Surrogate Modelling on Time of Flight Diffraction using Deep Learning

**Paialunga Piero, Corcoran Joseph**

1. Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati, OH 45221-0070, USA

2. Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati, OH 45221-0070, USA

\* Corresponding author, email@email

# Abstract

# Introduction

Stochatstic Modelling

A picture containing histogram

Description automatically generated

A picture containing building, window

Description automatically generated

Figure 1.1 (Upper Figure) A collection of multiple scans are plot vs the index datapoint. As it is possible to see, some areas (e.g. index around 400) are more uncertain than other (e.g. index around 1900). (Lower Figure) At a given index datapoint of the upper collection of scans, the signal is plotted against the temperature. It is possible to see that there is a clear trend between the temperature and the location, but the range can be very different (e.g. the range is 0.35 in location=400 and more than 0.50 in location=1970).

# Experiment and addition of damage

### Experimental Setup

The experiment that has been considered is a Guided Ultrasonic wave Tomography (GUWT) experiment (Simonetti & Alqaradawi, 2019).

In this experiment a carbon steel pipe has been monitored for 21 months at a rate of 22 datasets per day. 32 EMATs, 16 transmitters and 16 receivers, have been built in the original experiment.

A picture containing engine, gear

Description automatically generatedDiagram

Description automatically generated

Figure 2.1 (Left) The carbon steel pipe that has been monitored, (Right) (a) A picture of the carbon steel pipe and the SHM sensors that has been used, (b) a section of the pipe and the Transmitters (in red) and Receivers (green) used for the GUWT. In our experiment Transmitter 4 and Receiver 3 have been considered.

In our case, only a couple (Transmitter 4 and Receiver 3) has been studied. Nonetheless, the presence of different modes ([\*\*\*]) and the helicoidal paths make the signals complex and the damage detection problem challenging.

Each raw signal is made of 2080 data points. As the maximum absolute amplitude was large   
( ), and it could have led to numerical problems, the signals have been normalized.

Chart, line chart

Description automatically generated

**Figure 2.2** Example of a raw signal at a given temperature.

Due to the complexity of the signal and the different modes, the baseline signal stretch (BSS) method has been proven not to be useful in this context. [\*\*\*]

More information about the experiment details can be found in the original paper.

### 2.2 Addition of real damage

The original experiment focused on a corrosion damage. Nonetheless, the true extent of natural occurring corrosion was difficult to assess with independent measurements. For this reason, corrosion was prevented by covering the monitored section with a tarp sheet (Simonetti & Alqaradawi, 2019). After monitoring the pipe for approximately five months, a defect was introduced on the side of the elbow with an angle grinder.

The traditional method, namely Optimal Baseline Subtraction, has been used to detect the added defect (see Figure 2.3). As it is possible to see from Figure 2.3, the difference between the OBS baseline and the damaged signal is evident while the Non Damaged signal and the OBS are much more similar.

Chart, timeline

Description automatically generated

Figure 2.3 (Top) The damaged signal (red line) and the OBS selected baseline (blue line) are remarkably different. An apparent phase shift between the damaged signal and the OBS one can be seen. (Bottom) The non damaged signal (red line) is close to the OBS baseline (blue line). This example shows that the damage detection problem for the real damage can be solved using the OBS method.

In order to properly test the OBS method in the (real) damage detection problem, the True Positive Rates (TPR) and the False Positive Rates (FPR) are computed at different thresholds lines. Then, the Area Under Curve (AUC) has been computed.

As it is possible to see from Figure 2.4, the Area Under Curve is . This means that the OBS is perfectly able to distinguish a damaged signal and a non damaged one. As the traditional method works perfectly in this case, the GPR method has not been tested.

Shape, square

Description automatically generated

Figure 2.4 The True Positive Rate (TPR) and False Positive Rate (FPR) are computed at different values of thresholds (shown as dots in the plot). As it is possible to see, the TPR is while the . This means that the OBS method is working almost perfectly in terms of (real) damage detection.

### 2.3 Addition of synthetic damage

A more controlled and challenging situation has been considered by creating a synthetic damage and adding it to the non damaged signals. In particular, a toneburst has been added. The frequency of the toneburst has been chosen in order to match with the frequency of the signal. This leads to a toneburst that is ­­points long.

Different kind of damage has been considered. In fact, both the amplitude and the location of the damage has been changed. The dataset is divided in 33 areas from to . A value has been randomly picked from and , thus getting 33 quasi-random locations of damage.

Four different amplitudes have been considered as well (Figure 2.5):

* Small Damage: Amplitude = 5% of the maximum = 0.05
* Medium-Small Damage: Amplitude = 10% of the maximum = 0.10
* Medium-Large Damage: Amplitude = 15% of the maximum = 0.15
* Large Damage: Amplitude = 20% of the maximum = 0.20

A picture containing timeline

Description automatically generated

Figure 2.5 Different kinds of damage had been plotted at a given amplitude. The Damaged Signal and Non Damaged Signal are equals except for the location of the damage, where there is a localized change.

# Method

The idea of building an adaptive threshold is justified by the evident non-uniform variance of the dataset. In Figure 1.1 different signals are plotted in the same graph. It is possible to see that in certain zones (e.g. around 1500) the range of the signals is larger than in other ones (e.g. before 500).

This observation does not match with the assumption of the uniform variance that is used in the fixed threshold method of the OBS.

Another property that is evident is the dependency of the signal with respect to the temperature. Thus, it is possible to fit a polynomial curve that fits the data points at all the temperature of the dataset for a fixed index point (Mariani et al., 2020).

The approach that we are proposing suggests to use a Machine Learning algorithm, namely the Gaussian Process Regression (GPR) to fit the signal-temperature dependency and to find the non constant uncertainty of the signal. The final goal is to use this predicted uncertainty to classify a new signal as damaged or non damaged.

## 3.1 Kernel Function

Between all the infinite functions that can fit a set of points, the GPR method identifies the mean   
and variance of all these functions given a certain Kernel function (J. Wang, 2020).

The most general assumption has been made and the kernel that has been considered is the sum of the RBF (Radial Basis Function) kernel:

and the White Noise Kernel:

So that the Kernel that has been used is:

The White Noise Kernel parameter is used to incorporate the idea of a noisy observation inside the predictive model. On the other hand, the parameter in the RBF Kernel indicates how quickly the correlation relationship between two points drops as their distance increases (J. Wang, 2020).

## 3.2 GPR Generated Signal

Given a certain index point, we are able to fit the signal-temperature dependency using the GPR model. A GPR model is trained for each Index Data Point, thus obtaining 2080 GPR models. Different number of training set signals (which will be the same number of signals used in the OBS method) has been tested. More information about this test can be found in the result section.

Each trained GPR model gives, for a given temperature, a mean value and a variance value. The variance value can be used to set the boundaries of your data points at a given probability confidence. For example, corresponds to a 95% probability confidence. In other words 95% of the points are expected to be between mean - and mean + .

As is it possible to see from Figure 3.1, the data points’ distribution is not uniform: the majority of points is between 0 and 15 degrees, while we have less points from 20 to 30 degrees. As a RBF+White Noise kernel has been used, this leads to more uncertainty (larger boundaries) in the area where we have less datapoints (high temperature), and less uncertainty (smaller boundaries) in the area where we have more datapoints (small temperature) (J. Wang, 2020).

Text

Description automatically generated

Figure 3.1 Four different models (on four different index points) are displayed. For each index, a GPR mean, that is a function of the temperature, is predicted (blue line). A GPR variance, that is again a function of the temperature, is predicted as well. In the four images above, the GPR mean ± 3.00 the GPR variance is plotted (blue shade). This corresponds to a 99.7% confidence interval. The red points are the datapoints at different temperature for a given index. As it is possible to see, in these 4 examples all the points are inside the 99.7% confidence boundaries.

Combining all the index points at a given temperature, we can reconstruct the whole signal at a given temperature (Figure 3.2).

A group of flags

Description automatically generated with low confidence

Chart

Description automatically generated Background pattern

Description automatically generated

Figure 3.2 (Upper) A signal at a given temperature has been considered (blue line), the GPR predicted mean is shown with a red line, while the (95%) confidence boundaries are plot using the red shade. (Lower) Two different zones are highlighted in the lower figures. It is possible to see that the left figure boundaries follow the signal frequency. On the other hand, in a more uncertain region (around 1600) the boundaries are large everywhere as the predicted mean doesn’t accurately match with the signal one due to the large uncertainty of the signals in that area.

## 3.3 GPR Method

The trained models have been used to generate a bank ofGPR means and variances for an equispaced temperature range (Figure 3.3). Given a new signal the closest match, according to a specific criterion, has been chosen from the GPR generated bank of signals.

In the OBS method, the Mean Squared Error (MSE) is a typical criterion to select the optimal baseline (Croxford et al., 2010). Following this idea, the MSE criterion has been tested for our method as well, where the optimal baseline is now the GPR generated signal with lowest MSE between the new signal and the GPR predicted mean.

A picture containing text, screenshot

Description automatically generated

Figure 3.3 The GPR mean and variance can be virtually generated for all the values of temperatures in the real domain. In our case the whole temperature range has been uniformly sampled. Four values of the equispaced range of temperature are plotted. The GPR mean is plotted with a red line, while the 99.7% variance is plotted with a red shade.

Another method that has been tested is the Z score. Let’s consider a new signal , the -th GPR predicted mean of the baseline ] and the -th GPR predicted variance ], with , that is the length of each signal. The following quantity has been computed:

for

Where is the number of GPR generated signals. The selected signal is the one with the lowest value. Using the score we not only consider the GPR generated mean but the GPR generated variance as well, selecting the signal which is more likely to represent the new signal.

Four different methods, considering the four combinations of the two different criterions  
 (and the number of signals in the GPR generated bank, have been tested. More information about it can be found in the Result section.

Once the optimal baseline is selected, we classify a new signal as “damaged” if it goes out of the GPR predicted boundaries and “non damaged” if it is inside the GPR predicted ones. The idea that is behind this method is that there is a very low probability of having the signal in a certain location at a certain value if it goes out of the boundaries. For this reason, the signal can considered to be “damaged”.

If the signal goes out of the boundaries in the area where the damage has been synthetically added, we can considered the signal to be a true positive. If the signal goes out of the boundaries in the area where the damage has not been added, we considered the signal to be a false positive.

# A picture containing chart Description automatically generatedA picture containing chart Description automatically generated

Figure 3.4 (Left) The (Optimal) GPR Predicted Mean and Boundaires are plot using a blue line and a blue shade. On the other hand, the damaged signal is plotted using an orange line and the zone where tha damage has been added is the one between the two black lines. A close look of that area is plotted in the right Figure. In this case, the signal is correctly classified as damage as it goes out of the boundaries in the zone where it has been previously added (True positive) while it is inside the boundaries in the rest of the signal (not a false positive).

## 3.4 OBS Method for comparison

The OBS method has been implemented, using the Mean Squared Error as a metric for the baseline selection. Given a set of non damaged signal and a new signal, the signal with the lowest MSE has been considered to be the baseline. Then, the difference between the baseline and the new signal has been considered, and the Hilbert Transform of this difference has been computed. Using a fixed threshold, the new signal has been classified as “damaged” if it is larger than the threshold and “non damaged” if it is lower than the threshold. The whole process is shown in Figure 3.5.

In order to have a fair comparison, the same number of baselines in the OBS and of training set instances in the GPR has been used. More information about it can be found in the Result section.

Timeline

Description automatically generated

Figure 3.5 (Upper) The new signal is shown in blue and the optimal baseline, selected with the MSE method, is shown in red. The zone where the damage has been artificially added is shown in orange (Middle) The difference between the optimal baseline and the new signal is shown as a black line. The zone where the damage has been artificially added is shown in orange and it is clearly larger than the rest of the signal. (Lower) The Hilbert Transform of the difference is shown with a black line. The zone where the damage has been artificially added is shown in orange, and the fixed threshold is shown as a orange dotted line. This fixed threshold is able, in this case, to correctly identify the damage zone.

The limit of this method is that the Operational and Environmental Conditions can mask damage. In this case, OBS is uneffective, as it classifies the non damaged area points as damage (false positive) but it classifies the damage area points as non damage (false negative). An example of this situation is shown in Figure 3.6.

Timeline

Description automatically generated

Figure 3.6 (Upper) The new signal is shown in blue and the optimal baseline, selected with the MSE method, is shown in red. The zone where the damage has been artificially added is shown in orange (Middle) The difference between the optimal baseline and the new signal is shown as a black line. The zone where the damage has been artificially added is shown in orange. In this case, the Operational and Environmental conditions mask the damage, which is lower than other residuals. (Lower) The Hilbert Transform of the difference is shown with a black line. The zone where the damage has been artificially added is shown in orange, and the fixed threshold is shown as a red dotted line. In this case, we have different false positives points (red lines) while the damage is not mistakenly classified as non damage, as it is below the red dotted line.

# Results

The OBS method and the GPR have been tested for different entities of damage and different locations.

In order to obtain statistically relevant results, 300 test signals have been considered. These signals have been artificially “damaged” by applying a toneburst in different locations and with different amplitudes. In particular, four different amplitudes have been considered (Figure 2.5):

* Small Damage: Amplitude = 5% of the maximum = 0.05
* Medium-Small Damage: Amplitude = 10% of the maximum = 0.10
* Medium-Large Damage: Amplitude = 15% of the maximum = 0.15
* Large Damage: Amplitude = 20% of the maximum = 0.20

Each signal in the dataset is divided in 33 areas from to . A value has been randomly picked from and , thus getting 33 quasi-random locations of damage. At a given damage entity, each signal has been damaged in the 33 quasi random pre-selected locations and classified using OBS and GPR.

In the OBS, the fixed threshold value has been changed. In particular, 100 thresholds have been tested (from 0 to 0.3). At a given threshold, a signal could either be a false positive (if the damage is mistakenly identified in a non damage location), a true positive (if the damage is correctly identified in the damage location) or both. As this analysis is repeated for all the 300 signals at a given damage in a given location, a True Positive Rate (TPR) and a False Positive Rate (FPR) can be computed, thus generating a Receiving Operating Characteristic (ROC) curve for each location and each damage. This means that 33 curves (one per location) have been obtained for each damage entity.

A similar test procedure has been considered for the GPR method. In this case, the upper and lower bound have been changed. In particular, given a ­­ value that varies from 0 to 10, the lower bound has been considered to be GPR Mean GPR Variance, the upper bound has been considered to be GPR Mean GPR Variance. This formally represents a change of the confidence interval that it is considered. For example, by setting , the 95% confidence interval is considered, while with , the 99.7% confidence interval is considered. To have a fair comparison, 100 values from 0 to 10 have been considered.

An example of ROC curves plot for the OBS and GPR method is shown in Figure 4.1

# Chart Description automatically generatedChart Description automatically generated with medium confidence

Figure 4.1 (Left) The ROC Curves are plotted for the OBS method at a fixed damage entity (damage = 0.2 or 20% of the maximum). Each line represents the ROC curve at a specific location. The TPR and FPR rate change according to the change of the threshold (from 0 to 0.3). The colored lines represent the quartiles. (Right) The ROC Curves are plotted for the GPR method at a fixed damage entity (damage = 0.2 or 20% of the maximum). Each line represents the ROC curve at a specific location. The TPR and FPR rate change according to the change of the confidence intervals ( varies from 0 to 10). The colored lines represent the quartiles. As it is possible to see, the GPR ROC curves are closer to the perfect squares than the OBS ones.

Chart, box and whisker chart

Description automatically generatedChart, box and whisker chart

Description automatically generated

Figure 4.2 The boxplot of the OBS method (Left figure) and of the GPR method (Right figure) for different damage amplitudes are shown. While both the methods are largely ineffective for damage = 0.05, the GPR method has consistently larger AUC for all the other damage (0.1, 0.15 and 0.2).

Given the ROC curves for the two methods at different locations and damage, the Area Under Curve (AUC) has been considered as a metric of the two methods. The result of the AUCs for the two methods is shown in Figure 4.2

As it is possible to see, the GPR method’s AUC values are larger than the ones of the GPR for all the damage. While the AUC is still too low for Damage 0.05 (GPR Median AUC = 0.23 vs OBS Mean   
AUC = 0.18) it is considerably higher for the other damage, where the median of the AUC are respectively 0.45, 0.76 and 0.94 for the OBS method and 0.65, 0.93 and 0.99 for the OBS. Even in terms of the worst case scenario, it is possible to see that, except for the smallest damage (damage = 0.05 where the minimum and median AUC is close to 0 for both the methods), the minimum AUC values for the GPR method is always larger than the one of the OBS (see Figure 4.2). More information about worst and best cases, median and mean are found in Table 4.1.

The GPR model shown in Figure 4.2 and in Table 4.1 has been trained using   training set instances. The same number of baselines have been used for the OBS method. Moreover, the criterion selection that has been used is the MSE (, and the number of the baselines that have been produced is  .

Nonetheless, four different methods have been tested, considering the four possible combinations of the two different number of generated baselines ( and the two different criterions Each method has been tested using different training instances  
 (). For comparison, the same number of baselines have been used for the OBS model as well.

A summary of the four GPR methods is shown in Table 4.2, while the tests of these methods using different values of namely , is shown in Figure 4.3.

As it is possible to see, when , the OBS method outperforms all the proposed methods. On the other hand, when , all the proposed methods outperform the OBS. The most robust method is the one with and , which outperforms the OBS (in terms of the median values) for all the values.

Table 4.1 The OBS and GPR methods are compared considering their minimum, maximum, mean and median values of the 33 AUC values (one per location) at different damage entities. As it is possible to see, except for the minimum value of Damage=0.05, which is close to 0 for both the methods, the GPR method has higher values of AUC both in terms of the mean and the median than in terms of the outliers (minimum/maximum) values.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| OBS Method Statistics (AUC) | Damage = 0.05 | Damage = 0.10 | Damage=0.15 | Damage=0.20 |
| Minimum | 0.076 | 0.182 | 0.310 | 0.498 |
| Maximum | 0.352 | 0.794 | 0.983 | 0.9997 |
| Mean | 0.192 | 0.451 | 0.727 | 0.878 |
| Median | 0.183 | 0.446 | 0.758 | 0.935 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| GPR Method Statistics (AUC) | Damage = 0.05 | Damage = 0.10 | Damage=0.15 | Damage=0.20 |
| Minimum | 0.065 | 0.193 | 0.536 | 0.834 |
| Maximum | 0.501 | 0.969 | 0.997 | 0.9999 |
| Mean | 0.245 | 0.602 | 0.859 | 0.963 |
| Median | 0.227 | 0.653 | 0.927 | 0.989 |

Table 4.2 Four different methods have been tested producing different baselines and using one of the two criterions   
 or . The four possible combinations of the two baselines and the two criterions are shown in the table

|  |  |  |
| --- | --- | --- |
| Method |  |  |
| GPR Method 1 | 10 | MSE |
| GPR Method 2 | 10 | Z |
| GPR Method 3 | 100 | MSE |
| GPR Method 4 | 100 | Z |

The largest improvement is verified when for all the methods. As all the methods have similar performance, the method with ( and =MSE) has been used to produce the results shown in Figure 4.2 and Table 4.1. This has been done because the median of GPR Method 1 is slightly higher than the one of GPR Method 2 for , and even if both these methods performs worse (up to median AUC) than the correspondent ones with (GPR Method 3 and GPR Method 4), they require times the computational time.

Timeline

Description automatically generated

Figure 4.3 The OBS method (for comparison) is shown in the first plot, while 4 different boxplots (one per method of Table 4.2) are shown in the last 4 plots. Each boxplot represents the distribution of the average (on the 300 test signals) AUC values for the 33 quasi random located damage. The median and the minimum values of the distributions are displayed in the plots. The damage entity is, in this case, 0.15. As it is possible to see, when , the OBS method outperforms all the proposed methods. Nonetheless, when , the proposed methods outperform the OBS in at least one of their version. In particular, all the methods outperform the OBS when . The largest improvement, with respect to OBS and in terms of the Median AUC, is verified when for all the methods.

# 5. Discussion

The GPR method we proposed improved the classification performance of the OBS method as it has been proven in the result section where the AUCs of the two methods have been compared.

An example of a new signal analyzed using both the OBS and the GPR method is shown in Figure 5.1.   
As it is possible to see, the fixed threshold of the OBS method is not able to properly classify the signal. In fact, the other residuals, that are not due to the damage, are larger than the residuals in the damage location. On the other hand, using the 99.7% confidence boundary, the new signal is inside the confidence boundaries everywhere except for some points in the damage location.

(a) (b)

Timeline

Description automatically generatedA picture containing shape

Description automatically generated

Figure 5.1 The same signal has been analyzed using the OBS method (a) and the GPR one (b). It is possible to see that the OBS method is not able to properly classify the signal. In fact, using a fixed threshold (lower figure of the (a) column) the other residuals (orange line) are larger than the one of the damage zone (red line). Thus, the damage zone is either considered to be noise, or considered to be signal with multiple false positives (orange line). On the other hand, the same signal can be analyzed using the GPR method. In this case, we have that the signal goes out of boundaries in the damage zone (green line of the lower figure of the (b) column), while it stays inside (orange and red line of the upper figure of the (b) column) in the other points of the signal.

The overall improvements are displayed in the result section (Figure 4.2 and Table 4.1), which shows how the GPR method outperforms the OBS both in terms of the worst case scenario and in terms of median and mean values.

The differences between the proposed GPR and the OBS methods are:

1. The generation of a bank of GPR mean signals at given temperatures
2. The usage of a non fixed threshold (or confidence boundaries) associated with each GPR mean signal

To understand if the improvements of our method are due to the first or second point, the GPR method has been tested using a fixed threshold. Comparing the GPR Method with the Fixed Threshold and the GPR Method with the Non-Fixed boundaries (Figure 5.3), it has been shown that the improvement of our method is due to the usage of the confidence boundaries.

Chart, box and whisker chart

Description automatically generatedChart, box and whisker chart

Description automatically generated

Figure 5.2 The same signal has been analyzed using the OBS method (a) and the GPR one (b). It is possible to see that the OBS method is not able to properly classify the signal. In fact, using a fixed threshold (lower figure of the (a) column) the other residuals (orange line) are larger than the one of the damage zone (red line). Thus, the damage zone is either considered to be noise, or considered to be signal with multiple false positives (orange line). On the other hand, the same signal can be analyzed using the GPR method. In this case, we have that the signal goes out of boundaries in the damage zone (green line of the lower figure of the (b) column), while it stays inside (orange and red line of the upper figure of the (b) column) in the other points of the signal.

An intuitive view of the boundaries role is shown in Figure 5.3. In this example, a damaged signal has been considered and the optimal GPR mean and boundaries have been found using the MSE criterion.

The squared score has been computed for each point of the signal (see section 3.2). As it is possible to see, the score is larger in the damaged zone. It means that the “damaged zone” is correctly identified as an outlier of the process and thus as damage.

Graphical user interface, chart

Description automatically generated with medium confidence

Figure 5.3 (Upper) A damaged signal has been shown in red, except for its “damaged zone” which is in orange. The optimal the signal goes out of boundaries in the damaged zone. (Lower) The squared score is shown in the y axis. As it is possible to see, the value increases in the damaged zone, while it is consistently lower (in the non damaged zone.

As it has been shown in the result section (Figure 4.3) the limit of this method is that it works worse than the OBS method when the number of training set instances is small ( = 10).

When only a small amount of training set instances is considered, the GPR predicted boundaries and mean are not able to correctly fit the dependency between the signal and the temperature. In fact, as it is possible to see in Figure 5.4, the predicted mean and variances are not smooth enough.

(a) (b)

A picture containing chart

Description automatically generatedText

Description automatically generated

Figure 5.4 The signal-temperature dependency is shown in 4 different locations. As it is possible to see, the GPR mean and boundaries are much smoother in the (b) figure ( than in the (a) figure (

# 6. Conclusion

In this paper, the damage detection problem has been studied, where the damage has been considered as a localized change of the signal. The experiment that has been considered is a Guided Ultrasonic wave Tomography (GUWT) experiment (Simonetti & Alqaradawi, 2019). For numerical reasons, the signal has been normalized and the damage has been synthetically added with different amplitudes (0.05, 0.10, 0.15 and 0.20) and in different locations.

A Machine Learning supervised method, namely the Gaussian Process Regression (GPR), has been used to generate new mean signals and the relative confidence boundaries given input temperature’s values. Using these confidence boundaries, an adaptive threshold has been generated, that is lower when the signal is more certain and larger in the areas where the signal is more uncertain. This adaptive threshold is proposed as an alternative to the traditional Optimal Baseline Subtraction (OBS) method, which uses a fixed threshold to classify a new signal.

The two methods are compared using the Receiving Operating Characteristic (ROC) curves and the correspondent Area Under Curve’s (AUC) values for all the location and entity of the damage. The proposed method outperforms the traditional one as the AUC goes from a median value of 0.74 (OBS or traditional method) to a median value of 0.93 (GPR or proposed method) for damage of amplitude 0.15. Even in terms of the worst case scenario, the minimum value of the AUC for the traditional method is 0.31 while it is 0.56 in the proposed method.

The limit of the proposed method is the number of training set instances which are required to efficiently perform the GPR training. In fact, it has been shown that the traditional method outperforms the proposed one when the number of training set is not sufficiently large (.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Method**  **(AUC median value)** | **10** | **20** | **50** | **70** | **100** | **150** | **200** |
| **OBS** |  |  |  |  |  |  |  |
| **GPR Method 1** |  |  |  |  |  |  |  |
| **GPR Method 2** |  |  |  |  |  |  |  |
| **GPR Method 3** |  |  |  |  |  |  |  |
| **GPR Method 4** |  |  |  |  |  |  |  |

TEST TABLE!!!

# References

Croxford, A. J., Moll, J., Wilcox, P. D., & Michaels, J. E. (2010). Efficient temperature compensation strategies for guided wave structural health monitoring. *Ultrasonics*, *50*(4–5), 517–528. <https://doi.org/10.1016/j.ultras.2009.11.002>

Mariani, S., Heinlein, S., & Cawley, P. (2020). Location specific temperature compensation of guided wave signals in structural health monitoring. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, *67*(1), 146–157. <https://doi.org/10.1109/TUFFC.2019.2940451>

Simonetti, F., & Alqaradawi, M. Y. (2019). Guided ultrasonic wave tomography of a pipe bend exposed to environmental conditions: A long-term monitoring experiment. *NDT and E International*, *105*, 1–10. <https://doi.org/10.1016/j.ndteint.2019.04.010>

Wang, J. (2020). *An Intuitive Tutorial to Gaussian Processes Regression*. <http://arxiv.org/abs/2009.10862>

Wang, K., Zhang, J., Shen, Y., Karkera, B., Croxford, A. J., & Wilcox, P. D. (2021). Defect detection in guided wave signals using nonlinear autoregressive exogenous method. *Structural Health Monitoring*. <https://doi.org/10.1177/14759217211018698>