction results by using F1 (kernel: vanilladot)					
3-7	3	0,693	599	0,988	382
3-7	3,7	0,703	580	0,988	364
3-7	3,4,5,6,7	0,702	583	0,988	384
2-4 , 6-7	3,7	0,639	705	0,986	71
2-4 , 6-7	7	0,616	750	0,985	131
all	3,7	0,709	569	0,988	364
all (laplace kernel)	3,7	0,718	552	0,989	89

Prediction results by using F2					
all (laplace kernel)	3,7	0,734	518	0,734	169
all (laplace kernel)	3	0,735	517	0,735	169
all (polydot kernel)	3,7	0,882	541	0,822	492
all (polydot kernel)	3	0,760	467	0,760	413

5 DISCUSSION

The reported study shows that an improvement in the detection of defective pixels is possible, by the fusion of data collected from different sensors. The results could further be improved by using additionally engineered features based on mathematical transformations. The prediction accuracy for the not detected defect pixels could be reduced by up to 43%, by using the designed feature F2. However, compared to the values without using F2, it must be noted that the number of total errors has been increased. In these cases, incorrect categorizations of the intact pixels were increased. For practical application, this would correspond to a false forecasting corrosion that would in fact be absent. The balance between the proportion of false reports compared to the highest possible rate of corrosion detection should be considered in every specific case. In principle, it is also possible to reduce the degree of false alarms while improving the detection rate of defect pixels. The best compromise between these two parameters might be achieved by using our engineering feature F1. In this case, the rate of not detected defect pixels was reduced by 33%, while the rate of not detected pixels was reduced by approximately 18%.

The practical applications of this work facilitate opportunities for future research into leveraging the potential of data fusion. Multi-sensor data fusion can be used for automated and improved measurement evaluation. The availability of large multisensory data sets, as well as the high degree of development of NDT methods forms the basis for new methods of damage detection. In the case of corrosion detection, the diagnosis is carried out in practice based on the standards of DGZfP (Merkblatt B 03 E) and ASTM (C876). However, Völker (2017) was able to show that the localization of corrosion based on these standards was unsuitable for the presented experiment.

The applicational limits of the SVM algorithm on new data sets for corrosion detection is not part of this work, which is why the generalization remains limited to measurements with similar constraints to the presented experiment. In particular, material properties such as concrete strength, pore composition, carbonation depth, as well as design parameters such as reinforcement thickness and concrete cover, might be parameters which can influence the prediction performance significantly. An extension of the data set by these parameters promises a higher generalization, as well as an improvement of prediction model itself. However, through the use of conventional and unmodified sensors it is conceivable to transfer the procedure of data fusion to new data set, by considering the restrictions to the generalizability.

Furthermore, data fusion could be used in other testing problems, whenever single measurement information cannot provide sufficient information. This applies, for example, to the determination of reinforcing diameters, the detection of alkali-silica reactions or the detection of gravel nests in concrete. Ideally, data fusion models will be used in combination with autonomous robot platforms to determine the building conditions more accurate.

In addition, the application of data fusion is also applicable outside structural engineering problems when automated tasks are to be performed on complex data sets with numerous boundary conditions.

6 CONCLUSION AND OUTLOOK

Taken together, our novel and interesting results demonstrate a significant improvement in corrosion detection in comparison to different international standards as well as single sensor measurements. Our data fusion model using the SVM machine learning technique, is applicable in automated and improved measurement evaluations. In this paper, the data fusion model has been successfully tested for detection of active corrosion in reinforced concrete.

The practical applications of this work facilitate opportunities for future research into leveraging the potential of data fusion. Subsequent studies must be carried out to investigate the performance of the SML approach based on new data sets with different experimental and environmental conditions.

7 REFERENCES

DGZFP B 03 E, 2014, Electrochemical Half-Cell Potential Measurements for the Detection of Reinforcement Corrosion, Berlin, Germany, German Society for Non-destructive Testing (DGZfP).

GSSI, "GSSI 2.0 GHz Antenna Manual," Geophysical Survey Systems, Inc., 2005. [Online]. Available: http://www.allied-associates.co.uk/pdfmanuals/MN31-079C%20Model%205100%20Manual.pdf. [Accessed 20 01 2015].

Nelson, S., 2013, Chloride induced corrosion of reinforcement steel in concrete: Threshold values and ion distributions at the concrete-steel interface. Chalmers Univ. of Technology.

Proceq 1, "Corrosion Analysis - Canin+," 2016. [Online]. Available: http://www.canin-concrete-corrosion.com. [Accessed 16 08 2016].

Proceq 2, "Resipod Resistivity Meter," 2016. [Online]. Available: http://www.proceq.com/nondestructivetestequipment/concrete-testing/moisture-corrosion-analysis/resipod.html. [Accessed 16 08 2016].

Shalev-Schwartz, S. and Ben-David, S., 2014, Understanding Machine Learning. Cambridge University Press, 179–189.

Völker, C., 2017, Datenfusion zur verbesserten Fehlstellendetektion bei der zerstörungsfreien Prüfung von Betonbauteilen, Universität des Saarlandes.

Völker, C., 2017, Labeled non-destructive testing data set for corrosion detection (down-sampled and balancedversion), Bundesanstalt für Materialforschung- und Prüfung.

Völker, C., Kruschwitz, S., Ebell, G., Shen, J., 2018, Towards Data Based Corrosion Analysis of Concrete with Supervised Machine Learning, NDE/NDT for Highway and Bridges: Structural Materials Technology (SMT 2018) and the International Symposium Non-Destructive Testing in Civil Engineering.