10th IMC
10th International Masonry Conference
G. Milani, A. Taliercio and S. Garrity (eds.)
Milan, Italy, July 9-11, 2018

APPLICATION OF A BIO-INSPIRED ANOMALY DETECTION ALGORITHM FOR UNSUPERVISED SHM OF A HISTORIC MASONRY CHURCH

Alberto Barontini[1], Maria-Giovanna Masciotta[1], Luís F. Ramos[1], Paulo Amado-Mendes[2], and Paulo B. Lourenço[1]

[1] ISISE, University of Minho, Department of Civil Engineering,
Campus de Azurém, 4800-058 Guimarães, Portugal
e-mails: albe.barontini@gmail.com, mg.masciotta@gmail.com, lramos@civil.uminho.pt, pbl@civil.uminho.pt

[2] ISISE, University of Coimbra, Department of Civil Engineering,
R. Luis Reis dos Santos 290, 3030 Coimbra, Portugal
e-mail: pamendes@dec.uc.pt

Keywords: Anomaly detection; Damage identification; Bio-inspired algorithm; Negative selection; Masonry structures

Abstract. *Variations in dynamic properties are commonly used in Structural Health Monitoring to assess the conditions of a structural system, being these parameters sensitive to damage-induced changes. Yet, such variations can also be due to changes in environmental parameters, like fluctuations in temperature, humidity, etc. By performing a continuous monitoring, the correlation between those factors appears and their variations, if no damage exists, result in a cyclic phenomenon. Negative selection, a bio-inspired classification algorithm, can be exploited to distinguish anomalous from normal changes, thus eliminating the influence of environmental effects on the assessment of the structural condition. This algorithm can be trained to relate specific extracted features (e.g. modal frequencies) and other monitored parameters (e.g. environmental conditions), allowing to identify damage when the registered value oversteps the confidence interval defined around the predicted value. Negative selection draws inspiration from the mammalian immune system, whose physiology demonstrates the efficiency of this process in discriminating non-self elements, despite the restricted number of receptors available to face a vast amount of aggressors. In this paper, a negative-selection algorithm based on a non-random strategy for detector generation is optimized and tested on the monitoring data of a prominent monument of the Portuguese architecture.*

10thIMC International Masonry Conference
POLITECNICO MILANO 1863
IMS International Masonry Society

1 INTRODUCTION

Developing methods for continuous real-time damage detection in aerospace, mechanical and civil engineering is one of the main goals of the Structural Health Monitoring (SHM). In this regard, anomaly detection algorithms offer great advantages as they allow a quick cost-saving strategy for the early-warning in case of fault by classifying the value of case-specific monitored features as healthy or anomalous through the solution of a forward problem. The algorithms succeed when they are able to correctly identify an anomaly in the system behavior, namely an abnormal variation in such features and their performance can be assessed in terms of false positive errors (a normal state labelled abnormal) and false negative errors (an abnormal state labelled as normal), see Table 1.

		LABELS	
		HEALTHY	*UNHEALTHY*
ACTUAL STATE	*HEALTHY*	*TRUE NEGATIVE*	*FALSE POSITIVE*
	UNHEALTHY	*FALSE NEGATIVE*	*TRUE POSITIVE*

Table 1: Possible results of the anomaly detection.

Mathematically, the problem stated above is a two-class or binary classification problem, whose aim is to label each element of a given data set as normal or abnormal and supervised "*pattern classification*" or "*pattern recognition*" algorithms are used to solve it [1]. The scope is to find a mapping function $f(X)$ between the input data, composed of the features $X_i=(x_i^1,\dots,x_i^n)$, and the output $Y=\{Non\text{-}self, Self\}$, where $S \subseteq \mathbb{R}^n$ is the feature space or space of the states of the system. The best mapping function is the one that gives the smallest possible error in terms of wrongly labelled X [1], [2]. The algorithm learns this mapping procedure, beforehand, by means of a given number of training data whose label is known.

Negative Selection Algorithm (NSA) states a family of classification algorithms based on a minimal common framework, firstly developed by Forrest et al. [3] and improved with different details [4], [5]. The NSA demonstrated to be a powerful anomaly detection tool, thanks to a very simple formulation which makes it extremely easy to implement and computationally inexpensive. Furthermore, the NSA does not require any "a priori" numerical or analytical model of the system and allows to carry out the classification just learning from data collected in the undamaged state. In fact, in many real cases, the anomaly detection must be formulated as a one-class problem, since only the elements belonging to one of the two classes, viz. to the normal state, are available in the learning stage. For these reasons, negative selection methods have been satisfactorily applied to fault diagnosis in mechanical engineering [6]. However, despite the promising results and the suitability for real case applications in the field of monitoring, so far, the use of NSA in the civil engineering practice is still very limited.

Experience demonstrates that the features used to assess the health state of a system are extremely sensitive to time-varying operational and environmental conditions as ambient factors can cause large variations in the monitored quantities, masking those induced by structural damage. The process of filtering out such effects with NSA has been addressed by Surace and Worden [7], [8], using the response transmissibility functions in terms of transverse displacement with respect to an excitation applied to a single node as feature for the classification. Though, in many real situations, data are collected under the operating condition of the system without recurring to forced excitation mechanisms, thus, it is more common to monitor the natural frequencies as state indicators [9].

In light of the above considerations, the first aim of the paper is to show how to set the key parameters of the NSA algorithm for anomaly detection in civil engineering applications, and how to test and analyze its performance using a statistical design of the experiments to normalize data and filter out environmental effects. The data used to carry out the experiments come from the monitoring systems installed in the Church of Monastery of Jerónimos in Lisbon, a prominent Portuguese monument which was object of an extensive experimental campaign in past years [10], [11]. The second aim of the work is to introduce a modified version of the canonic Real-valued Negative Selection Algorithm with constant radius (RNSA) and discuss the preliminary results of this new formulation.

The reminder of the paper is organized as follows: after a brief presentation of the proposed version of the algorithm in Section 2, Section 3 explains the design of the experiment conducted, whereas the results are reported in Section 4. Finally, in Section 5, the main conclusions drawn from the work are summarized.

2 NOVEL DETERMINISTICALLY GENERATED NSA

The NSA distinguishes self from non-self elements by assessing the matching between any monitored element and a set of detectors. Each detector is an element $d = (c_{det}, r_{det})$ of the space belonging to *Non-self* defined by a n dimensional vector c_{det}, namely the center, and a real value r_{det}, namely the radius or matching threshold.

The feature space is usually normalized. The unitary space $U = [0,1]^n$, where n is the dimensionality of the search space, is divided into two complementary subsets, *Self* and *Non-self*, such that:

$$Self \cup Non-self = U \qquad Self \cap Non-self = \emptyset \tag{1}$$

To map the elements of the problem space onto the unitary space, it is necessary to introduce a normalization algorithm. Normalization is essential for some data mining processes and increases both the speed of learning and the effectiveness. It consists in rescaling the input value into a specific output range, usually [0,1]. However, the performance of the algorithm can be affected by the type of normalization used [12]. In the present study, the conventional min-max normalization algorithm is employed. Min-max normalization maps each value x_{ij} of the attribute m into a new interval which ranges between 0 and 1 according to the following formula:

$$\bar{\bar{x}}_{nm} = \frac{x_{ij} - \min(m)}{\max(m) - \min(m)} \tag{2}$$

The main issue is the "out of range" problem, since during the monitoring there might be elements of the feature space whose values exceed the bounds known during the train. Some authors suggest a 20% increase of the bounds and it is evident that problem-specific knowledge helps defining a more suitable limit. Usually, all the values which lay outside the boundary in the problem space are normalized into the limit of the bound, so that any inverse mapping becomes impossible. In this study, the "out of range" values are classified into two specific sets: (1) suspension of the anomaly evaluation if the temperature range is exceeded, (2) "unhealthy" element labelling if the frequency range is exceeded. The bounds of the temperature are limited to the values available in the training data, whereas for the frequencies a 20% variation margin is allowed.

The NSAs are all composed of two stages: (1) censoring and (2) monitoring. During the first stage of the process, a given number of detectors are generated rejecting anyone that matches an element in *Self*. The matching threshold can be set in order to allow an overlapping of the detector (r_{det}) with the self set (r_{self}). In the second stage, new collected data z is matched against the detectors. The closest detector is identified and the element is classified as *Non-self* if the *distance* (c_{det}, z) < r_{det}.

The formulation of the distance and the definition of the matching thresholds form the matching rule that in the unitary space defines the shape of the detector.

Gonzalez et al. [1] developed the first real-valued NSA with many similarities to the binary coded greedy algorithm [13]. Each detector had a fixed radius and its center was defined according to two strategies: (1) to move it away from the training data and (2) to keep it separated from the other detectors (coverage maximization). Randomly generated detectors are likely to present blind holes of different size and shape between them or between them and the self elements. Li et al. [14] faced such issue by exchanging the random generation with a deterministic generation based on the division of the problem space in hypercubes.

In this paper, a similar concept is followed and adapted to the specific problem space. The difference is in the inclusion of the moving away process. In fact, operating in a two-dimensional space (normalized temperature – normalized first frequency) with Euclidean distance, the problem of covering the feature space with the detectors corresponds to the problem of covering a rectangle of side length equal to 1 with circles, allowing overlapping and avoiding any holes. Although this version of the algorithm loses generality, being suitable only for the 2-D space, it results easier to implement and still largely applicable to many monitoring problems since it is common to work with two main parameters: (1) the fundamental frequency of the system, easy to measure, and (2) the temperature, whose influence on the frequencies is well-known.

Despite in discrete geometry a few studies have already addressed the optimization of the distribution of a given number of circles to cover a square [15]–[17], here a regular grid is preferred because it allows a variable number of detectors, thereby guaranteeing a complete coverage of the feature space at the cost of a bigger overlapping area. In detail, the unitary flat surface is subdivided into a given number of squares and the detector is assumed as the circumscribed circle, considering two well-known results of discrete geometry, see Figure 1 [16]: (1) a circle covers the maximum total length of two parallel lines when its position is symmetric and (2) a circle covers the maximum total length of two perpendicular lines when its position is symmetric and the intersection of the lines is on the boundary of the circle.

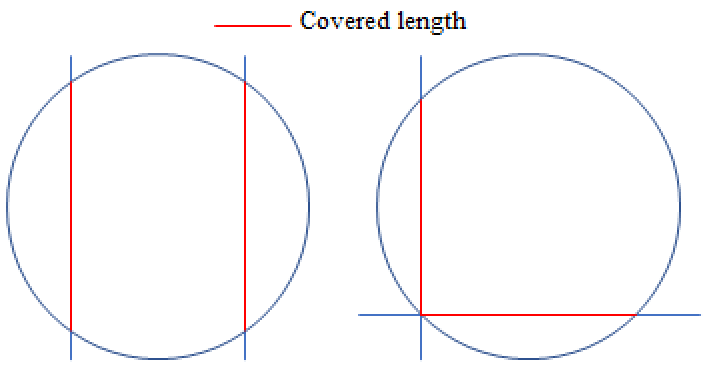


Figure 1: Maximum coverage of parallel and perpendicular lines with a circle.

Once the number of division n of the side of the space is defined, the side of each inscribed square can be computed as:

$$b = 1/n \tag{3}$$

and the radius of the detector results:

$$r_{det} = \sqrt{\frac{b^2}{2}} \tag{4}$$

The detectors so generated are matched with the training set and the ones that match a self-element are collected into a different temporary set. From this set, only the external elements are saved, namely the ones with maximum and minimum normalized frequency for the same normalized temperature, whereas the others are deleted. Considering the distribution of the features, this approach allows to surround the samples. Then, the saved elements are moved away from the self data as in the canonical form of the NSA and, finally, added back to the detectors set. When moving away from the self elements, the adaptation rate automatically reduces the step size of the detector to avoid an excessive jump and to help the convergence to the minimum allowable distance. In detail, the following formulation is adopted [18]:

$$\eta_i = \eta_0 e^{-i/t} \tag{5}$$

where η_0 is the initial value of the adaptation rate and t is a parameter that controls the decay.

The pseudo-code of the censoring follows below:

1. Define (r_{self}, b, η_0, t, $dist$)
   Generate the set D as a regular grid of detectors every b from $b/2$ to (1-$b/2$) whose radius is $r = \sqrt{b^2/2}$
   Initialize as 0 the counter *age* of the number of attempts of moving the detector
2. Calculate, for each detector d_i the closest self element, *NearestSelf*
   2.1. If $distance(d, NearestSelf) < Detselfdist$
   Move d_i to a temporary set TD
3. For each element vertical of the grid save the elements of TD with maximum and minimum normalized frequency and discard the others.
4. Calculate *NearestSelf* for each detector d_i in TD
   4.1. If $distance(d, NearestSelf) < Detselfdist$
   age = *age* + 1
   $dir = \frac{d_i - NearestSelf}{distance(d, NearestSelf)}$
   4.1.1. If *age* > *t*
   Move d_i to the set *D*
   4.1.2. Else
   $d_i = d_i + \eta_i \cdot dir$
   4.2. Else
   age = 0
   Move d_i to the set D
5. End

3 DESIGN OF THE EXPERIMENT

The NSA algorithm depends on several parameters that must be appropriately set before running the analysis, like: (1) the number of samples in the training set, (2) the self radius r_{self}, (3) the detector radius r_{det}, here related both to the number of divisions n of the feature space and to the side of the inscribed squares b, (4) the factors to define the process of "moving away", namely the adaptation rate η_0 and the maturity age t, and (5) the threshold or censoring distance between self-element and detector. The definition of the distance is a non-quantitative parameter.

Many authors have denounced a general lack of attention to the parameter setting in the application of Soft Computing methods despite the significance of this aspect for the outcome of the process [19]. Therefore, in the present study, the Design Of Experiment (DOE) approach is followed to draw objective conclusions. The DOE is a procedure for planning the experiments by changing the values of selected factors and observing the consequences in the dependent variables. As for this specific case, a three-level full factorial design is applied, identifying three main factors: the number of divisions of the side of the unitary space, n, which is related to the detector radius r_{det} through Eq. (4); the maturity age t; and the censoring distance $dist$, defined as a function of the two radii, i.e. self radius r_{self} and detector radius r_{det}. Table 2 summarizes the values adopted for all the three identified parameters.

LEVELS	n	t	$dist$
1	*40*	*5*	r_{det}
2	*20*	*15*	$0.5\cdot(r_{self} + r_{det})$
3	*10*	*30*	$(r_{self} + r_{det})$

Table 2: Values adopted in the three-level factorial design.

As for the adaptation rate η_0, a value equal 0.005 is chosen, using only one parameter to assess the influence of the moving process. Finally, the self radius is set to 0.05 that corresponds to a temperature variation of 1.2°C and a frequency variation of 0.03 Hz, while all the distances are computed using the Euclidean formulation. It is worth noting that in the three-level factorial design the levels are usually low, intermediate and high, although the first factor, viz. the number of divisions, is decreasing and is inversely proportional to the detector radius which is, thus, increasing.

The detector generation proposed in this work avoids any random effect, except for the selection of the training data (set to 300 elements), thus the random generation of the training set is repeated 6 times. After the training stage, the remaining data (480 healthy-states and 2 anomalies) are used to simulate the monitoring stage. For each combination of the levels and each repetition, the number of generated detectors, as well as the right classifications, True Positive (TP) – True Negative (TN), and the wrong classifications, False Positive (FP) – False Negative (FN), are collected and used for the analysis.

In order to assess the robustness, the optimized version of the algorithm is further tested over a set of random problem instances artificially created according to the probability distribution of the recorded data set. All the numerical experiences are conducted using OCTAVE, an open source software which uses the same programming language as MATLAB.

4 DISCUSSION OF THE RESULTS

As stated in the Introduction, the revisited version of the NSA algorithm is here implemented making use of the data collected from an extensive experimental campaign carried out on the Church of Monastery of Jerónimos in Lisbon [10], [11]. The focus was on the external temperature measurement and the first natural frequency extracted from the automated dynamic identification process (Figure 2). As a clear correlation between these two quantities has been observed [10], continuing the monitoring it is essential to assess whether any new couple of recorded values (Temperature, First Frequency) belongs to the same distribution or it is an anomaly linked to a possible damage condition. The scope of the analysis is to train the classification algorithm over a set of self elements (i.e. a sub set of measurements belonging to the health state of the system), setting the best values for the intrinsic parameters of the algorithm itself, and to validate its performance using all the remaining data of the experimental campaign as well as other numerically generated data.

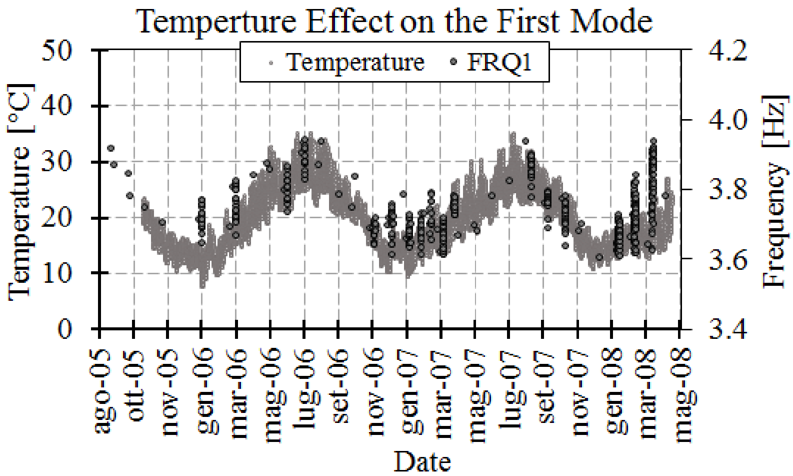


Figure 2: Trend of temperature and first natural frequency during the monitoring period.

The examination of the results allows to understand that the algorithm produces no false negatives, but the number of the anomalies within the data is not sufficient to draw any further conclusions on this aspect. Indeed, in many real applications, there is no a priori knowledge of the anomalies.

The cube plot of the algorithm outcome (Figure 3) shows, for each combination of levels, the average number of detectors and false positives over the 6 repetitions. In a cube plot the coordinates of each node correspond to the level (low, intermediate or high) of the factors which represent in turn the axes of the three-dimensional space. It is intuitive to see that decreasing the overlapping during the training (moving from level 1 to 3) leads to a reduction of false positive errors, as increasing the detector radius strongly affects the reduction of the number of detectors. The maturity age has no influence on the number of detectors, which is coherent with the algorithm since this number is defined before moving the detectors away from the self training set, but it is interesting to notice that the effect on the false positive results is also small.

The algorithm performance is assessed in terms of detectors generated and false positives normalized to the highest values found during the experiment (1423 detectors and 269 FP). For this purpose, a weighted objective function which gives more importance to the FP (75%) and less to the number of detectors (25%) is used as it is better to compromise on the memory allocation than on the error reduction. It is noted that the percentage weights have been arbitrarily chosen. Different percentage values or more sophisticated multi-objective approaches can be adopted, but these aspects are not addressed in the present work. For each combination of the factors, the following objective function is evaluated:

$$f = 0.25 \cdot \frac{detectors}{1423} + 0.75 \cdot \frac{FP}{269} \tag{6}$$

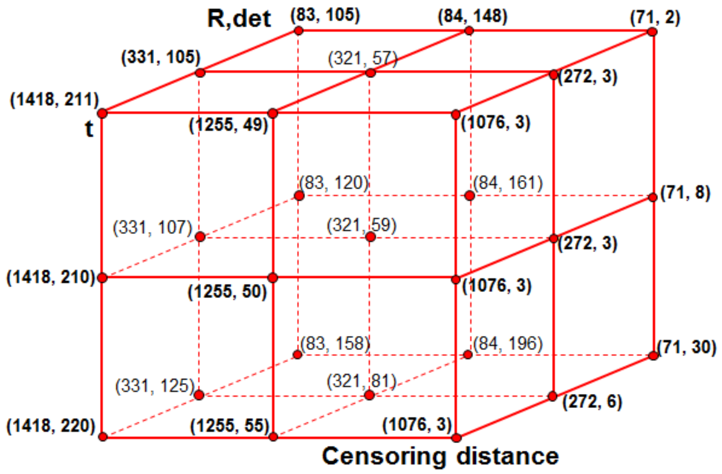
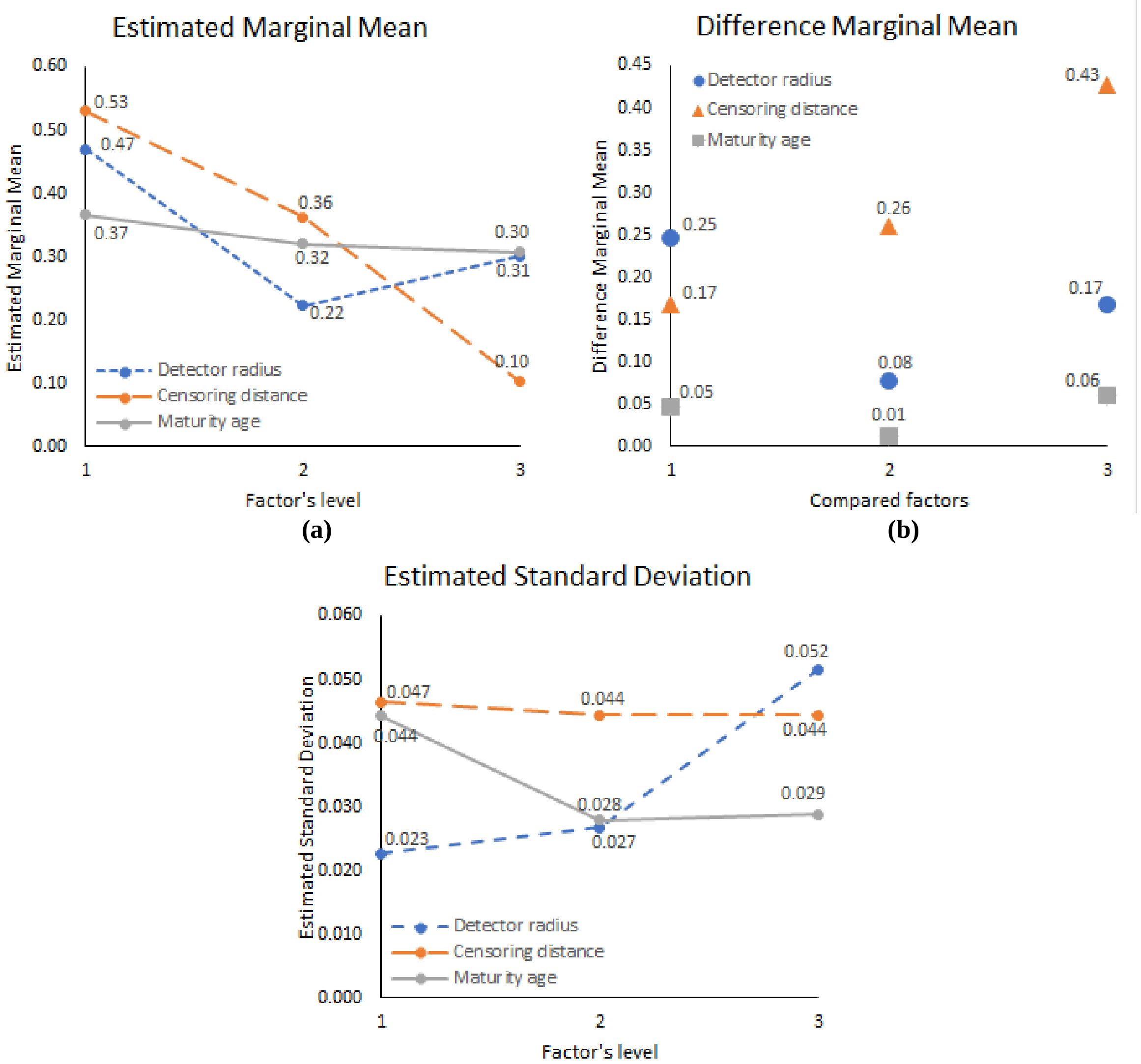


Figure 3: NSA cube plot. The axes represent the factors of the 3-D space; the nodal coordinates correspond to the factors' levels; and the numbers between brackets indicate the average number of detectors and false positives for each combination of levels.



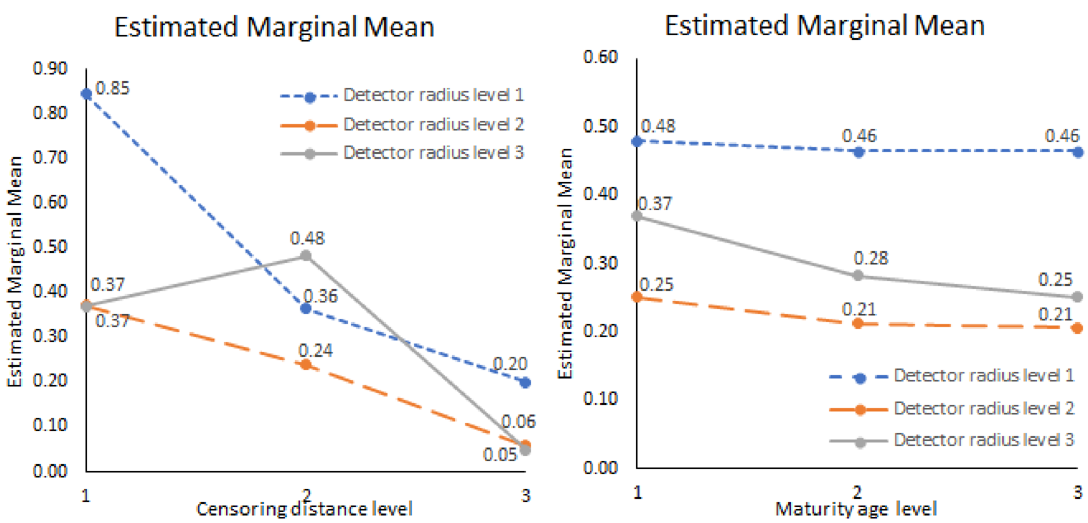
(a) (b)

(c)

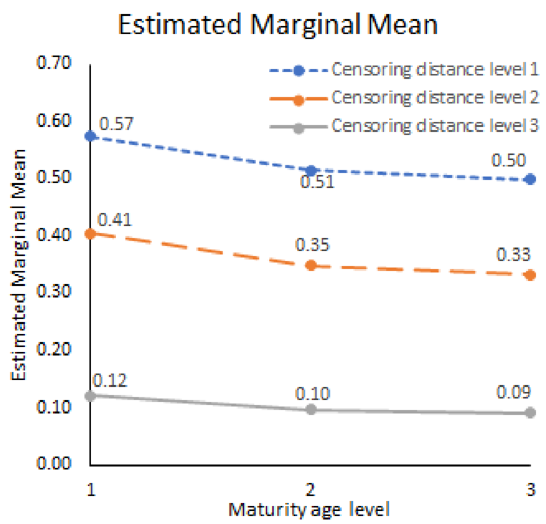
Figure 4: Statistical evaluation of the influence of setting parameters on the algorithm performance: (a) estimated marginal means, (b) differences between marginal means and (c) standard deviations for the three independent variables.

The marginal means, namely the means of the objective function *f* averaged across all levels of the factors (Figure 4a), allow to better weigh their effects and to define the best level to minimize the objective function. According to the graph, the best values of the setting parameters are obtained for: detector radius level 2, censoring distance level 3 and maturity age level 3. In what concerns the difference between the marginal means (Figure 4b) for the different factors' levels, this quantity is a good indicator of the significance of a certain variable. In this particular case, the maturity age results to be the least significant, meaning that the other two variables have higher influence on the algorithm performance. Finally, the standard deviations estimated for the different factors at the different levels (Figure 4c) provide an estimate of the effect of randomness on the factors. Unlike the censoring distance, which features minor variations in the standard deviation values over the different factors' levels, both detector radius and maturity age show higher dispersions. The interaction plots in Figure 5 show the marginal means for different values representing the interaction between two factors. This further comparison enables to improve the setting based on the marginal means of Figure 3a. In fact, although the marginal mean for the intermediate level of the detector radius is lower than for the high level, an evident interaction with the censoring distance exists. Assuming the high level of the censoring distance, the marginal means for the high level of the detector radius is the lowest (Figure 5a). Therefore, the best combination for the factors is the one with the high level also for the other variables.

It is remarked that the analysis of the interaction plots is essential when the experiment follows a fractional design and the best combination of setting parameters needs to be found to minimize the objective function *f* and maximize the algorithm performance. Since, in the present study, the design of the experiment is full factorial it is possible to confirm that the lowest value of *f* is obtained from the combination with high levels of the three variables, as clearly highlighted by the inspection of the interaction plots.



(a) (b)



(c)

Figure 5: Interaction plots of the three setting parameters over the different levels: (a) detector radius vs censoring distance; (b) detector radius vs maturity age; and (c) censoring distance vs maturity age.

The last step of this work consists in monitoring a series of 1000 health-state data, randomly generated according to the joint probability mass function of the two variables (temperature and frequency) discretized into interval of 2°C and 0.02 Hz respectively, and a series of 520 outliers, 40 per each interval of 2°C. The outliers are randomly generated according to the marginal distribution of the frequencies into a temperature interval to be out of the following range:

$$sup = Q_3 + 1.5(Q_3 - Q_1) \tag{7}$$

$$inf = Q_1 - 1.5(Q_3 - Q_1) \tag{8}$$

where Q_1 and Q_3 are the first and the third quartile, respectively.

Due to the consistency of the results of the previous analysis, the selection of the detector set is unlikely to affect the outcome. The detectors generated in the second repetition of the experiment are used as an example for this test. In Figure 6 such a set of detectors is displayed, compared with self and non-self acquired data.

The new simulated data are normalized to the same range of the acquired data and plotted together in Figure 7. All the outliers normalized to values bigger than 1 or lower than 0 are rejected. Therefore, 38 elements are discarded in total.

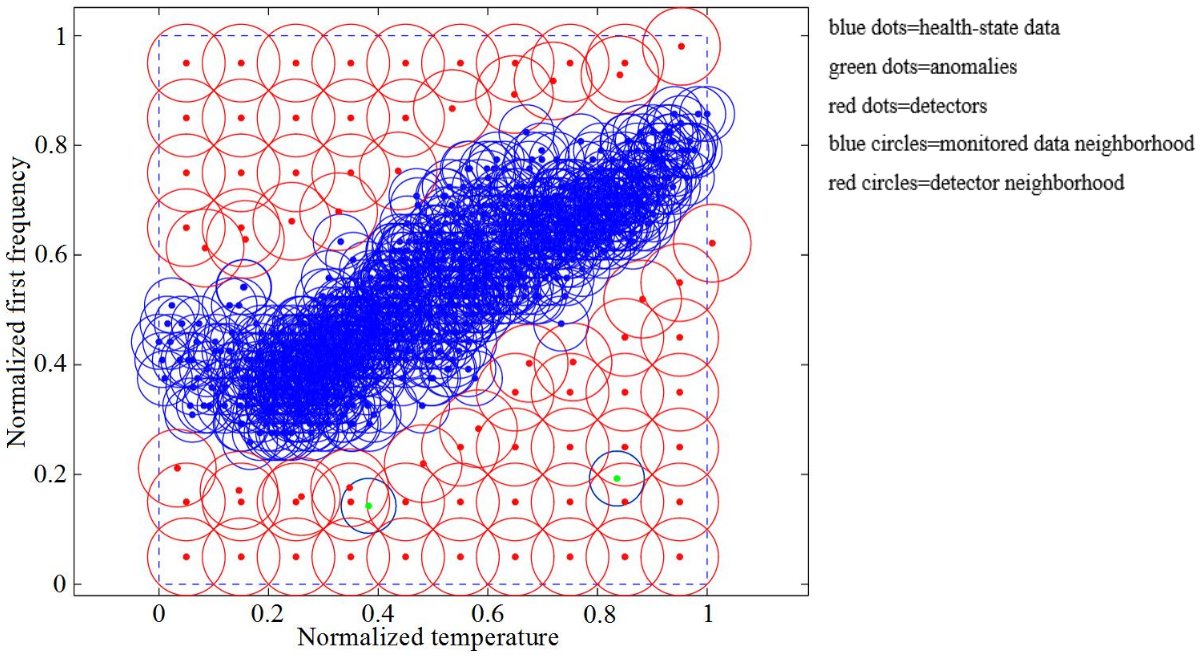


Figure 6: Result of the training and monitoring stages with the optimal setting (level 3 of all the variables).

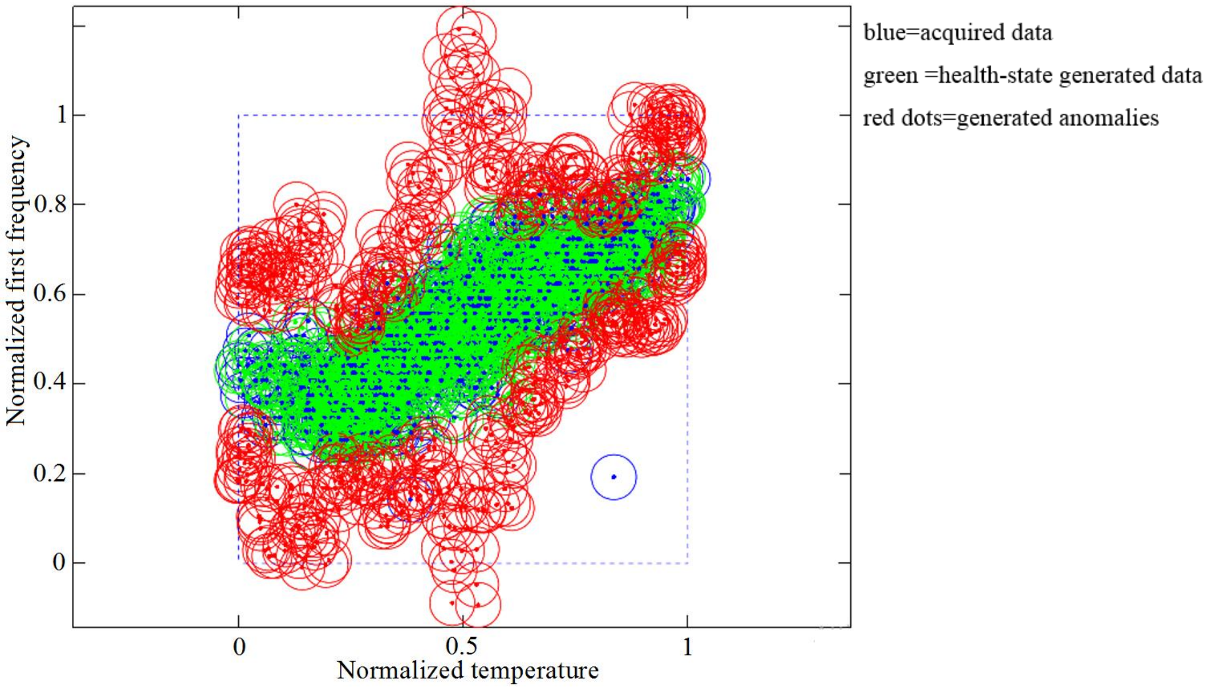


Figure 7: Comparison between acquired and generated monitoring data.

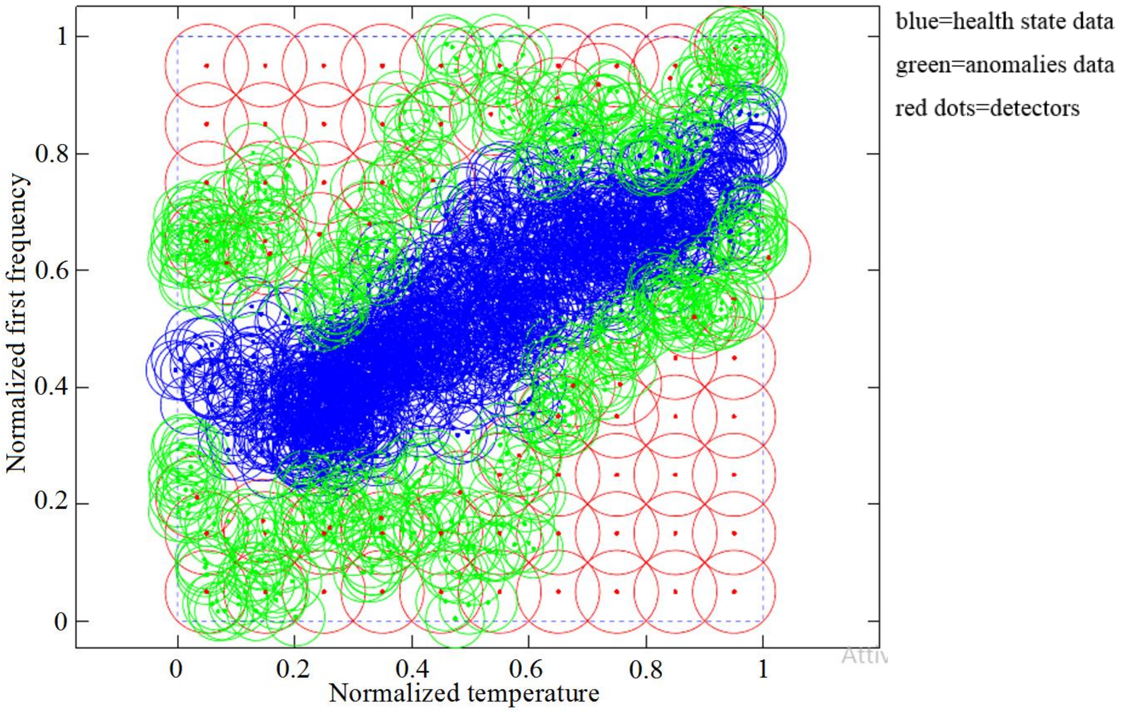


Figure 8: Simulated monitoring stage.

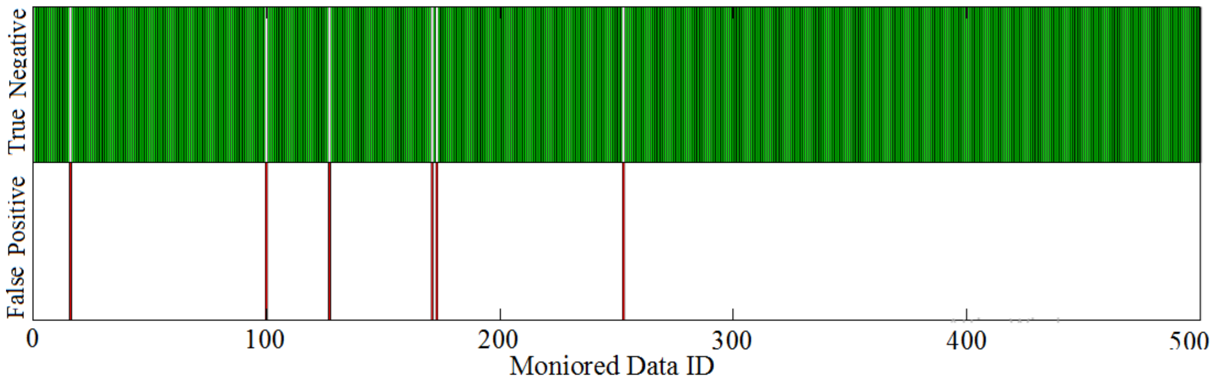


Figure 9: True Negative (green) against False Positive (red).

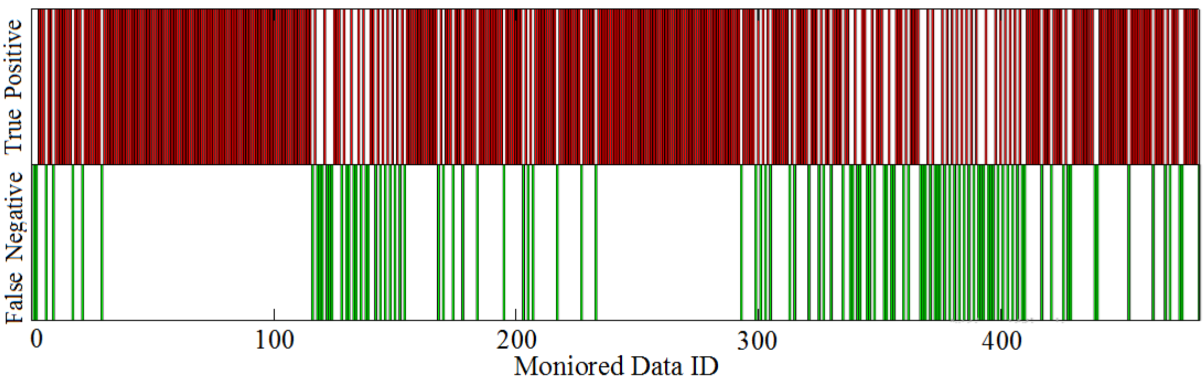


Figure 10: True Positive (red) against False Negative (green).

In Figure 8, the detectors and the data used for the simulated monitoring are plotted together. It is remarked that among the Health-state data only 10 false positive errors emerge (1%), see Figure 9, whereas among the Anomalies the false negatives are 107 (about 22% reduced to 21% considering the out of bound data rejected), see Figure 10. Such a value, clearly relevant, is likely due to the formula used to randomly generate the anomalies, which lack any reference to real anomalous conditions. Indeed, analysing Figure 8, several green dots (anomalies) can be noticed between the blue ones (normal states).

5 CONCLUSIONS

NSA is a family of classification algorithms suitable for one-class classification problems (anomaly detection). In this paper, a new NSA based on a deterministic generation of the detectors for a 2-dimensional feature space is applied to filter out the environmental effects from the monitoring data collected from the Church of Monastery of Jerónimos, in Lisbon.

The influence of three selected parameters is analysed to assess the algorithm performance and to define the best setting. It is found that two are the parameters that result relevant for the performance of the algorithm: matching threshold and detector radius. Afterward, the algorithm with the best setting is tested over artificially generated data of both healthy and anomalous states. If appropriately formulated, the proposed NSA shows promising results: indeed, the algorithm configuration selected for this study allows to reduce the false positive answers to a negligible 1% of the available data, despite a still relevant 21% of false negative classifications. Plausibly, such a value has been affected by the method adopted to generate artificial non-self elements lacking any information about real non-self conditions. Although the nearly-complete absence of false positives is already a great achievement in terms of algorithm reliability, the NSA formulation will be further tested on lab specimens so that self and non-self samples can be generated under controlled conditions and the influence of other parameters as the self radius and the number of training elements will be also assessed.

REFERENCES

[1] F. Gonzalez, D. Dasgupta, and R. Kozma, "Combining negative selection and classification techniques for anomaly detection," *Evolutionary Computation, 2002. CEC '02. Proceedings of the 2002 Congress on*, vol. **1**. pp. 705–710, 2002.

[2] T. Stibor, P. Mohr, J. Timmis, and C. Eckert, "Is Negative Selection Appropriate for Anomaly Detection?," in *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*, 2005, pp. 321–328.

[3] S. Forrest, A. S. Perelson, L. Allen, and R. Cherukuri, "Self-Nonself Discrimination in a Computer," in *Proceedings of the 1994 IEEE Symposium on Security and Privacy*, 1994, p. 202--.

[4] Z. Ji and D. Dasgupta, "Applicability Issues of the Real-valued Negative Selection Algorithms," in *Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation*, 2006, pp. 111–118.

[5] D. Dasgupta, S. Yu, and F. Nino, "Recent Advances in Artificial Immune Systems: Models and Applications," *Appl. Soft Comput.*, vol. **11**, no. 2, pp. 1574–1587, 2011.

[6] N. Bayar, S. Darmoul, S. Hajri-Gabouj, and H. Pierreval, "Fault detection, diagnosis and recovery using Artificial Immune Systems: A review," *Eng. Appl. Artif. Intell.*, vol. **46**,

no. Part A, pp. 43–57, 2015.

[7] C. Surace and K. Worden, “A negative selection approach to novelty detection in a changing environment,” in *Third European Workshop on Structural Health Monitoring,* 2006.

[8] C. Surace, G. Surace, and K. Worden, “A Negative Selection Approach to detect damage in aeronautical structures with changing Operating Conditions,” in *26th International Congress of the Aeronautical Sciences, Anchorage, USA*, 2008.

[9] M. Li and W. X. Ren, “Negative Selection Algorithm Using Natural Frequency for Novelty Detection under Temperature Variations,” *Adv. Mater. Res.*, vol. **163–167**, pp. 2747–2750, 2011.

[10] M.-G. Masciotta, J. C. A. Roque, L. F. Ramos, and P. B. Lourenço, “A multidisciplinary approach to assess the health state of heritage structures: The case study of the Church of Monastery of Jerónimos in Lisbon,” *Constr. Build. Mater.*, vol. **116**, pp. 169–187, 2016.

[11] P. B. Lourenço, K. J. Krakowiak, F. M. Fernandes, and L. F. Ramos, “Failure analysis of Monastery of Jerónimos, Lisbon: How to learn from sophisticated numerical models,” *Eng. Fail. Anal.*, vol. **14**, no. 2, pp. 280–300, 2007.

[12] M. F. Mohamad Mohsin, A. R. Hamdan, and A. Abu Bakar, “The Effect of Normalization for Real Value Negative Selection Algorithm BT - Soft Computing Applications and Intelligent Systems: Second International Multi-Conference on Artificial Intelligence Technology, M-CAIT 2013, Shah Alam, August 28-29, 2013. Pro,” S. A. Noah, A. Abdullah, H. Arshad, A. Abu Bakar, Z. A. Othman, S. Sahran, N. Omar, and Z. Othman, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 194–205.

[13] P. D’haeseleer, S. Forrest, and P. Helman, “An immunological approach to change detection: algorithms, analysis and implications,” in *Proceedings 1996 IEEE Symposium on Security and Privacy,* 1996, pp. 110–119.

[14] D. Li, S. Liu, and H. Zhang, “Negative selection algorithm with constant detectors for anomaly detection,” *Appl. Soft Comput.*, vol. **36**, pp. 618–632, 2015.

[15] K. J. Nurmela and P. R. J. Östergård, “Covering A Square With Up To 30 Equal Circles.” 2000.

[16] A. Heppes and H. Melissen, “Covering a Rectangle With Equal Circles,” *Period. Math. Hungarica,* vol. **34**, no. 1, pp. 65–81, 1997.

[17] M. C. Markót, “Robust Designs for Circle Coverings of a Square BT - Optimized Packings with Applications,” G. Fasano and J. D. Pintér, Eds. Cham: Springer International Publishing, 2015, pp. 225–242.

[18] F. A. González and D. Dasgupta, “Anomaly Detection Using Real-Valued Negative Selection,” *Genet. Program. Evolvable Mach.*, vol. **4**, no. 4, pp. 383–403, 2003.

[19] B. Adenso-Díaz and M. Laguna, “Fine-Tuning of Algorithms Using Fractional Experimental Designs and Local Search,” *Oper. Res.*, vol. **54**, no. 1, pp. 99–114, Feb. 2006.