npj | materials degradation

Article

Published in partnership with CSCP and USTB

https://doi.org/10.1038/s41529-025-00594-7

Identifying stable pitting pathways in 316 L stainless steel via fractal-inspired PCA-based clustering

Check for updates

Leonardo Bertolucci Coelho[1,2,3,4] ✉, Thibaut Amand[3], Daniel Torres[1], Marjorie Olivier[3] & Jon Ustarroz[1,2]

This study introduces a fractal-inspired PCA-based framework to distinguish stable pitting pathways in 316 L stainless steel in chloride media. By transforming potentiodynamic polarisation data into a Mandelbrot space, the approach reveals two distinct scenarios of stable pitting growth: directly following passivity breakdown (case I) and preceded by metastable activity (case II). Clustering identifies critical pitting potentials (E_pit and E_sp) with high accuracy, effectively capturing rare metastability-driven events often overlooked in traditional analyses. The method demonstrates robust performance across varying chloride concentrations, with classification metrics highlighting its ability to detect low-frequency E_sp events. Results show that metastability-driven stable pitting (case II) occurs at higher activity levels and potentially at lower potentials. This work advances the understanding of the probabilistic nature of pitting and provides a scalable, data-driven strategy for predicting stable growth regimes.

Pitting corrosion is a localised and often unpredictable form of metallic degradation that can lead to catastrophic failures, particularly in chloride-containing environments. Under such conditions, the passive oxide film that normally protects metals like stainless steel can be compromised, effectively lowering the passivity breakdown potential and facilitating pit initiation[1]. Once pits begin to form, the transition from mere initiation to stable propagation regimes often determines the severity and long-term implications of the corrosion process.

A key parameter in assessing pitting susceptibility is the critical pitting potential (E_pit). Unlike a pit initiation potential, E_pit is traditionally associated with the threshold beyond which stable pits can propagate in a self-sustaining manner[2]. Another important descriptor is the metastable pitting potential (E_m), identified at the onset of the first significant current peak in a polarisation curve. This peak marks the formation of metastable pits - features that may or may not evolve into stable pits, as they can repassivate and disappear[3,4]. Although chloride ions lower the passivity breakdown potential in various ways[1], E_pit and E_m serve as crucial markers for understanding how a material transitions from a passive state to one featuring active, stable pit growth.

Despite the recognised importance of these parameters, the literature often treats stable pitting propagation as a single, uniform phenomenon. For example, it has been stated that "the early growth of pits which repassivate (metastable pits) and that of ones which continue to grow in a destructive fashion (stable pits') appear to be identical"[3]. However, our observations using classic potentiodynamic polarisation (PP) reveal that stable pitting can arise through two clearly distinguishable pathways: (i) stable growth occurring directly after passivity breakdown, and (ii) stable growth preceded by a phase of metastable activity. While it is likely true that all pits start out metastable[3], the conventional macro-scale polarisation resolution is often insufficient to detect metastability in the first scenario (case I), where stable growth follows passivity so quickly that no intervening metastable event is evident.

In contrast, the second scenario (case II) is characterised by stable growth emerging from an already active baseline established by metastable events. Although the literature acknowledges different types of metastable occurrences, such as transient current spikes or metastable growth phases[5], it does not explicitly distinguish these two stable-growth pathways. Indeed, documented transitions in stable pit growth rates[6] may correspond to cases where the defined E_pit is convoluted with metastability. Common definitions of E_pit - be they scientific ("E_pit is defined as a potential above which there is a rapid increase in the current on a polarisation curve"[7]) or normative ("the potential at which a sharp rise in current is observed"[8]) - implicitly assume the "classic" scenario (case I) as the general case.

This assumption overlooks case II, where stable pitting follows metastability and, consequently, can occur at higher activity levels. Figure 1 schematically illustrates both scenarios, highlighting that at the E_pit of case

[1]ChemSIN – Chemistry of Surfaces, Interfaces and Nanomaterials, Université libre de Bruxelles (ULB), Brussels, Belgium. [2]Research Group Electrochemical and Surface Engineering (SURF), Vrije Universiteit Brussel, Brussels, Belgium. [3]Materials Science Department, Faculty of Engineering, Place du Parc, University of Mons-UMONS, 20, Mons, Belgium. [4]IAlto Ltda, Taió, Santa Catarina, Brazil. ✉e-mail: leonardo.bertolucci.coelho@ulb.be

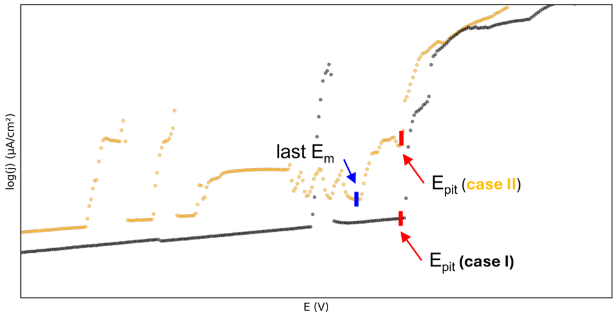


Fig. 1 | Schematic representation of the two stable pitting growth scenarios that occur at the critical pitting potential (E_pit) during anodic potentiodynamic polarisation. In case I (the "classic" scenario), stable pitting growth follows directly after a passive regime. In case II, stable growth (E_sp) is immediately preceded by a metastable event. Although metastable events can occur in both cases, the last metastable pitting potential (last E_m) is a key descriptor in case II, as it identifies the metastable pit that ultimately becomes stable (adapted from 2 examples of polarisation curves obtained with 316 L SS in 0.01 M NaCl, 0.5 mV/s scan rate).

II, the current density is considerably greater than that observed at the E_pit of case I. This higher level of activity makes sense because, in case II, stable growth starts from a surface already damaged by metastable pits. Frequent metastable events increase cumulative damage and enhance susceptibility within the passivity range, thus potentially accelerating the onset of stable growth[3,9–12]. Although a surface exhibiting case II behaviour would be readily identified as highly susceptible to localised corrosion, detecting and quantifying the exact E_pit in such a scenario is more challenging. It is considerably simpler to recognise a sudden current increase from a flat baseline (case I) than to detect an incremental rise from an already active baseline (case II).

A data-driven model, such as one trained via supervised machine learning techniques[13], could accurately detect E_pit values for both scenarios if properly labelled examples - representing straightforward (case I) and metastability-influenced (case II) stable growth - are available. Indeed, the large and diverse dataset employed in previous work[13], obtained at the micro-scale using Scanning Electrochemical Cell Microscopy (SECCM), likely contains numerous examples of case II stable growth. However, even if such regression models can estimate E_pit reliably for both scenarios, they alone cannot indicate which stable growth pathway is occurring.

Distinguishing between case I and case II stable pitting is critical for both corrosion prevention and monitoring. In case I, stable pitting occurs abruptly after passivity breakdown, making it easier to detect but harder to prevent due to its rapid onset. In contrast, case II involves stable pitting preceded by metastable activity, where the surface is already damaged by metastable pits, potentially leading to higher current densities. Identifying case II is particularly challenging yet essential, as these events potentially leading to catastrophic failure are more difficult to detect. To address this challenge, we employ an unsupervised clustering approach capable of automatically identifying and characterising rare events associated with metastable activity. This strategy reveals two distinct groups of critical pitting potentials without prior labelling, allowing us to map these clusters onto the observed stable pitting pathways and effectively "classify" stable pitting scenarios post hoc. By differentiating these pathways, we not only refine predictive models but also enable tailored monitoring strategies to mitigate the risks associated with localised corrosion.

While the current study focuses on 316 L stainless steel in chloride-containing environments, the fractal-inspired PCA-based clustering approach is broadly applicable to other passive alloys and aggressive media. By transforming PP data into a fractal-inspired Mandelbrot space, this framework provides a universal tool for identifying rare but critical events, such as the onset of stable pitting, across diverse material-environment combinations. This transformation not only facilitates the detection of these events but also reveals their multifractal nature, capturing the complexity of underlying electrochemical processes.

To achieve this, we employ Principal Component Analysis (PCA) within the Mandelbrot space, enabling the segmentation of data into distinct regions that correspond to different corrosion regimes. By focusing on the frequency of log(j) values rather than relying on arbitrary current density thresholds, this method enables the detection of the rarer, more harmful stable pitting scenario associated with metastability. In other words, our approach can identify and label the stable pitting potential in case II (E_sp) without presupposing simple current jumps. Conceptualising PP curves as multifractal distributions acknowledges their autocatalytic nature and scale invariance, thus revealing hidden corrosion regimes. Additionally, by uncovering rare but crucial events, fractal-based analysis aligns with the broader aim in electrochemistry to move beyond ensemble-averaged responses and capture the diversity of local conditions and transient phenomena[14–18].

In summary, this work leverages a fractal-based, data-driven clustering approach that, when coupled with post-hoc interpretation of the clusters, effectively classifies stable pitting pathways. By doing so, it refines our predictive capabilities for localised corrosion and provides a practical, generalisable framework for identifying and distinguishing between stable pitting growth scenarios.

Results

Populations of polarisation curves

Figure 2 shows the populations of PP curves obtained for the three investigated electrolytes (0.005 M, 0.01 M, and 0.05 M NaCl). The critical pitting potentials (E_pit or E_sp), as well as the passive potentials, were estimated using the machine learning (ML) model described previously[13]. Based on these estimates, the PP datasets were segmented into three corrosion regimes characteristic of a passive system: activity (yellow), passivity (green), and pitting (orange), as illustrated in Fig. 2a, d, g. The critical pitting potentials are displayed in red, with E_sp cases highlighted in panels (b), (e), and (h) and E_pit cases highlighted in panels (c), (f), and (i). The blue markers, referred to as "last E_m," were manually assigned and represent the last metastable pitting event observed prior to stable pit growth.

First, it is worth noting that the ML model[13] accurately estimated the critical pitting potentials in both E_pit and E_sp cases. This was expected since the present estimation task was less challenging than the one performed on the more diverse micro-scale datasets upon which the model was originally trained.

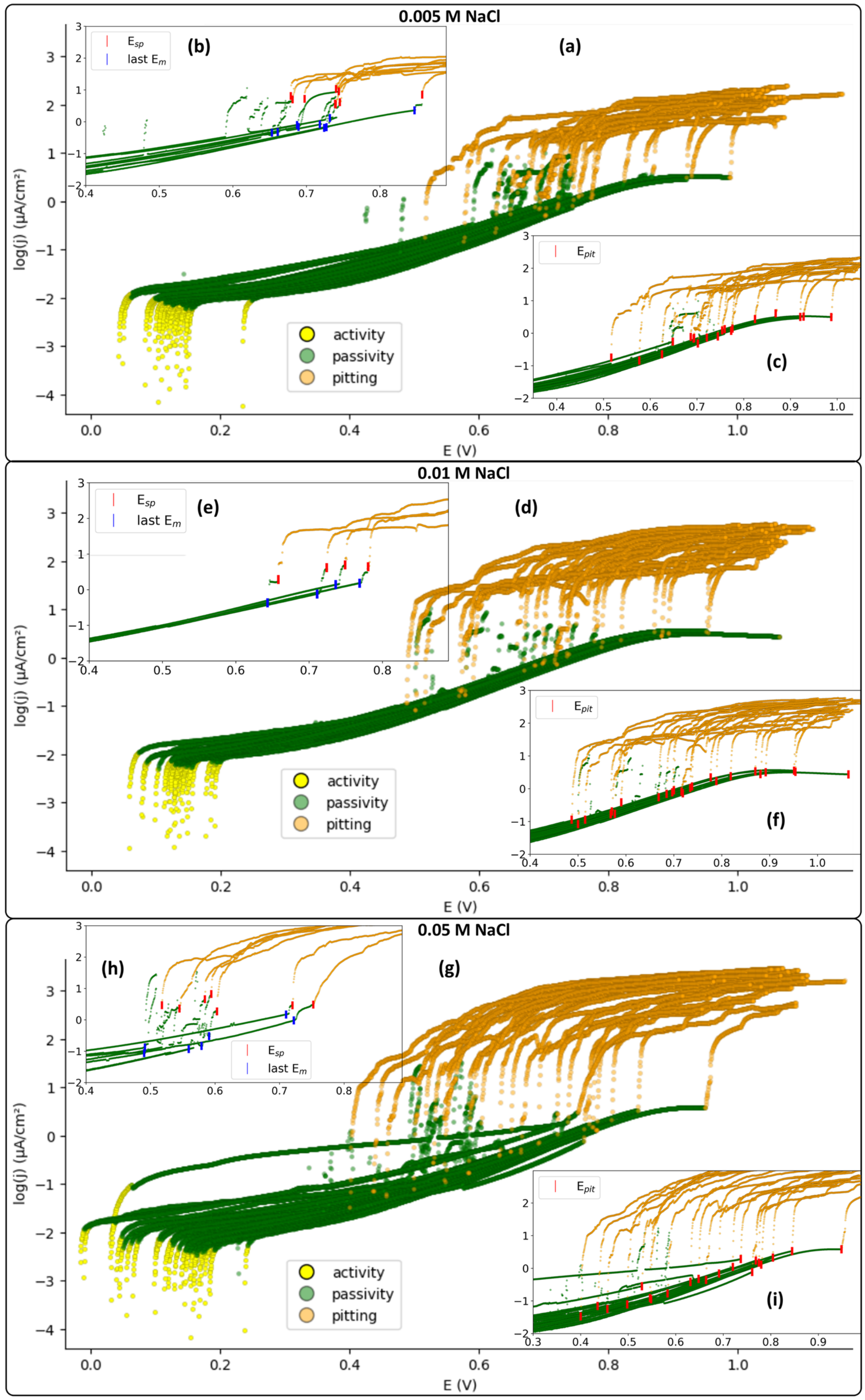
Second, the sub-populations displaying E_sp (Fig. 2b, e, h are fewer in number compared to those displaying E_pit (Fig. 2c, f, i. This observation is not surprising and may help explain why stable pitting potential is predominantly described as the E_pit case in the literature; it appears to be the more common scenario. While the numbers of E_sp events for 0.005 M, 0.01 M, and 0.05 M NaCl are 9, 4, and 7 respectively, no clear correlation emerges between E_sp occurrence and chloride concentration. Although this does not confirm the absence of a $[Cl^-]$ effect, increasing the dataset size - particularly the rare E_sp cases - would be needed for more robust statistical assessment.

As the NaCl concentration increases from 0.005 M to 0.01 M and 0.05 M, the spread of log(j) values along the potential axis broadens (Fig. 2a, d, g, especially within the pitting regime. This indicates greater variability in electrochemical activity at higher $[Cl^-]$ concentrations[19].

The observed variability in pitting potentials reflects the inherent probabilistic nature of pitting corrosion, influenced by factors such as material heterogeneities, passive film quality and local chemical conditions. This variability is particularly pronounced at higher chloride concentrations (0.05 M NaCl), where passive film destabilisation amplifies the range of breakdown potentials. A more comprehensive description of the log(j) distributions is provided in Supplementary Notes,2(Figs. S1 and S2).

Qualitatively, it appears that the amount of "passive" (green) data points exhibiting metastable pit events increases with $[Cl^-]$. Counting these

Fig. 2 | Populations of potentiodynamic polarisation curves obtained on 316 L SS immersed in NaCl media. **a**–**c** 0.005 M; **d**–**f** 0.01 M; and **g**–**i** 0.05 M NaCl. Markers are coloured according to the corrosion regime: activity (yellow), passivity (green), and pitting (orange). **a**, **d**, **g** show entire populations; inset plots **b**, **e**, **h** show sub-populations of curves displaying E_sp (stable pitting potential following metastability), with the last E_m points (final metastable pitting events) indicated; and inset plots **c**, **f**, **i** show sub-populations of curves displaying E_pit (stable pitting potential following passivity). For each NaCl concentration, the inset plots complement each other to fully reconstruct the complete population of curves shown in (**a**, **d**, **g**).



metastable occurrences is challenging due to the wide spectrum of intermediate pit states that lie between transient and fully metastable conditions, but the trend aligns with literature reports. Higher chloride concentrations increase the total number of active sites for metastable pitting on stainless steel[5,12,20]. Additionally, chloride ions hinder repassivation, thus decreasing the likelihood of healing and increasing the frequency and duration of metastable breakdowns [21,22].

As presented in Fig. 3a a remarkably high correlation ($R^2 \geq 0.94$) is observed between the last E_m and subsequent E_sp values, independent of the [Cl^-] concentration. This outcome is consistent with the understanding that the applied overpotential drives pitting processes[4,22–24]. Specifically, metastable pits that initiate at higher overpotentials are more likely to transition into stable growth at even higher potentials, whereas those nucleating at lower overpotentials tend to stabilise at lower E_sp values.

Histogram analysis

As illustrated in Fig. 4, at the lowest NaCl concentration (0.005 M), both E_pit and E_sp distributions are nearly unimodal, sharing a similar central

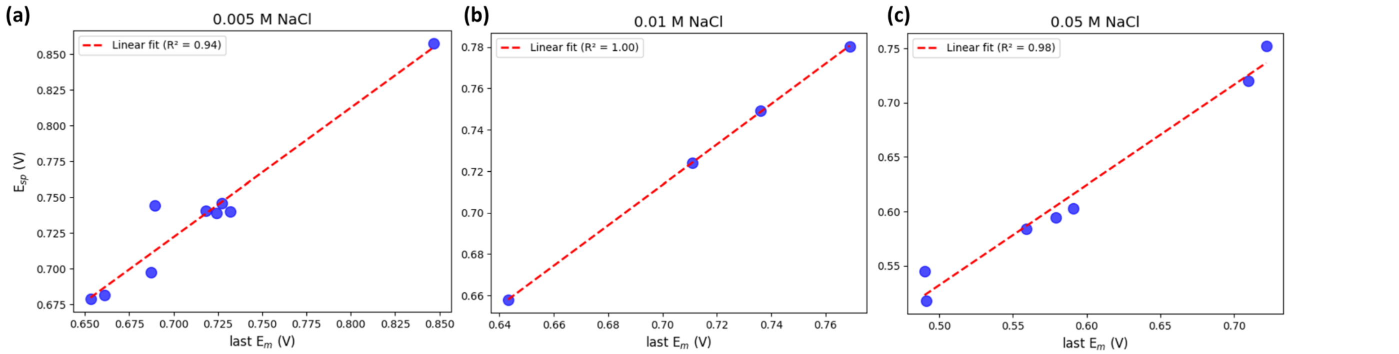


Fig. 3 | Correlation between the last E_m (last metastable pitting potential) and the critical pitting potential E_sp (stable pitting growth following metastability) for three NaCl concentrations. The data for each electrolyte is fitted with a linear regression line, and the corresponding R^2 (coefficient of determination) values are shown. **a** 0.005 M, **b** 0.01 M, and **c** 0.05 M NaCl.

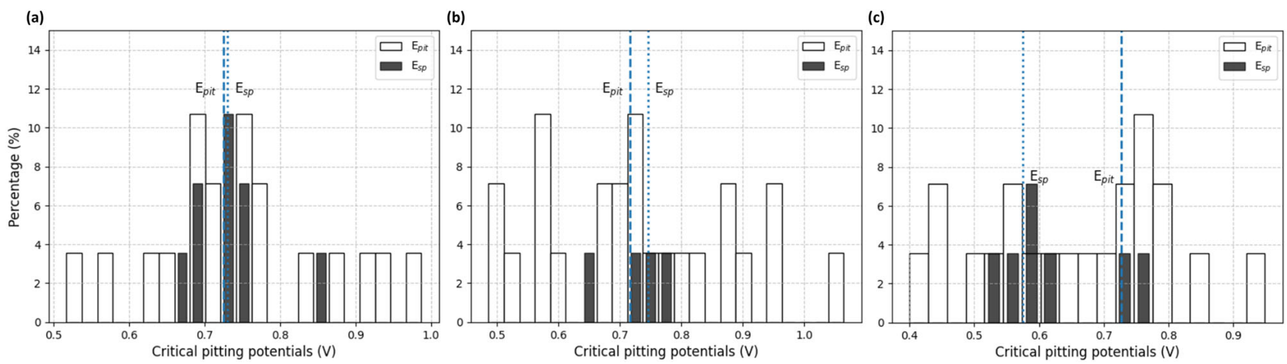


Fig. 4 | Distributions of the E_sp (stable pitting potential following metastability) and E_pit (stable pitting potential following passivity) features with central tendency estimations indicated. **a** 0.005 M, **b** 0.01 M, and **c** 0.05 M NaCl.

tendency (~0.73 V). At the highest concentration (0.05 M), while the central tendency of E_pit remains unchanged (~0.73 V), the E_sp distribution shifts notably downward (~0.57 V). Thus, at 0.05 M NaCl, stable growth preceded by metastable pitting tends to occur at lower overpotentials compared to stable growth preceded directly by passivity.

Previously, we noted that more data would be required to conclusively evaluate the influence of [Cl^-] on the probability of critical pitting transitioning through the metastable pathway. Nevertheless, the histograms at the highest chloride concentration suggest that increased [Cl^-] enhances the likelihood of stable growth occurring from a metastable state. On average, E_sp values at 0.05 M NaCl are ~0.16 V lower than at lower concentrations, reinforcing this hypothesis.

For didactic purposes, we consider two extreme scenarios: low and high [Cl^-]. Thus, the 0.005 M and 0.01 M NaCl cases are grouped as "low [Cl^-]", and the 0.05 M NaCl case is designated as "high [Cl^-]". As schematically illustrated in Fig. 5, under low [Cl^-] conditions (panel (a)), both types of stable pitting growth (E_pit and E_sp) tend to occur at similar potential values. In contrast, at high [Cl^-] (Fig. 5b), stable growth that follows metastability (E_sp) occurs, on average, at lower potentials than the growth that initiates directly from passivity breakdown (E_pit).

In the following section, we demonstrate how a PCA-based clustering Mandelbrot approach can accurately detect E_pit and, crucially, identify the rarer E_sp points.

Mandelbrot PCA modelling

Figure 6 exemplifies the workflow of the Mandelbrot PCA-based approach applied to the population of polarisation curves obtained in 0.005 M NaCl. In Fig. 6a, the anodic branch data are transformed into the Mandelbrot space, and each data point (defined by a Frequency-Rank pair of log(j)) is colour-mapped according to its average E value. For a complete description of the curves in Mandelbrot space, the reader is referred to Supplementary Notes 3 (Figs. S3, S4, and S5). The critical pitting potentials are indicated on the colourbar, and their respective types (E_pit or E_sp) are explicitly noted on the plot. Notably, within Mandelbrot space, the distributions of E_sp and E_pit points appear well separated.

In Fig. 6b, the Mandelbrot data are projected onto the first two principal components (PC1 and PC2), revealing a natural segmentation into five clusters: head, neck, core, tail, and tail's tip (further details on the clustering method are provided in Fig. S6). It is immediately evident that the tail's tip cluster contains all E_sp points, while the E_pit points are predominantly located within the tail and core clusters.

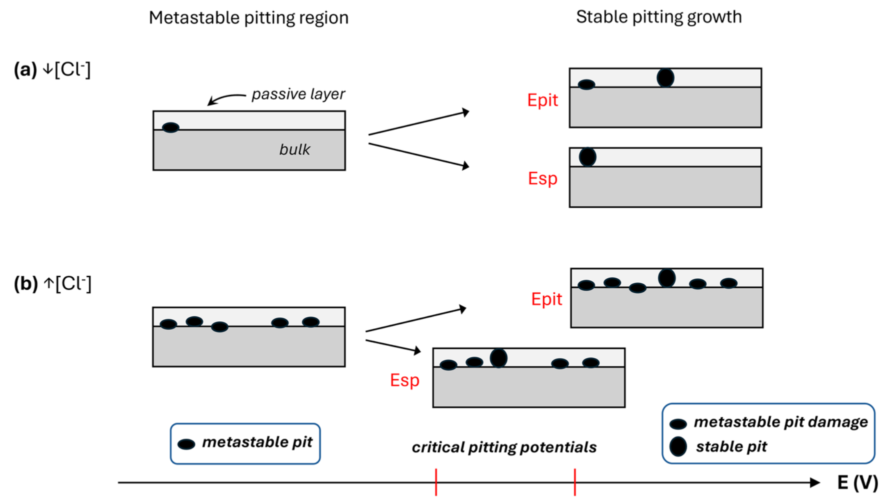
Figure 6c returns to the Mandelbrot space representation, now colouring the data points according to the clusters defined in Fig. 6b. Additional explanations regarding cluster assignments and curve-segment fitting can be found in Figs. S7 and S8. Finally, in Fig. 6d, these clusters are mapped back onto the original log(j) vs. E domain. This step clearly shows that the tail and tail's tip clusters respectively encompass most of the E_pit and E_sp distributions. More details on how the clusters influence the original data distributions are provided in Figs. S9, S10 and S11.

As shown in Fig. 7, the Mandelbrot approach efficiently assigns E_sp and E_pit points to distinct classes, corresponding to the "tail's tip" and "tail" clusters in the Mandelbrot space.

After the clusters were identified via the unsupervised approach, we assigned them to the E_pit or E_sp classes known from external information. This post-hoc mapping enabled us to compute classification metrics (Accuracy, Precision, Recall, and F1 Score) as a measure of how well the unsupervised clusters corresponded to the stable pitting classes of interest.

First, consider the original ("ground truth") distributions of both E_sp and E_pit classes (Fig. 7a, c, e). One can clearly see that E_sp points form a

Fig. 5 | Pathways to stable pitting growth as influenced by chloride concentration and applied potential. Schematic representation for **a** low (0.005, 0.01 M) and **b** high (0.05 M) [Cl⁻]. The transition to stable growth occurs at critical pitting potentials (indicated in red) and can be preceded either by passivity (E_pit) or metastability (E_sp). Metastable damage is shown as small black pits at the passive layer/substrate interface, while stable growth is depicted as larger black pits breaching the passive layer surface.



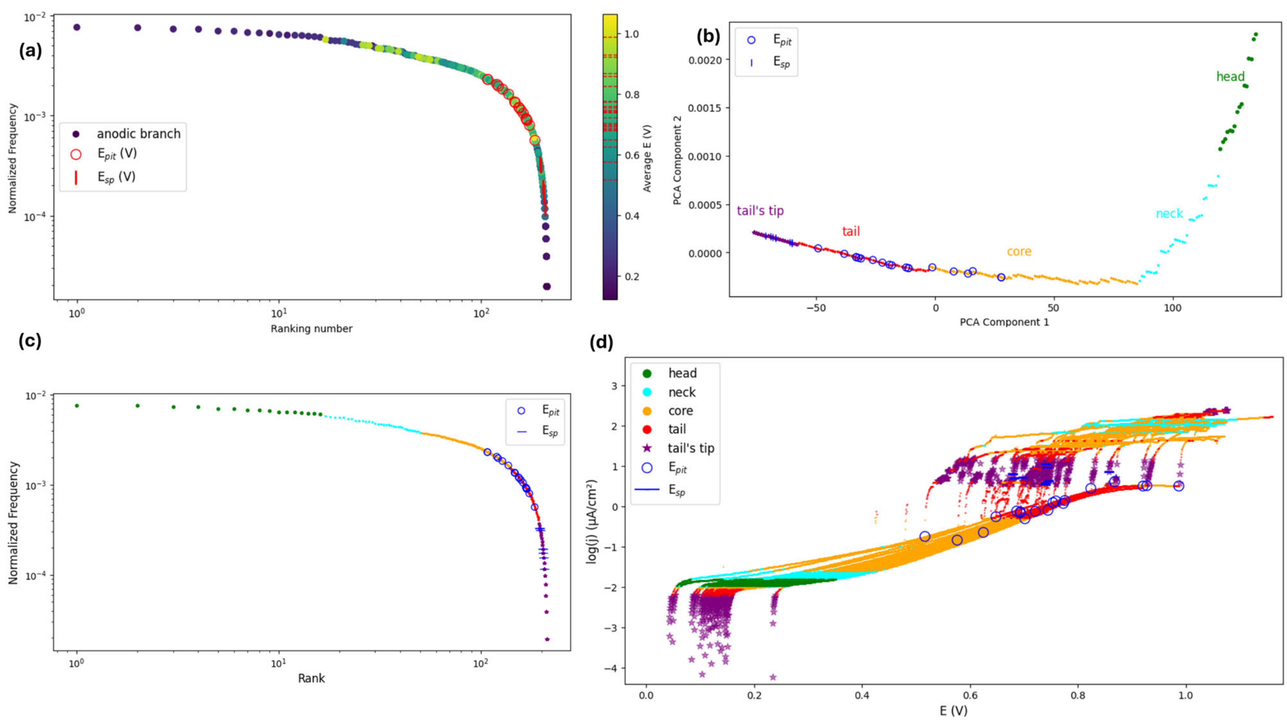


Fig. 6 | Workflow of the Mandelbrot PCA-based approach illustrated for the 0.005 M NaCl population of polarisation curves. E_pit (stable pitting potential following passivity) and E_sp (stable pitting potential following metastability) points are superimposed on all plots. **a** Mandelbrot plot (Rank-Frequency) of log(j), with data points colour-mapped to average E. Red markers on the colourbar indicate the critical pitting potentials, distinguished as E_pit or E_sp on the plot. **b** Mandelbrot data in the PC1-PC2 space (PCA Component 1 capturing maximum variance in the data), showing the defined clusters that form a "worm-like" structure with distinct regions (head, neck, core, tail, and tail's tip). **c** Mandelbrot plot coloured according to the clusters identified in (**b**). **d** Polarisation curves coloured by cluster assignment, illustrating how each cluster relates to E_pit and E_sp distributions.

distinct group in the log(j) vs. E domain, exhibiting significantly higher current densities than their E_pit counterparts. It is expected that the critical potential at which stable growth occurs from metastability (E_sp) would be associated with a higher level of activity than when stable growth directly follows passivity (E_pit). Moreover, the E_sp points show little to no correlation between the achieved log(j) and the corresponding potential, indicating that the high activity reached when a metastable pit transitions into a stable pit is not dependent on the applied potential. This finding aligns with established literature, where metastable and stable pit growth rates are diffusion-controlled and generally independent of electrode potential[4,20].

Examining the E_pit points in the ground truth classes (Fig. 7a, c, e) reveals that their distributions form relatively straight lines in the log(j) vs. E space. Linear regression fits exhibit strong correlations ($R^2 \geq 0.89$) with low associated errors (NRMSE ≤ 0.11), reflecting the near-baseline passive current densities. Although stainless steels ideally maintain a relatively constant passive current density over a range of potentials, in aggressive media, a slight upward slope can appear, suggesting incremental increases in dissolution and/or thinning of the passive film[1,25]. Notably, at high [Cl⁻] (0.05 M), the slope of the regression line (4.1 log(μA/cm²)/V) is larger than in the low [Cl⁻] media (3.2 and 3.3 log(μA/cm²)/V for 0.005 and 0.01 M, respectively), indicating greater anodic dissolution or passive film thinning at higher chloride concentrations.

At 0.005 M NaCl (Fig. 7a, b), the binary classifier perfectly identified all E_sp points; however, 5 E_pit points were incorrectly assigned to the "core"

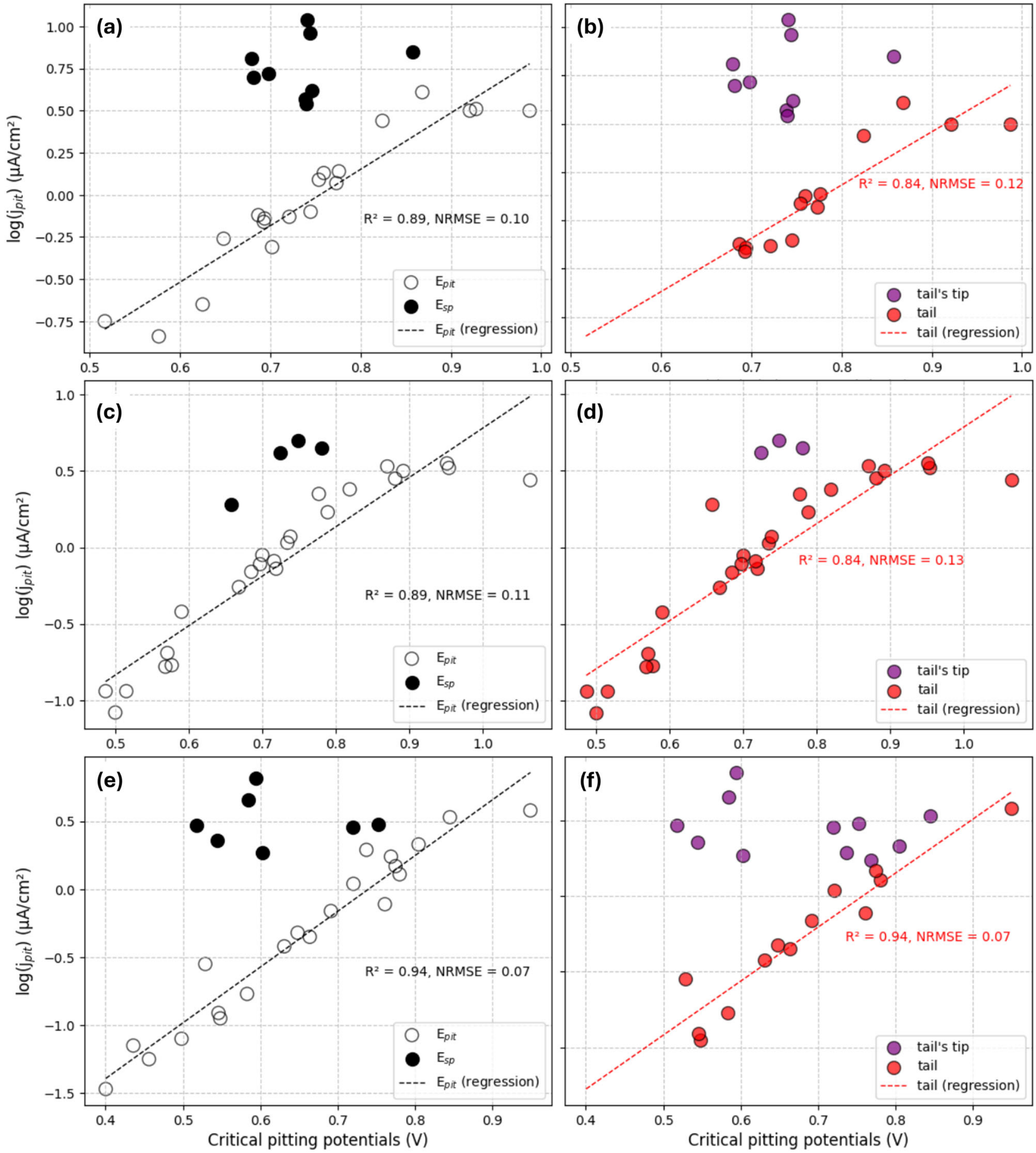


Fig. 7 | Correlation plots between critical pitting potentials (E_sp or E_pit) and the corresponding log(j_pit) values. a, **c**, **e** represent the ground truth classes of critical pitting, and **b**, **d**, **f** show the classes assigned by the clustering approach. **a**, **b** 0.005 M, **c**, **d** 0.01 M, and **e**, **f** 0.05 M NaCl. Regression lines for E_pit points highlight the intra-class relationship between potential and activity (with corresponding R^2 and NRMSE values). E_pit and E_sp are the stable pitting potentials following passivity and metastability, respectively; tail and tail's tip: distinct segments within the "worm-like" structure of the PCA-transformed Mandelbrot representation; R^2 coefficient of determination, NRMSE normalised root mean square error.

cluster instead of the "tail" cluster. At 0.01 M NaCl (Fig. 7c, d), the classifier was less precise in detecting E_sp points, likely due to their scarcity (only 4 events). Although all true E_pit points were correctly identified, one E_sp point was misclassified as E_pit. At 0.05 M NaCl (Fig. 7e, f), all E_sp points were correctly grouped, but 4 actual E_pit points were misclassified as E_sp, and an additional 4 E_pit points were not identified. These misclassifications not only degraded the regression fit (as evidenced by lower R^2 and higher NRMSE) but also highlight the sensitivity of the activity-potential correlations to classification errors.

Below (Table 1), the performance metrics of the binary classifier are presented for each electrolyte, separating the results for E_sp and E_pit classes. For clarity, accuracy reflects the overall proportion of correctly classified instances; precision indicates the fraction of positive (E_sp) predictions that are correct (minimising false positives); recall measures the

Table 1 | Classification performance of the Mandelbrot PCA-based method in determining the E_sp and E_pit classes for the three NaCl electrolytes

Electrolyte (NaCl)	Accuracy	Class	Precision	Recall	F1 Score
0.005 M	1.000	E_{pit}	1.000	1.000	1.000
		E_{sp}	1.000	1.000	1.000
0.01 M	0.964	E_{pit}	0.960	1.000	0.980
		E_{sp}	1.000	0.750	0.857
0.05 M	0.857	E_{pit}	1.000	0.810	0.895
		E_{sp}	0.636	1.000	0.778

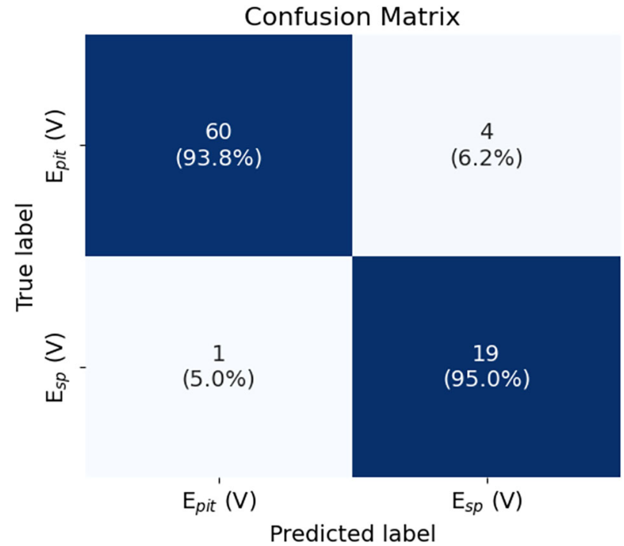


Fig. 8 | Confusion matrix summarising the overall performance of the binary classifier. The Y-axis indicates the true labels of the critical pitting potentials (E_pit and E_sp are the stable pitting potentials following passivity and metastability, respectively), while the X-axis shows the predicted labels (classifications determined by the Mandelbrot clustering approach). All metrics were computed using the aggregate dataset of 84 curves obtained from the three NaCl concentrations.

proportion of true E_{sp} events that are successfully detected (minimising false negatives); and the F1 score provides a balanced measure by harmonically combining precision and recall (see Supplementary Notes 4 for full definitions). For instance, in the 0.05 M NaCl dataset, the classifier achieved a recall of 1.000 for the E_sp class - indicating that all true E_sp events were correctly identified (i.e., no false negatives) - although some E_pit events were misclassified as E_sp, reflected in a precision of 0.636. High recall values for the critical E_sp class are particularly important in corrosion monitoring, ensuring that even rare metastability-driven pitting events are not overlooked.

In the confusion matrix shown in Fig. 8, the overall performance of the classifier is summarised, combining results from all three investigated electrolytes. The classifier successfully identified the sub-represented E_sp class in 19 out of 20 instances (95%), which was the primary objective. It also accurately identified the E_pit class in 60 out of 64 instances (93.8%).

Beyond these counts, the overall classification metrics confirm strong model performance. Specifically, the classifier achieved an accuracy of 0.941, a precision of 0.826, a recall of 0.950, and an F1 score of 0.884. These metrics indicate that the model correctly identifies most samples (high accuracy) and excels at detecting the less frequent E_sp class (high recall), underscoring its robustness.

Discussion

It is well documented that the critical pitting potential correlates with metastable pitting potentials in general terms. Specifically, as the frequency of metastable events increases, E_m tends to occur at lower potentials, thereby lowering the potential at which stable growth is triggered[3,9–12]. Although in this study a poor correlation (R^2 values between 0.02 and 0.2) was observed between the first metastable pitting potential (E_m) and the corresponding stable pitting potential, a strong relationship was confirmed when considering the last metastable pitting potential. Plotting these last E_m values against the corresponding E_sp values (see Fig. 3) clearly demonstrate this correlation.

The strong correlation between the last metastable pitting potential and E_sp highlights the critical role of chloride ions in facilitating metastable-to-stable transitions. Chloride ions not only initiate metastable events by destabilising the passive film but also create a corrosive microenvironment that sustains pit growth[26]. At higher chloride concentrations, the increased frequency of metastable events leads to greater cumulative damage, lowering the potential at which stable growth occurs. This phenomenon aligns with Williams's pitting initiation model [21,27], which advances that higher chloride levels enhance the nucleation rates of both unstable and propagating pits.

In contrast, when examining the corresponding current densities, only a low correlation (~0.29 on average) was found between log(j_m) and log(j_sp) (Supplementary Figs. 1,2). Here, j_m corresponds to the current density at the last metastable potential, and j_sp corresponds to the current density at the stable pitting potential. Although log(j_m) is influenced by the passive-state activity baseline, the actual current density reached at E_sp is less predictable. This variability likely arises from the diverse activity pathways that may occur between the last E_m and E_sp points (Fig. 2b, e, h, which may include diffusion-controlled processes and localised chemical changes within the pit environment[2].

It remains challenging to determine from the present data whether chloride directly affects the initiation potential of metastable pits (Fig. 2), even though the overall metastable activity appears to increase with [Cl^-]. Moreover, the observed drop in E_sp at 0.05 M NaCl (Fig. 4) cannot be solely attributed to increased metastability, as we did not observe a clear rise in the frequency of Type II pitting events directly correlated with chloride concentration Fig. 2b, e, h.

Roughly, metastable events occur above ~0.6 V at 0.005 M, around ~0.6 V at 0.01 M, and below ~0.6 V at 0.05 M NaCl, indicating a shift toward lower potentials with increasing chloride concentration (Fig. 2). While this observation aligns qualitatively with the noted increase in metastable activity at higher [Cl^-], it does not conclusively establish a direct effect of chloride on the initiation potential of metastable pits[28,29]. The greater likelihood for stable pitting to occur at lower E_sp values under higher [Cl^-] likely reflects the broader distribution of polarisation curves observed at 0.05 M NaCl (Fig. 2g–i).

Since chloride lowers the onset potential of metastable events[3,9–12] and increases their likelihood[21,22], the drop in E_sp at 0.05 M NaCl (Fig. 4c) is statistically plausible. In other words, if chloride promotes more frequent metastable pit nucleation and impedes repassivation, it is reasonable to expect E_sp events to manifest, on average, at lower potentials.

This reasoning aligns with Williams's pitting initiation model[21,27], which postulates a proportional relationship between the nucleation rates of unstable pits and propagating pits[30]. Since the model does not differentiate between the two types of stable propagation (E_pit and E_sp), it implies a general correlation between the frequency of metastable events and the likelihood of both types of stable pitting.

Our rationale on the effect of increasing [Cl-] aligns with Williams's model, which explains the enhanced likelihood of metastable pitting and the observed reduction in overall critical pitting potential (which effectively represents the average of the observed E_pit and E_sp values). Notably, Fig. 4 shows that E_pit remains constant irrespective of [Cl^-]. Comparing this observation with Williams's foundational assumption is insightful: "the model assumes that pit nucleation is a stochastic process, and that subsequent growth is deterministic"[30] ("stochastic" here refers to the temporal randomness of metastable initiation events[27]). When chloride concentration was considered[21], the effects align with expectations: higher [Cl⁻] leads to more frequent metastable events. Consequently, if metastable nucleation is stochastic in time and chloride-dependent, yet stable pit growth following these events is deterministic, then higher [Cl⁻] should decrease the E_sp, as our results suggest.

Although E_pit appears relatively insensitive to chloride concentration, E_sp exhibits a clear dependence on [Cl⁻], occurring at lower potentials as chloride levels increase. This behaviour can be attributed to the dual role of chloride ions in both initiating and sustaining metastable pits. Chloride ions lower the energy barrier for pit nucleation by destabilising the passive oxide film, thereby increasing the frequency of metastable events. Additionally, they hinder repassivation by maintaining an aggressive local chemistry within metastable pits, facilitating their transition to stable growth[22,26,31]. Future work should explore the interplay between chloride concentration, pit geometry, and local pH changes to better understand the physical mechanisms driving metastable-to-stable transitions.

Moreover, our analysis is based on a relatively modest dataset (28 measurements per concentration, totalling 84 curves) and a limited chloride concentration range (approximately one log-decade, from 0.005 to 0.05 M NaCl). These constraints may not fully capture more subtle shifts in E_pit that could emerge over wider parameter spaces. Further experiments with larger datasets and broader chloride ranges would be needed before drawing definitive conclusions on the chloride effect on E_pit. These observations stress the importance of E_sp. Since E_sp can occur at lower potentials than E_pit, it may represent a more easily triggered form of stable pitting, potentially overlooked by models that consider only a sharp transition from passivity to stable growth[7,8]. Historically, studies have defined the critical pitting potential (E_pit) as a single parameter[21,27,30], encompassing both possible stable pitting pathways. By averaging these two pathways, conventional models would fail to account for the lower E_sp values observed in our study, potentially underestimating the risk posed by metastability-driven stable pitting.

The classification results demonstrate strong to exceptional performance, depending on the chloride concentration. For the low [Cl⁻] case at 0.005 M NaCl, the classifier achieves perfect scores across all metrics, indicating outstanding classification accuracy. At 0.01 M NaCl, the overall accuracy remains high (96.43%), with strong F1 scores, although a slight reduction in recall is noted for the E_sp (tail's tip) class. Under high [Cl⁻] conditions (0.05 M NaCl), precision remains reliable but recall for the E_sp class is more challenging, resulting in a moderate F1 score. These findings highlight the robustness of the classifier at lower chloride concentrations while suggesting that improved strategies may be needed to handle the more aggressive (high [Cl⁻]) scenarios.

The accurate identification of E_sp and E_pit has significant implications for corrosion prevention and mitigation strategies. For instance, knowing the range of potentials at which metastable-to-stable transitions occur allows engineers to define safer operating conditions for materials in aggressive environments. In industries such as oil and gas or marine infrastructure, where chloride-induced pitting is a major concern, this information can guide the selection of protective coatings, inhibitors, or cathodic protection systems. Furthermore, the ability to distinguish between case I and case II stable pitting pathways enables targeted interventions: surfaces exhibiting frequent metastable activity (case II) may require more stringent monitoring or surface treatments to mitigate cumulative damage, while those prone to direct passivity breakdown (case I) may benefit from alloy modifications to enhance passive film stability.

A remaining question is whether determining the E_sp distribution of a given dataset (as performed here through manual labelling) would suffice to establish a suitable working potential range with appropriate margins of error for a material. The answer is no: acquiring multiple datasets under evolving conditions quickly becomes impractical, highlighting the need for scalable, data-driven methodologies. To address this, we propose the following Mandelbrot PCA-based approach: by converting the entire anodic branch (log(j) values) into the Mandelbrot space, the separation between rare E_pit events and even rarer E_sp events becomes more apparent. Then, applying PCA to the Mandelbrot-transformed data reveals discontinuities that facilitate clustering. Preliminary results suggest that this clustering approach, when extended to micro-scale polarisation curves, is highly reproducible. Mapping these clusters back onto the original polarisation curves confirms a strong correspondence between E_pit and E_sp classes and two specific clusters. Notably, the cluster associated with the rarest log(j) values ("tail's tip") captures 95% of the true E_sp points. These findings indicate that employing Mandelbrot and PCA transformations enables faster, automated determination of a safe potential range, which can be dynamically updated as new datasets are acquired under changing conditions.

Why begin with Mandelbrot modelling? Although applying data reduction methods directly to polarisation curves might suffice for clustering, starting with Mandelbrot provides a probabilistic framework based on a theoretical law that is both well-suited for detecting rare events and straightforward to fit. By enabling a quantitative description of the entire distribution, Mandelbrot modelling allows us to accurately determine the frequency of even the rarest occurrences. For instance, in the present case, the rarest E_sp events occur at frequencies below 0.0003 of the total log(j) values (Fig. 6c). For further probabilistic insights in other media, please refer to Figs. S3 and S4. Mandelbrot modelling may thus offer a probabilistic perspective that could extend to other frequency-based analyses, including wavelet spectrograms and Electrochemical Impedance Spectroscopy.

While the current dataset is limited in size and chloride concentration range, it provides robust insights into the influence of [Cl⁻] on pitting behaviour, with ongoing and planned efforts to expand these findings to broader parameter spaces, including wider chloride concentration ranges.

Looking ahead, an ideal pitting predictive tool would focus on the earliest possible detection of pitting, such as identifying the final metastable event before stable growth. Future work will aim to develop automated methods for detecting this critical last E_m point.

Building upon these foundational findings, our current research is extending the fractal analysis to micro-scale polarisation curves obtained via local electrochemical scanning techniques (SECCM). These larger, more complex datasets - characterised by higher variance due to finer spatial resolution - promise to further elucidate the probabilistic interplay between passivity and pitting initiation and growth at smaller scales.

Finally, bridging the gap between micro- and macro-scales through a multifractal lens holds promise for capturing the diversity and frequency of underlying phenomena at multiple scales. While the current work already demonstrates the transformative potential of fractal-based methodologies in analysing and predicting localised corrosion processes at the macro-scale, future multiscale analyses are expected to provide even deeper insights into the mechanisms governing passivity breakdown and stable pitting growth.

In summary, our fractal-inspired PCA-based clustering framework not only enables the rapid, automated identification of distinct stable pitting regimes at the macro-scale but also lays the groundwork for scalable, real-time corrosion monitoring strategies for passive metals and alloys.

Methods

Material

The selected passive alloy was the 316 L stainless steel produced by Aperam (France) and surface treated by Packo Inox (Belgium). The industrial surface treatment consists of electropolishing (bath temperature = 65 °C, current densities between 20 and 40 A/dm²) followed by passivation and water rinsing[32]. Squared 316 L SS specimens of ~5 × 5 cm² were cut from the same plate (same industrial batch) with a lever cutter and degreased in an acetone bath under ultrasound for 10 min (followed by rinsing in demineralised water). The resulting surface finishes have been thoroughly investigated previously[32], demonstrating high homogeneity and reproducibility at the microscale.

Polarisation curves

The PP tests were performed at varying NaCl concentrations (0.005 M, 0.01 M and 0.05 M) chosen to match the electrolytes of our Micro PP datasets[13]. Using an SP-300 (Bio-logic) potentiostat, the electrochemical cell (inside a Faraday cage) comprised the stainless steel specimen as the working electrode (WE, ~1 cm² of exposed area), an Ag/AgCl/KClsat as the reference electrode (inside a Luggin capillary) and a platinum foil as the counter electrode (CE). The electrochemical cell contained a large volume of electrolyte (150 mL), ensuring that the bulk pH remained stable despite localised changes due to pitting. All experiments were conducted at room temperature (23 °C ± 2 °C).

The polarisation tests started from − 30 mV to +900 mV (Vs OCP) after 60 min of immersion to guarantee relatively stable OCP values. The measurements were repeated 28 times for each NaCl electrolyte concentration, generating a complete dataset of 84 polarisation curves (the dataset can be integrally downloaded at https://doi.org/10.5281/zenodo.15284154)[33]. A scan rate of 0.5 mV/s was chosen to match our previous Macro PP study on the same 316 L SS sample[32].

Data processing

Data preprocessing was an integral aspect of our study, aimed at refining the datasets to achieve optimal analytical granularity. This process involved evaluating signal-to-noise ratios to ensure the integrity of the overall behavioural patterns was maintained while avoiding underfitting and overfitting. Our approach is analogous to the binning strategy commonly employed in histogram analysis, where bin sizes are carefully selected to balance detail and clarity. For this reason, we adjusted the precision of the log(j) values, setting it to 2 digits. To determine the optimal precision for log(j) values, we conducted a preliminary sensitivity analysis comparing 1-, 2-, 3- and 4-decimal resolutions. One decimal masked important features, while three and four decimals introduced excessive noise or overfitting, making two decimals the best compromise. Adjusting the precision to two digits ensured that the resulting data representation retained essential detail without preserving excessive noise, thereby maintaining the meaningful variance intrinsic to each dataset.

To facilitate comparability across diverse datasets, the frequency axis was normalised for all distributions, whether depicted in traditional histograms or Mandelbrot plots. This normalisation converts the frequencies so that the sum of all data points equalled unity, standardising the datasets for subsequent comparative analysis.

The entire code for data analysis and visualisation was developed in Python 3.8 and is available on GitHub https://github.com/bcoelholeonardo/pitting_pathways_pca_clustering. Various Python libraries were used throughout the data processing and modelling workflow (*Numpy*, *pandas*, *Matplotlib*, *sklearn.decomposition*, *sklearn.metrics*, *sklearn.cluster*, *scipy.optimise*).

Mandelbrot-Zipf modelling

Our model uses the Mandelbrot-Zipf introduced by Mandelbrot arguments on the fractal structure of lexical trees[34]. Let us identify a particular log(j) value by an index i equal to its rank, where 'rank' refers to the position of the value in a sorted list ordered by frequency of occurrence. The most frequent log(j) value is assigned rank 1, while progressively rarer values receive higher ranks. By $f(i)$, we denote the normalised frequency of occurrence of that value, that is, the number of times it appears in a given population of polarisation curves divided by the total number of values N[35] (see 'Clarification of "Rank"' in the Supplementary Information for additional details). The only improvement of the Mandelbrot-Zipf law over the original form of the law (Zipf) is that it fits more adequately the region corresponding to the lowest ranks, that is, Ranks <100.[36]

The Mandelbrot-Zipf distribution defines the probability of accessing an object at rank i out of N available values[37]. The generalised form proposed by Mandelbrot can be written as follows (Eq. 1):

$$f(i) = \frac{K}{(Ci + q)^{\alpha}} \quad \text{(Eq1)}$$

Where K is the normalising constant, C is a second parameter that needs to be adjusted to fit the data, α is the skewness factor, and $q \geq 0$ is the plateau factor. α controls the skewness (slope) of the curve. q is called the plateau factor because it is the reason behind the plateau shape near the left-most part of the distribution. Notice that the higher the value of q, the more flattened the head of the distribution will be. When $q = 0$, the Mandelbrot-Zipf distribution degenerates to a Zipf-like distribution with a skewness factor α[37]. The Mandelbrot curves were fitted using the *curve_fit* module (*scipy.optimise*), and their goodness of fit was evaluated by the R-squared metric (*sklearn.metrics*).

For comparative clarity, when plotting multiple Mandelbrot curves, the normalised frequencies were numerically adjusted to yield curves with an integrated area of one (including points with identical ranks in the Mandelbrot space). This standardisation includes accounting for points with identical ranks in the Mandelbrot space, which is critical for analysing datasets of varying sizes and concentrations on equal terms. Focusing on the relative occurrence of events removes biases introduced by the absolute number of data points, as indicated by the multiplicity bars in the plots. Such adjustment elucidates the distributional dynamics, allowing for a direct comparison of trends and patterns irrespective of sample size differences. The available code [persistent link will be added upon the article acceptance] includes all steps related to the Mandelbrot modelling.

To provide a schematic theoretical foundation for the Mandelbrot-Zipf model, we introduce Fig. 9 here, which illustrates the concept of power-law distributions and the distinction between Zipf and Mandelbrot-Zipf models. Figure 9a shows a schematic example of a data distribution with a wide range of values, extending into a fat-tail region indicative of rare events. Figure 9b compares the Zipf (dashed line) and Mandelbrot-Zipf (solid line) distributions, highlighting how the latter incorporates parameters α (skewness factor) and q (plateau factor) to account for curvature in the head region, offering a more flexible representation than the strictly linear Zipf distribution. This schematic representation is essential for understanding how the Mandelbrot-Zipf model captures the heterogeneity of electrochemical data, particularly in the context of rare events like pitting. For a more complete theoretical background on the Mandelbrot-Zipf model, the reader is referred to the Supplementary Notes. For simplicity, the term Mandelbrot-Zipf will often be named as Mandelbrot from now on.

While other fractal or probability-based models may also be viable, the rank-based and parameter-flexible structure of the Mandelbrot-Zipf model is particularly suited to the highly skewed nature of pitting events, where rare but critical occurrences unfold across multiple orders of magnitude. This makes it a compelling choice for modelling the complex distributions observed in localised corrosion phenomena.

PCA and clustering

The Mandelbrot plots exhibited pronounced arc-shaped deviations from canonical Mandelbrot-Zipf distributions expected relatively linear trend. This divergence from anticipated patterns suggested a potential multifractality, indicating the simultaneous occurrence of multiple fractal processes instead of a singular, overarching fractal dynamic. This realisation required a shift towards a multifractal analytical approach, wherein a series of Mandelbrot curves could be sequentially combined to represent the underlying electrochemical processes more accurately.

Confronted with the Mandelbrot plots' complex, potentially multifractal nature, manual segmentation efforts to define coherent segments correlating with specific corrosion regions (activity, passivity, and pitting) proved inadequate and arbitrary. Consequently, a more objective and systematic approach with improved precision was warranted.

Therefore, *PCA* from *sklearn.decomposition* was deployed to reduce the complexity of the data, aiming to transform the original (2D) Mandelbrot space into a new representation (2D) where segmentation into different regions could naturally emerge - addressing the challenge of arbitrary segmentation due to the continuous nature of the original Mandelbrot plots. PCA identifies the direction (principal component) that captures the most variance in the data, helping filter out noise or less

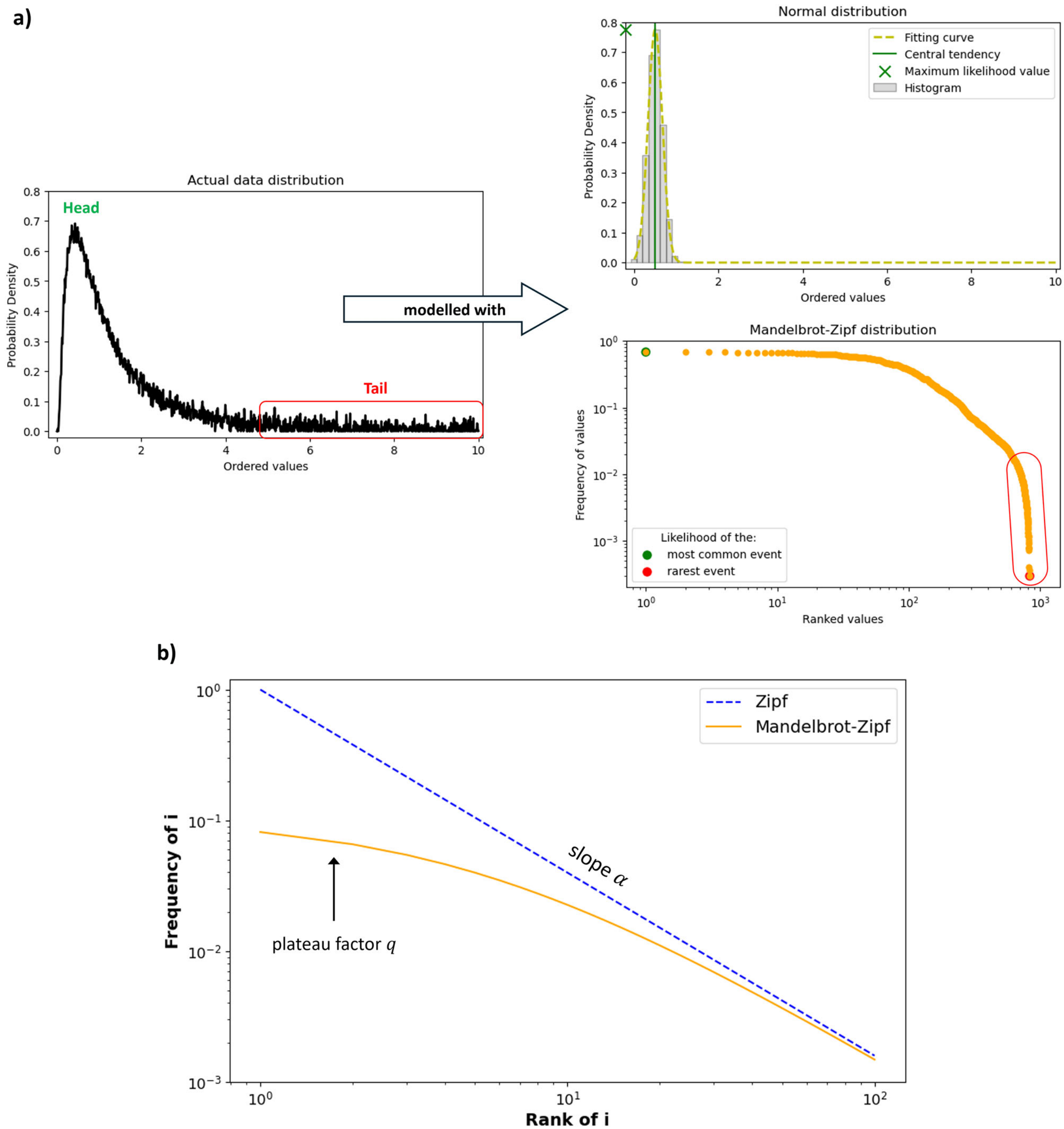


Fig. 9 | Modelling rare events with Mandelbrot-Zipf distributions. **a** A schematic example of a data distribution with a wide range of values, extending into a fat-tail region and indicative of rare events. Applied to a log-log scale, the Mandelbrot-Zipf model effectively captures the intrinsic heterogeneity of such systems, contrasting with normal distributions better suited to narrower value ranges. **b** A comparison of Zipf (dashed line) and Mandelbrot-Zipf (solid line) distributions, plotting the rank of an element *i* against its frequency of occurrence. The Mandelbrot-Zipf model incorporates parameters *α* (skewness factor) and *q* (plateau factor) to account for curvature in the head region, offering a more flexible representation than the strictly linear Zipf distribution.

informative variations. Moreover, by reorienting the data along the principal components, the method offers new perspectives from which patterns might be more easily recognisable. Finally, PCA is often used as a preparing technique for clustering (clusters might be more distinguishable in the space defined by the principal components than in the original space).

We apply PCA to transform the Mandelbrot representation of the data, followed by clustering algorithms (K-Means and Agglomerative Clustering) to reveal natural segments within the transformed data space. It is important to note that no E_sp or E_pit labels are provided during this process; the clusters emerge purely from the underlying data structure. Only after these clusters are identified do we relate them to the known stable pitting events, effectively performing a post-hoc classification.

This strategic application of PCA (based on the two principal components), followed by clustering techniques, facilitated a systematic identification of distinct groupings within the transformed data space. Next, the in-parallel application of clustering algorithms - K-Means and Agglomerative Clustering (*KMeans* and *AgglomerativeClustering* from *sklearn.cluster*) - enabled the robust identification of distinct data groupings. Algorithmic outputs informed an optimal cluster number selection, further corroborated through visual examination. It became apparent that the PCA-transformed Mandelbrot representations could be described as containing discernible segments similar to the parts of a "worm-like" figure (e.g., "head", "core", "tail"). This structured segmentation framework routinely revealed five to six distinguishable regions within each dataset, each demarcated by noticeable discontinuities. Then, each segmented region could be separately modelled with a Mandelbrot curve ("divide and conquer" approach), aiming to capture the multifractal nature underlying the broader electrochemical processes defined in the polarisation curves.

Model evaluation and metrics

Passive potentials (E_pass) were identified following previously described protocols[13].

Although our approach initially operates in an unsupervised manner - identifying clusters without any prior labelling - once these clusters were formed, we associated them post hoc with known stable pitting classes (E_pit or E_sp). This retrospective assignment enabled a classification-like evaluation. Specifically, when evaluating how well the derived clusters corresponded to the E_sp and E_pit points identified independently, we calculated Accuracy (Eq. S1), Precision (Eq. S2), Recall (Eq. S3) and F1 Score (Eq. S4) using *sklearn.metrics*. These metrics reflect how effectively the unsupervised segmentation, after labelling, distinguishes between the stable pitting scenarios of interest.

Detailed definitions of these metrics, along with their computational formulas, are provided in the Supplementary Information (Supplementary Notes 4).

For tasks involving regression-like predictions (e.g., fitting the Mandelbrot curves or assessing potential-current relationships), the Normalised Root Mean Square Error (NRMSE) was employed. NRMSE provides a normalised measure of model error, facilitating easier comparisons across different models or datasets.

Data availability

All data generated or analysed during this study are available in the Zenodo Data repository: https://doi.org/10.5281/zenodo.15284154.

Code availability

The code required to reproduce these findings is available to download from GitHub: https://github.com/bcoelho-leonardo/pitting_pathways_pca_clustering.

Received: 20 February 2025; Accepted: 21 April 2025;
Published online: 03 May 2025

References

1. Långberg, M. et al. Redefining passivity breakdown of super duplex stainless steel by electrochemical operando synchrotron near surface X-ray analyses. *Npj Mater. Degrad.* **3**, 1–11 (2019).
2. Soltis, J. Passivity breakdown, pit initiation and propagation of pits in metallic materials - Review. *Corros. Sci.* **90**, 5–22 (2015).
3. Tang, Y. et al. The metastable pitting potential and its relation to the pitting potential for four materials in chloride solutions. *Corros. Sci.* **80**, 111–119 (2014).
4. Pistorius, P. C. & Burstein, G. T. Growth of corrosion pits on stainless steel in chloride solution containing dilute sulphate. *Corros. Sci.* **33**, 1885–1897 (1992).
5. Burstein, G. T. & Mattin, S. P. Nucleation of corrosion pits on stainless steel. *Philos. Mag. Lett.* **66**, 127–131 (1992).
6. Christiaan, P. & St John's College Cambridge, P. *Stability and Metastability of Corrosion Pits on Stainless Steel*. https://www.repository.cam.ac.uk/items/fe350612-97a2-43bc-a0df-2a14f9ee82af (1991).
7. Yi, Y., Cho, P., Al Zaabi, A., Addad, Y. & Jang, C. Potentiodynamic polarization behaviour of AISI type 316 stainless steel in NaCl solution. *Corros. Sci.* **74**, 92–97 (2013).
8. ASTM. ASTM G61-86(2018). Standard Test Method for Conducting Cyclic Potentiodynamic Polarization Measurements for Localized Corrosion Susceptibility of Iron-, Nickel-, or Cobalt-Based Alloys. https://store.astm.org/g0061-86r18.html (2018).
9. Yang, Z. et al. A statistical study on the effect of hydrostatic pressure on metastable pitting corrosion of X70 pipeline steel. *Materials* **10**, 1307 (2017).
10. Sander, G. et al. On the corrosion and metastable pitting characteristics of 316L Stainless Steel Produced by Selective Laser Melting. *J. Electrochem. Soc.* **164**, C250–C257 (2017).
11. Williams, D. E., Stewart, J. & Balkwill, P. H. The nucleation, growth and stability of micropits in stainless steel. *Corros. Sci.* **36**, 1213–1235 (1994).
12. Zuo, Y., Wang, H., Zhao, J. & Xiong, J. *The Effects of Some Anions on Metastable Pitting of 316L Stainless Steel*. www.elsevier.com/locate/corsci.
13. Coelho, L. B. et al. Estimating pitting descriptors of 316 L stainless steel by machine learning and statistical analysis. *Npj Mater. Degrad.* **7**, 82 (2023).
14. Torres, D., Bailly, J., Bernal, M., Coelho, L. B. & Ustarroz, J. Electrochemical nucleation and the role of the surface state: unraveling activity distributions with a cross-system examination and a local electrochemistry approach. *J. Solid State Electrochem.* https://doi.org/10.1007/s10008-023-05760-y (2023).
15. Bernal, M. et al. A microscopic view on the electrochemical deposition and dissolution of Au with scanning electrochemical cell microscopy – Part I. *Electrochim. Acta* **445**, 142023 (2023).
16. Bernal, M. et al. A microscopic view on the electrochemical deposition and dissolution of Au with scanning electrochemical cell microscopy – Part II: potentiostatic dissolution and correlation with in-situ EC-TEM. *Electrochim. Acta* **492**, 144302 (2024).
17. Kang, M., Bentley, C. L., Mefford, J. T., Chueh, W. C. & Unwin, P. R. Multiscale analysis of electrocatalytic particle activities: linking nanoscale measurements and ensemble behavior. *ACS Nano* **17**, 21493–21505 (2023).
18. Torres, D. et al. Distribution of copper electrochemical nucleation activities on glassy carbon: a new perspective based on local electrochemistry. *J. Electrochem. Soc.* **169**, 102513 (2022).
19. Coelho, L. B. et al. Probing the randomness of the local current distributions of 316 L stainless steel corrosion in NaCl solution. *Corros. Sci.* **217**, 111104 (2023).
20. Pistorius, P. C. & Burstein, G. T. Aspects of the effects of electrolyte composition on the occurrence of metastable pitting on stainless steel. *Corros. Sci.* **36**, 525–538 (1994).
21. Williams, D. E., Westcott, C. & Fleischmann, M. Stochastic models of pitting corrosion of stainless steels: II. Measurement and interpretation of data at constant potential. *J. Electrochem. Soc.* **132**, 1804–1811 (1985).
22. Frankel, G. S. Pitting corrosion of metals: a review of the critical factors. *J. Electrochem Soc.* **145**, 2186–2198 (1998).
23. Zhou, H., Chhin, D., Morel, A., Gallant, D. & Mauzeroll, J. Potentiodynamic polarization curves of AA7075 at high scan rates interpreted using the high field model. *Npj Mater. Degrad.* **6**, 20 (2022).
24. Ansari, T. Q. et al. Phase-field model of pitting corrosion kinetics in metallic materials. *NPJ Comput. Mater.* **4**, 38 (2018).
25. Kiey, S. A. A., Meguid, E. A. A., El & Rehim, S. S. A. El. Electrochemical investigations on the corrosion behavior of 904L stainless steel in LiBr solutions. *J. Mater. Eng. Perform.* **32**, 9163–9173 (2023).
26. Burstein, G. T. & Vines, S. P. Repetitive nucleation of corrosion pits on stainless steel and the effects of surface roughness. *J. Electrochem. Soc.* **148**, B504 (2001).
27. Williams, D. E., Westcott, C. & Fleischmann, M. Stochastic models of pitting corrosion of stainless steels: I. Modeling of the initiation and growth of pits at constant potential. *J. Electrochem. Soc.* **132**, 1796–1804 (1985).
28. Zhang, H. et al. Effect of HAc on the metastable pitting corrosion of 304 SS in NaCl solution. *Materials* **15**, 3618 (2022).

29. Zeng, Y. et al. Effects of Clconcentration, temperature and AC interference on metastable pitting behavior of CoCrFeMnNi High-entropy alloy in a simulated concrete pore solution. *Int. J. Electrochem. Sci.* **17**, 221181 (2022).
30. Stewart, J. & Williams, D. E. The initiation of pitting corrosion on austenitic stainless steel: on the role and importance of sulphide inclusions. *Corros. Sci.* **33**, 457–474 (1992).
31. Zhang, G. A. & Cheng, Y. F. Micro-electrochemical characterization of corrosion of pre-cracked X70 pipeline steel in a concentrated carbonate/bicarbonate solution. *Corros. Sci.* **52**, 960–968 (2010).
32. Coelho, L. B. et al. Mechanical and corrosion characterization of industrially treated 316L stainless steel surfaces. *Surf. Coat. Technol.* **382**, 125175 (2020).
33. Amand, T., Olivier, M. & B. Coelho, L. Potentiodynamic polarisation curves and pitting descriptors of 316L stainless steel in NaCl electrolyte. *Zenodo*. https://doi.org/10.5281/zenodo.15284154 (2025).
34. Mandelbrot, B. Information theory and psycholinguistics: a theory of words frequencies. in *Readings in Mathematical Social Science*. (eds P. Lazafeld and N. Henry) (MIT Press, 1966).
35. Montemurro, M. A. Beyond the Zipf–Mandelbrot law in quantitative linguistics. *Phys. A Stat. Mech. Appl.* **300**, 567–578 (2001).
36. Silagadze, Z. K. Citations and the Zipf – Mandelbrot Law. **1997**, 487–499 (1997).
37. Hefeeda, M. Modeling and caching of peer-to-peer traffic modeling and caching of peer-to-peer. https://doi.org/10.1109/ICNP.2006.320218 (2014).

Acknowledgements

The author L.B.C. is a Postdoctoral Researcher of the Fonds de la Recherche Scientifique – FNRS (Belgium) which is gratefully acknowledged. D.T. acknowledges financial support to the Fonds de Recherche dans l'Industrie et dans l'Agriculture (FRIA). The author, L.B.C. expresses gratitude to Mr. Nassim N. Taleb for his book The Black Swan, which inspired the application of Mandelbrot modelling to the study of pitting corrosion. J.U. acknowledges financial support to the Fédération Wallonie-Bruxelles for the ARC-Consolidator grant RENEGADE.

Author contributions

L.B.C.: Conceptualisation, Methodology, Software, Formal analysis, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualisation, Project administration, Funding acquisition. T.A.: Data Curation. D.T.: Validation, Writing - Review & Editing. M.O.: Validation, Resources, Writing - Review & Editing, Supervision. J.U.: Validation, Formal analysis, Resources, Writing - Review & Editing, Supervision.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41529-025-00594-7.

Correspondence and requests for materials should be addressed to Leonardo Bertolucci Coelho.