# An intelligent pattern recognition model to automate the categorization of pile damage

## Un modèle intelligent de reconnaissance de formes pour automatiser la catégorisation du dommage en pile

Silvia García
*Geotechnical Department, Instituto de Ingeniería UNAM, México, sgab@pumas.iingen.unam.mx*

Jessica Romero
*Graduate Programs Department, Benemérita Universidad Autónoma de Puebla, México*

Jorge López-Molina
*Rock Mechanics, Comisión Federal de Electricidad, México*

ABSTRACT: Deep foundations integrity testing is employed for assessing the soundness of in-place constructed elements. The intensified use of these structures but especially the latest evidence of their failure under loading situations not even under extreme levels, has resulted in an increased demand for quality control testing. In this paper an automatic identification routine that focuses on the recognition of patterns and regularities in data (small strain integrity tests) from piles is presented. Chaos theory, particularly Recurrence Plots RPs, is the alternative tool for exploration and interpretation of time series while Machine Learning is applied for providing reasonable answers for all possible inputs (waves from piles) and to perform "most likely" matching of these inputs, taking into account their statistical variation (detailed quantification of defects). With the use of RPs and automatic-intelligent analysis of their structures, the interpretation conflicts are solved and the uncertainties associated with criteria biased are reduced.

RÉSUMÉ : Des essais d'intégrité des fondations profondes sont utilisés pour évaluer la solidité des éléments construits sur place. L'utilisation intensive de ces structures, mais spécifiquement les dernières données de leur échec sous situations de chargement même pas sous des niveaux extrêmes, a entraîné une demande additionnelle de tests de contrôle de qualité. Dans cet article, on présente une routine d'identification automatique qui met l'accent sur la reconnaissance des formes et des régularités dans les données (petits essais d'intégrité des contraintes) des piles. La théorie du chaos, particulièrement les Graphiques de Récurrence GsR, est un outil alternatif d'exploration et d'interprétation des séries chronologiques tandis que l'Apprentissage Automatique est appliqué pour apporter des réponses raisonnables à toutes les entrées possibles (ondes provenant des piles) et pour effectuer l'appariement «le plus probable», compte tenu de leur variation statistique (quantification détaillée des défauts). Avec l'utilisation de GsR et l'analyse intelligente-automatique de leurs structures, les conflits d'interprétation sont résolus et les incertitudes associées aux critères biaisés sont réduites.

KEYWORDS: pile integrity test PIT, pile damage, pattern recognition, chaos theory, recurrence plots, time series analysis.

## 1 INTRODUCTION.

The increasing use of large diameter bored piles as deep foundation elements demands definitive and economical test procedures that can be used shortly after construction to assess their structural integrity. Low strain integrity testing requires the recording of the pile top acceleration caused by a blow of a hand-held impact device. Strains will be small but acceleration can be measured by very sensitive motion transducers. Large imperfections or irregularities in the pile may be detected and the wave velocity gives a relative indication of concrete quality.

The PIT test, sometimes referred to as the SIT, TNO, Sonic Integrity Test, Acoustic integrity test, TDR (Time domain Reflectometry) or Pulse Echo Test, is a non-destructive method to evaluate a pile foundation element in order to confirm its integrity (Bungenstaba and Beimb 2015). The test detects potentially dangerous defects such as major cracks, necking, soil inclusions or voids and, in some situations, it can determine unknown lengths of piles that support existing buildings.

Difficulties in record interpretation may arise from a lack of systematic investigations. Despite the increasing demand, its efficiency has become a controversial issue in the geotechnical community, due to methodology limitations and its applicability to different types of piles. In some cases, the collected force and velocity data has an unusual shape, making it hard or even impossible to be interpreted with adequate reliability (Likins 2015). In this investigation, an alternative process to interpret the characteristics of a pile foundation, is presented. The chaotic tool for analyzing the testing recordings is the Recurrence Plots RPs (Eckmann et al. 1986).

The RPs-methodology proposed here, robust and practical, uses the projection of the one-dimensional time series to the topological space of much higher dimensions, where behaviors and trends are discovered in the reconstructed facet. The chaotic configurations, manifestations of the excited pile foundation (heterogeneous system vibrating), are studied and codified in order to automatically-*intelligently* analyze their structures. Using machine learning concepts (classification trees), the interpretation conflicts are solved and the uncertainties associated with criteria biased are reduced.

## 2 RECURRENCE PLOTS

Recurrence is a fundamental property of dynamical systems, which can be exploited to characterise the system's behaviour in phase space. A powerful tool for their visualisation and analysis is called recurrence plot (Eckmann et al. 1986). RPs are intricate and visually appealing.

They are also useful for finding hidden correlations in highly complicated data. In this work the RP-analysis is extended, formalized, and systematized in a meaningful way that is based both in theory and experiments and that targets both quantitative and qualitative properties for its geotechnical application.

In this section, we briefly outline some of the basic features of RPs and describe how an RP of an experimental data set can be generated. The standard first step in this procedure is to reconstruct the dynamics by embedding the one-dimensional time series in a $d_E$-dimensional reconstruction space using the method of delay coordinates. Given a system whose topological dimension is d, the sampling of a single state variable is equivalent to projecting the d-dimensional phase-space dynamics down onto one axis.

Loosely speaking, embedding is akin to "unfolding" those dynamics, albeit on different axes (Packard et al, 1980; Takens, 1981). Given a trajectory in the embedded space, finally, an RP is constructed by computing the distance between every pair of points ($y_i$,$y_j$) using an appropriate norm and then shading each pixel (i,j) according to that distance. The process of constructing a correct embedding is the subject of a large body of literature and numerous heuristic algorithms and arguments. (Abarbanel 1995, Marwan et al. 2007) gives a good summary of this extremely active field.

### 2.1 *Delay Coordinate Embedding*

To reconstruct the dynamics, we begin with experimental data consisting of a time series:

$$\{x_1, x_2, \dots x_n\} \tag{1}$$

delay-coordinate reconstruction of the unobserved and possibly multi-dimensional phase space dynamics from this single observable $x$ is governed by two parameters, embedding dimension $d_E$ and time delay τ. The resultant trajectory in $R^{(dE)}$ is:

$$\{y_1, y_2, \dots y_m\} \tag{2}$$

where m=N-($d_E$-1)τ and

$$y_k = \left(x_k, x_{k+\tau}, x_{k+2\tau}, \dots x_{k+(d_E-1)\tau}\right) \tag{3}$$

for k=1,2,…,m. Note that using $d_E$=1 merely returns the original time series; one dimensional embedding is equivalent to not embedding at all.

Proper choice of $d_E$ and τ is critical to this type of phase-space reconstruction and must therefore be done wisely; only "correct" values of these two parameters yield embeddings that are guaranteed by the Takens Theorem (Takens, 1981; Packard et al., 1980; Sauer et al., 1991) to be topologically equivalent to the original (unobserved) phase-space dynamics.

Assuming that the delay-coordinate embedding has been correctly carried out, it is natural to assume that the RP of a reconstructed trajectory bears great similarity to an RP of the true dynamics.

Furthermore, we expect any properties of the reconstructed trajectory inferred from this RP to be true of the underlying system as well. This is, in fact, the rationale behind the standard procedure of embedding the data before constructing a RP.

### 2.2 *Constructing the Recurrence Plot.*

RPs are based upon the mutual distances between points on a trajectory, so the first step in their construction is to choose a norm D. In this work the maximum norm is used, although in one dimension the maximum norm is, of course, equivalent to the Euclidean p-norm. The maximum norm was chosen for two reasons: for ease of implementation and because the maximum distance arising in the recurrence calculations (the difference between the largest and smallest measurements in the time series) is independent of embedding dimension $d_E$ for this particular norm. This means that, direct comparisons between RPs generated using different values of $d_E$ without first having to re-scale the plots, can be made. The recurrence matrix A i defined as:

$$A(i,j) = D(y_i, y_j), \quad 1 \le i,j \le m \tag{3}$$

$$D(y_i, y_j) = \max_{1\le k\le d_E} \left|x_{i+(k-1)_\tau} - x_{j+(k-1)_\tau}\right| \tag{4}$$

The time series spans both ordinate and abscissa and each point (*i,j*) on the plane is shaded according to the distance between the two corresponding trajectory points $y_i$ and $y_j$ (Figure 1). The pixel lying at (*i,j*) is color-coded according to the distance. For instance, if the 117th point on the trajectory is 14 distance units away from the 9435th point, the pixel lying at (117, 9435) on the RP will be shaded with the color that corresponds to a spacing of 14.

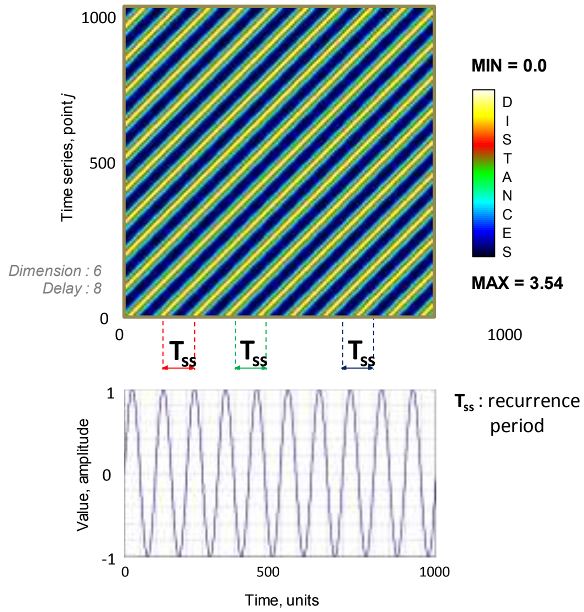


Figure 1. Graphical description of RPs: sinusoidal wave example; $T_{ss}$: natural period of sine or recurrence period of deterministic structures.

## 3 RECURRENCES AND INTEGRITY

It is important to state that for a consistent test interpretation achievement, the pile logs, soil borehole logs and all other available specifications of the pile foundation have to be clearly understood (PIT Manual 2005, Alonso 2013).

The inverse problem of taking a pile shape in a given soil profile and drawing the characteristics (the reflectogram) has not a unique solution. So it is important the additional information about the piling method, the soil environment and the supervisors' field notes. While testing can be learned fairly quickly by qualified personnel, interpretation should be left to geotechnical engineers with thorough knowledge of wave propagation theory, soil mechanics and piling techniques (Amir 2009). This is why the importance of developing an automatic pattern detection system that significantly minimizes errors due to lack of experience and contaminated data.

The analysis of a reflectogram is done by mentally comparing the graph to a catalogue of various pile characteristics and their respective reflectograms (Rausche et. al 1992). If a graph falls into one of these categories, it means that we can explain the significance of the data. If the opposite is true, another spot on the pile must be tested.

The first stage of the RPs-automatic recognition device proposed is on the basis of a qualitative interpretation. As the most basic attempt the records collected are divided into four categories:

- Category A – Clear indication of a sound pile shaft
- Category B – Clear indication of a serious defect
- Category C – Indication of a possibly defective pile shaft
- Category D – Inconclusive data

RPs calculated from a massive set of velocigrams are used to visualize the trajectories in phase space, which is especially advantageous in the case of high dimensional systems. RPs yield important insights into the evolution of the trajectories, because typical patterns in RPs can be linked to specific behaviors of the piles systems. The piles have quasi-chaotic recurrent structures (checkerboard) and abrupt changes in the dynamics. The evolution causes white areas and darker bands; the combination of vertical and horizontal lines forms rectangular *clusters* and also single, isolated dots are presented.

From RPs in which pile conditions are fully known (additional inspections, integrity post-verification) the recurrences in the plots are categorized. Firstly, data must be labeled as *meaningful* or *inconclusive*. Examples of typical reflectograms of Categories A to C and Category D are shown in Figure 2.

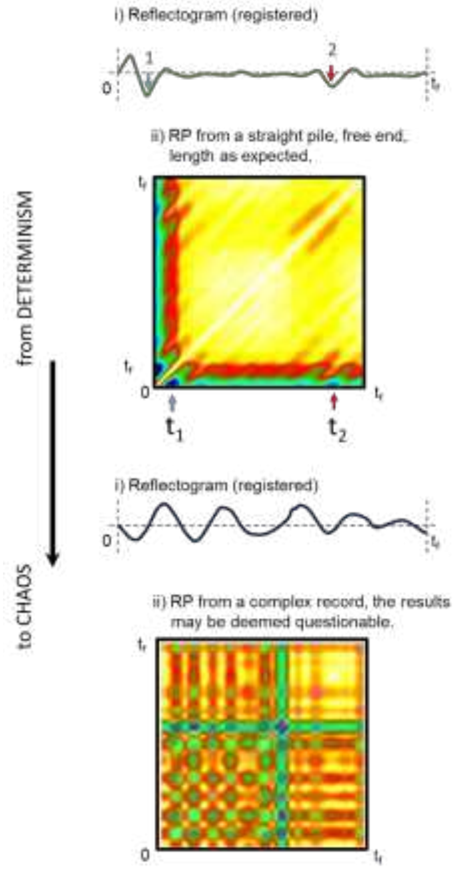


Figure 2. 1st Stage of classification of data: meaningful or inconclusive (deterministic to chaotic RPs-structures).

Well-defined checkerboard structures are indicator of inconclusive data, while sharp-abrupt changes in the dynamics of the pile (i.e. reflection from the toe, crack or necking, enlargement/bulb) that generate significant white areas, thin bands and spots in the RPs are considered meaningful. Darker dots (see Figure 3) are related to the sharp defined reflections attributed to impedance changes, whereas slower changing (green or red) is attributed to the soil. An impedance decrease resulting in a positive wave, usually means the presence of soft toe, while an impedance increase causes a negative wave and is considered as hard toe. Thus, necking or inclusions appears as a positive-negative cycle at the pile shaft.

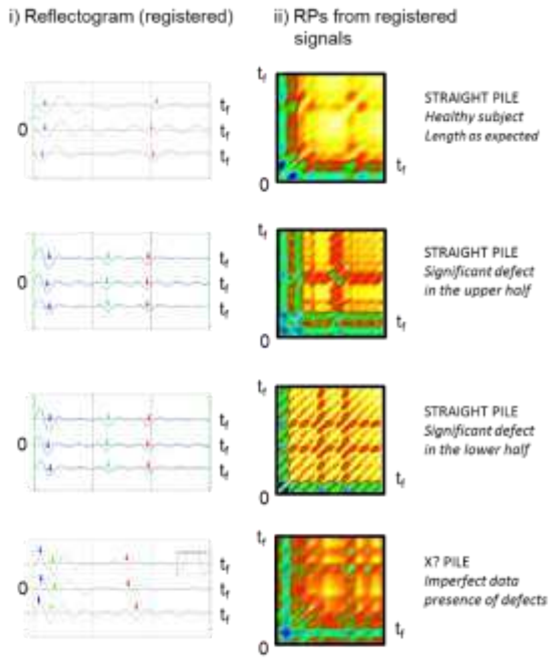


Figure 3. 2nd Stage of classification of data: taking in consideration impedance length and changes in geometry are detected.

The pile toe represents a reduction in pile impedance. Therefore, a tension wave is reflected from the toe, which is detected by the accelerometer and recorded as a single dark dot in the RP. Similar structures, but in different locations, are founded for a crack or necking in the pile: the trace will dip below then immediately rises above the zero line at the defect location. The initial dip is a characteristic response in pile impedance and occurs as the stress wave passes from the original into the reduced cross sectional area, so it is very clear as a single isolated structure in a RP. For an enlargement/bulb, the trace rises above the zero line and then immediately dips below. The dip is caused by a reflected tension wave which is generated by the relative decrease in impedance as the wave propagates out of the local increase in pile cross sectional area; this kind of behavior produces to orange-red fine bands.

A block diagram of the decision-theoretic pattern recognition system is shown in Figure 4. The area of image processing consists not only of coding, filtering, enhancement, and restoration of the reflectograms, but also analysis and recognition of images. On the other hand, the area of pattern recognition includes not only feature extraction and classification (texture in RPs), but also preprocessing and description of patterns. In this investigation the image processing considers only two-dimensional pictorial patterns but it must be linked to pattern recognition in order to deal with one-dimensional (the vector-reflectogram) and two-dimensional (the projection of the vector in 2D, the RP).

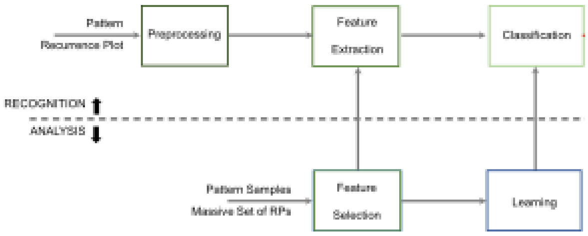


Figure 4. Block diagram of the decision-theoretical pattern recognition system proposed in this investigation.

Figures 5a and 5b show examples of recognition of recurrence points (isolated or bands) and the feature selection from these specific results. The two cases are piles of 1.0 m in diameter and 15.1 and 13.5 m in length, respectively. Since their length are less than 30 times the diameter, the empirical rules of the PIT are fulfilled and the velocigrams are useful.

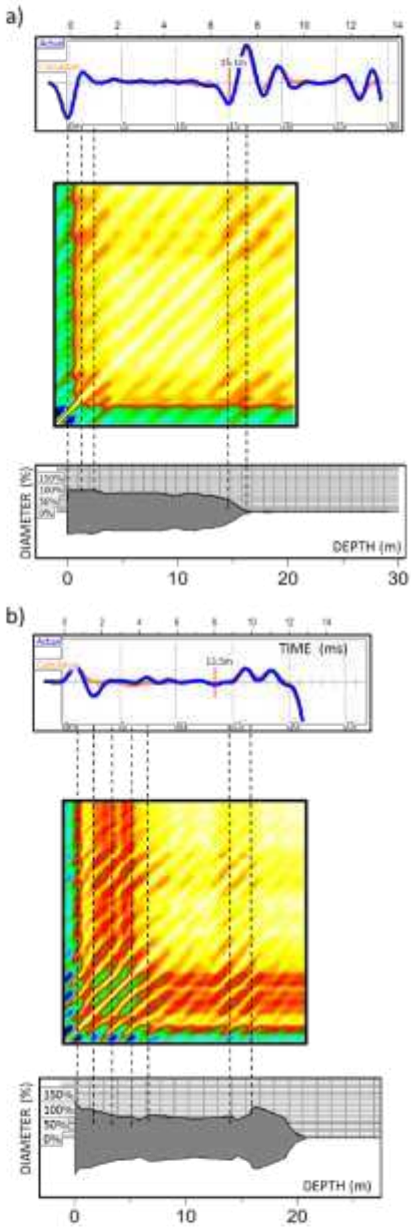


Figure 5. Recognition and analysis of RPs, a) healthy pile, b) defective pile.

The impedance profiles, the radiagrams as well as their RPs permit to show the exploratory evidences about the advantageous interpretation of the pile characteristics with RPs. Due to the quality of the structures and the definition of clusters, the velocigram was used successfully to qualify the integrity of the piles.

For the RP of Figure 5a, a healthy initial cluster is identified and the one generated for the toe is located at ≈15m, no more structures are detected (nor bands, nor dots), which allows to qualify the pile as healthy, of adequate length and without section problems.

On the contrary, in Figure 5b a series of clusters were identified and because of their size and the structures they form when intersect, they are labeled as anomalies-imperfections. Note that these details are more evident in the RP reading than in the time series (registered reflectogram). Some of the facts detected are: i) the length of the pile is not the expected (13.5 m), (ii) one possible defect is at ~ 4.5 m and it is about the diameter, and (iii) the quality of the base of the pile is doubtful, also the attributes of the concrete at the toe probably are not as in the rest of the foundation. In both cases, the conditions defined from the analysis of the RPs coincided with the *in situ* identification (additional tests and pile-extraction / section verification).

In the decision-theoretic approach, we are still looking for a more effective and efficient feature extraction and selection technique, particularly for nonparametric and small sample situations (strange or particular anomalies, external effects on the readings). The computational complexity of this pattern recognition system, in terms of time and memory, is at the present time subject for further investigation.

## 4 CONCLUSIONS

The proposed methodology is simple, based on accepted theory but refined from the experiences of engineers in the field. Its inclusion as an algorithm that enriches the testing-tools in the field is direct and affordable. Based on the presented results, the RPs-methodology can be very useful for rejecting or questioning piles with consistency and repeatability. Further testing and remedial actions can be programmed with lower expenses.

## 5 REFERENCES


Bungenstab F.C.I. and Beim, J.W. 2015. Piles – A Review of the Execution Process and Integrity Evaluation by Low Strain Test. *Proc. of the 15th Pan-American Conference on Soil Mechanics and Geotechnical Engineering (XV PCSMGE) Argentina.* 414-421.

Likins G.E. 2015. Pile Testing – State-of-the-Art. (Keynote Lecture). *Proc. Seminário de Engenharia de Fundações Especiais e Geotecnia: São Paulo, Brazil.*

Marwan N., Carmen-Romano M., Thiel M. and Kurths J. 2007. Recurrence plots for the analysis of complex systems. *Physics Reports* 438 (2007) 237–329

Sauer T., Yorke J.A., and Casdagli M., 1991, Embedology, *J. Stat. Phys.* 65(3/4), 579-616.

Packard N.H., Crutchfield J.P., Farmer J.D, and Shaw R.S. 1980. Geometry from a Time Series, *Phys. Rev. Lett*, 45- 9, 712-716.

Takens F. 1981. Detecting strange attractors in turbulence, *in: D. Rand, L.-S. Young (Eds.), Dynamical Systems and Turbulence, Lecture Notes in Mathematics*, 898, 366–381.

Eckmann J.P., Kamphorst S.O., and Ruelle D. 1987. Recurrence plots of dynamical systems. *Europhys. Lett.* 5, 973–977.

PIT Manual. 2005. *PIT-W Manual–Software for PDI's Pile Integrity Tester.* Pile Dynamics, USA.

Alonso U. R. 2013. Avaliação de Ensaios PIT nas Estacas Hélice Contínuas. ABEF (2007). Available in http://www.abef.org.br.

Rausche F., Likins G.E. and Kung S.R. 1992. Pile integrity testing and analysis. *Proc. 4rd Conf. Application of Stress Waves to Piles*, 613-617.