Lawrence Berkeley National Laboratory
Chemical Sciences

Title
Electrochemical metrics for corrosion resistant alloys

Permalink
https://escholarship.org/uc/item/5jd9p8sp

Journal
Scientific Data, 8(1)

ISSN
2052-4463

Authors
Nyby, Clara
Guo, Xiaolei
Saal, James E
et al.

Publication Date
2021-02-01

DOI
10.1038/s41597-021-00840-y

Peer reviewed

Check for updates

SCIENTIFIC DATA

OPEN

DATA DESCRIPTOR

Electrochemical metrics for corrosion resistant alloys

Clara Nyby [1,5], Xiaolei Guo[2,5], James E. Saal[1 ✉], Szu-Chia Chien[2], Angela Y. Gerard[3], Huibin Ke[2], Tianshu Li[2], Pin Lu[4], Christian Oberdorfer[2], Sarita Sahu[2], Sirui Li[2], Christopher D. Taylor[2], Wolfgang Windl [2], John R. Scully [3] & Gerald S. Frankel [2 ✉]

Corrosion is an electrochemical phenomenon. It can occur via different modes of attack, each having its own mechanisms, and therefore there are multiple metrics for evaluating corrosion resistance. In corrosion resistant alloys (CRAs), the rate of localized corrosion can exceed that of uniform corrosion by orders of magnitude. Therefore, instead of uniform corrosion rate, more complex electrochemical parameters are required to capture the salient features of corrosion phenomena. Here, we collect a database with an emphasis on metrics related to localized corrosion. The six sections of the database include data on various metal alloys with measurements of (1) pitting potential, E_{pit}, (2) repassivation potential, E_{rp}, (3) crevice corrosion potential, E_{crev}, (4) pitting temperature, T_{pit}, (5) crevice corrosion temperature, T_{crev}, and (6) corrosion potential, E_{corr}, corrosion current density, i_{corr}, passivation current density, i_{pass}, and corrosion rate. The experimental data were collected from 85 publications and include Al- and Fe-based alloys, high entropy alloys (HEAs), and a Ni-Cr-Mo ternary system. This dataset could be used in the design of highly corrosion resistant alloys.

Background & Summary

Corrosion of metallic structures has a substantial impact on the economy. A 1998 study estimates the direct cost of corrosion to be 276 billion USD each year in the U.S. alone, or about three percent of the gross domestic product (GDP)0[1–3]. Therefore, the accurate assessment, control, and prediction of corrosion is of paramount importance. For metals that undergo uniform corrosion, the corrosion rate can often be obtained by mass loss measurements as a function of exposure time to a corrosive environment or by electrochemical measurements. However, for many corrosion resistant alloys (CRAs), either homogeneous solid solutions or multi-phase alloys, this assessment is more complicated because of the existence of a thin oxide on the alloy surface known as a passive film[4]. This film decreases the rate of uniform corrosion but can be susceptible to accelerated localized corrosion, including pitting corrosion, crevice corrosion, and stress corrosion cracking, all of which are associated with the local breakdown of passive films. These corrosion modes can result in unexpected, catastrophic failure, which should be avoided by all means.

Electrochemical approaches that are simple and fast have been widely used to assess the corrosion rate of CRAs, including linear polarization, potentiodynamic polarization, and electrochemical impedance spectroscopy (EIS)[5]. A number of electrochemical metrics derived from these techniques have been adopted to describe the corrosion resistance of alloys, such as corrosion current density (i_{corr}), passive current density (i_{pass}), corrosion potential (E_{corr}), pitting potential (E_{pit}), repassivation potential (E_{rp}), crevice corrosion potential (E_{crev}), pitting temperature (T_{pit}, more commonly abbreviated CPT), and crevice corrosion temperature (T_{crev}, more commonly abbreviated CCT). These metrics are defined in Table 1 and depend on the alloy composition, structure and defects, environmental factors including the temperature and chemical composition of the electrolyte, as well as physical factors such as the crevice former. A typical curve resulting from the cyclic potentiodynamic polarization of a passive metal is schematically illustrated in Fig. 1. Metals are electrochemically polarized from the active region toward the noble region by incrementally stepping the potential while the corresponding current density is recorded. The E_{corr} corresponds to the potential at which no external current flows. At this potential, the metal corrodes at a rate defined by i_{corr}, which is the corrosion current density. This value is generally obtained by extrapolating the linear portion of the anodic and cathodic branches of the polarization curves to E_{corr}. When the

[1]Citrine Informatics, Redwood City, CA, 94063, USA. [2]Department of Materials Science and Engineering, Ohio State University, Columbus, OH, 43210, USA. [3]Department of Materials Science and Engineering, University of Virginia, Charlottesville, VA, 22904, USA. [4]QuesTek Innovations LLC, Evanston, IL, 60201, USA. [5]These authors contributed equally: Clara Nyby, Xiaolei Guo. ✉e-mail: jsaal@citrine.io; frankel.10@osu.edu

	Electrochemical Metric	Definition
i_{corr}	Corrosion current density	Current density on the metal surface describing the general corrosion rate of metals.
i_{pass}	Passive current density	Current density in the passive region
E_{corr}	Corrosion potential	Potential at which the anodic and cathodic reaction rates are equal so that the measured current changes sign at this potential when the potential is scanned as in potentiodynamic polarization
E_{pit}	Pitting potential	Lowest potential at which stable pits form
E_{rp}	Repassivation potential	Potential at which stably growing localized corrosion (pits or crevices) repassivates and stops growing
E_{crev}	Crevice corrosion potential	Lowest potential at which stable crevice corrosion forms
T_{pit} (CPT)	Pitting temperature	Temperature above which stable pitting corrosion initiates
T_{crev} (CCT)	Crevice corrosion temperature	Temperature above which stable crevice corrosion initiates
	Corrosion rate	Sample thickness loss per unit time, sometimes converted from weight loss per unit time normalized by sample area

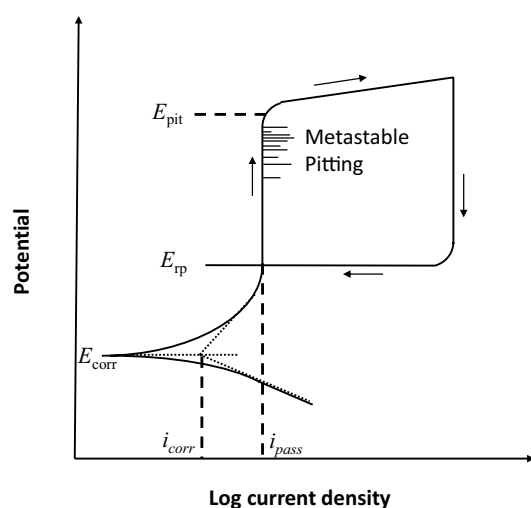Table 1. Definitions of various electrochemical metrics for metal corrosion.

Fig. 1 A typical cyclic potentiodynamic polarization curve, after Frankel, Journal of The Electrochemical Society, 1998, 145, 6[97]. Reproduced with permission.

potential is scanned toward the more noble direction, a passive region exists where the current density is independent of the applied potential. The corresponding current density for this region represents the i_{pass}. Although metastable pitting could occur in this region, no stable pit could form. When the applied potential is more noble than a specific range of values, the passive film breaks down and stable pits form, a process accompanied by a rapid increase of the current density. This characteristic potential is defined as E_{pit}, which has been broadly used to determine the tendency of a given metal or alloy to breakdown. During cyclic polarization, the scan direction of the potential is reversed when the current density exceeds a predetermined value. Subsequently, the E_{rp} value can be frequently but not always achieved when the current density drops substantially, indicating the repassivation of the pit. This value is generally lower than E_{pit}. Since stable pits can only form when the potential is more noble than E_{pit}, and they can only propagate at a potential more noble than E_{rp}, higher values of E_{pit} and E_{rp} suggest that the material is more resistant to pitting corrosion. Compared to the weight loss measurement, these electrochemical metrics possess a number of advantages: (1) They can be obtained much faster, particularly for CRAs that are intentionally designed with high corrosion resistance. (2) For parameters such as E_{pit}, E_{rp}, and T_{pit}, the results can be reproducible with well-controlled experimental conditions. (3) They help to gain an understanding of the corrosion mechanism. Therefore, these electrochemical metrics have been extensively used in the metal corrosion community, and a vast resource of data is available in the literature.

Given the hidden damage that localized corrosion introduces, with its frequent difficulty of detection and consequent potential for unforeseen catastrophic failure of high-value assets, designing materials with resistance to localized corrosion is of high priority. Various factors should be taken into account for such designs, including material composition, microstructure, tendency for passivity, and electrochemical activity of the surface. However, current approaches for the development of CRAs are primarily based on intuition, history of past successes, and trial and error. Although empirical models capable of predicting the corrosion resistance of a specific group of materials exist in the literature, they are strictly constrained to a very limited and particular range of compositional space. For instance, the pitting resistance equivalence number (PREN) was developed as a figure of merit by correlating the pitting corrosion resistance to alloy compositions for Fe-Cr-Ni alloys[6]. One relation characterizing the resistance of austenitic stainless steels to pitting corrosion is[7]:
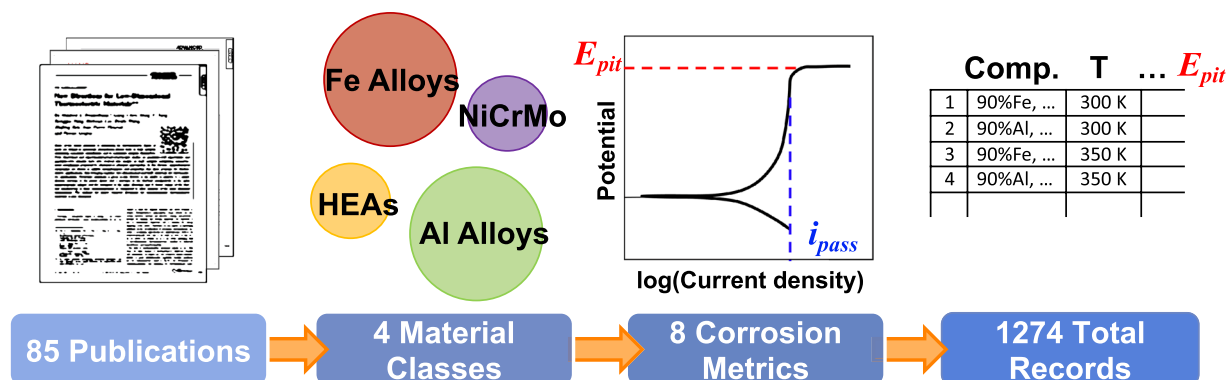
Fig. 2 A schematic overview of the dataset. The data were collected from 85 publications, with materials which fall into 4 material classes. There are 6 datasets reporting a total of 8 different corrosion metrics, with 1274 total records.

$$PREN = \%Cr + 3.3 \times \%Mo + 19.4 \times \%N \tag{1}$$

where the coefficients simply describe the relative effects of Mo and N to that of Cr and the concentrations are in weight percent. Based on this equation, the pitting corrosion resistance of austenitic stainless steel can be primarily controlled by the amount of beneficial components, i.e. Cr, Mo, and N. While the PREN has been broadly used in the CRA industry, this equation cannot be extrapolated to compositions outside of those used to fit the equation, including high entropy alloys (HEAs) and aluminum alloys.

In this study, we create a CRA database with electrochemical metrics that could be fed into machine learning (ML) based models to allow for exploration of the compositional space beyond what was used to create the current empirical models. To our knowledge, this is the first published database of this type. Existing large-scale databases only contain the uniform corrosion rate of certain types of alloys[8,9]. However, for CRAs, the uniform corrosion rate is not very meaningful because the dominant corrosion mechanism is localized corrosion as has been introduced above. The rate of the localized corrosion can be many orders of magnitude higher than that of uniform corrosion. Therefore, more complex electrochemical parameters, such as the pitting potential and repassivation potential or other metrics reported in this database, are required to capture these corrosion phenomena. A high level overview of the dataset is shown in Fig. 2. This database not only allows us to link the corrosion resistance of CRAs to various experimental parameters, including materials composition and environmental attributes (temperature, pH, and chloride ion concentration), but also enables the development of calculable matrices that could shed light on the fundamental physical processes that govern the corrosion performance. For instance, the corrosion resistance of metals could be correlated to the bonding strengths of metal-metal and metal-oxygen bonds[6,7], chloride ion adsorption susceptibility[10,11] and oxide enrichment and depletion factors[12], which could be calculated through density functional theory or molecular dynamics approaches. ML models can also be integrated into a generic multi-physics modeling framework to bridge gaps where there is not yet mechanistic theory that enables the prediction of complex corrosion phenomena. Therefore, development of corrosion databases such as this may be crucial for the future of alloy design and optimization.

Methods

The data were collected from experimental results reported across 85 literature sources[7,10,11,13–94]. Each composition was assigned a "Material class" based on the elements present and the source of the data. The Fe- and Al-based alloys are compositions that contain >35 wt. % Fe or Al, respectively. High entropy alloys are compositions that were reported as such in the source literature. The Ni-Cr-Mo ternary system contains alloys with all three elements present as well as some compositions containing just one or two of the three base elements. Alloy compositions which fall outside these bounds are classified as "Other". The data for Fe-based alloys were collected from the references of a book[95], while the data for the Ni-Cr-Mo ternary system of alloys were extracted from a doctoral dissertation[89]. The data for HEAs and Al alloys were collected from existing literature by searching the Corrosion journal, The Journal of Science and Technology, Science Direct, and Google Scholar. Some values were reported directly in each source, while the others were extracted from the figures using the WebPlotDigitizer software[96]. The database consists of six datasets containing different electrochemical metrics for CRAs including corrosion current density (i_{corr}), passive current density (i_{pass}), corrosion potential (E_{corr}), pitting potential (E_{pit}), repassivation potential (E_{rp}), crevice corrosion potential (E_{crev}), pitting temperature (T_{pit}), and crevice corrosion temperature (T_{crev}) (see Table 1 for definitions). Some sources provided average values of these descriptors, while others reported the maximum and minimum values. All of these values were recorded.

In addition to the electrochemical metrics, metadata related to details of the experiments were recorded. For every measurement reporting a corrosion potential, the experimental parameters include the alloy composition and environmental attributes of the corrosion experiment (temperature, test solution, pH, and chloride ion concentration). When provided, the alloy microstructure, alloy heat treatment, electrochemical testing method, and the electrochemical scanning rate were also recorded. For every measurement reporting a critical temperature

Dataset	total	Fe	Al	HEA	NiCrMo	Other
E_{pit}	810	523	118	27	92	50
E_{rp}	189	45	41	9	92	2
E_{crev}	31	31	0	0	0	0
T_{pit}	117	116	0	0	0	1
T_{crev}	70	35	0	0	0	35
E_{corr}, i_{corr}, and i_{pass}	51	0	0	51	0	0

Table 2. The number of data records by electrochemical metric and material class. Here, "Fe" = Fe-based alloy, "Al" = Al-based alloy, "NiCrMo" = Ni-Cr-Mo ternary system alloy, "Other" = Other alloy.

as the corrosion metric, the experimental parameters reported include the alloy composition, test solution, and test method, with some measurements also including the scan rate of the temperature, test potential, alloy heat treatment, and alloy microstructure.

Data Records

The data were collected from sources studying Fe-based alloys, HEAs, a Ni-Cr-Mo ternary system of alloys, and Al-based alloys. A description of the number of data records based on the electrochemical metric and material class is provided in Table 2. Beyond the reported electrochemical metrics and elemental composition, other fields that may be populated are:

- [Cl-]: Chloride ion concentration. It has been well established that the presence of chloride ions can break down the passive film of various CRAs and the extent of corrosion strongly depends on the concentration of these ions[97]. Therefore, it is crucial to include this parameter as an input for any model designed for the prediction of metal corrosion.

- Microstructures: Description of reported microstructures. In general, homogeneous solid solutions are more resistant to localized corrosion compared to multi-phase alloys. Localized corrosion can be initiated preferentially at heterogeneity on the surface, including dislocations and defects, secondary phase particles, inclusions, and grain boundaries, all of which are closely related to the microstructure of a given alloy. For example, the S-phase (Al_2CuMg) particles existing on the surface of many Al-based alloys are more active than the matrix, so they are susceptible to localized breakdown. Similarly, pitting corrosion is usually initiated on MnS precipitates that exist in most stainless steels[97].

- Oxide: Oxide layer composition, if present and reported. As described above, localized corrosion is a phenomena associated with the localized breakdown of the surface oxide, which strongly depends on the chemical composition of the oxide. The most widely used CRAs to date are primarily based on Fe-Cr and Ni-Cr, which rely on the formation of a Cr-rich surface layer that is highly resistant to aqueous corrosion. Therefore, the oxide composition, if known, will be beneficial for the understanding and prediction of localized corrosion.

- pH: pH of bulk test solution. Metal corrosion is influenced not only by the type of materials and potential, but also by the solution pH. The thermodynamic stability of a given metal or alloy in different pH environments can be found in the potential-pH diagram, also known as the Pourbaix diagram[98]. For example, under open circuit potential at room temperature, Fe can be strongly dissolved at acidic pH whereas passivisation occurs under basic conditions. Al corrodes slowly in a neutral environment but it is soluble in both acidic and alkaline environments.

- Scan rate of temperature: For temperature-dependent scans, this reports the scan rate of temperature in °C/min. For example, the determination of the T_{pit} relies on progressively increasing or decreasing the temperature at a specific step. This step sizes has been known to influence the measured T_{pit} values and their range of distribution[99].

- Test Solution: Chemical composition of test solution. Corrosion is not an intrinsic property that only correlates to the material itself; it is also significantly influenced by the environment. The chemistry of the test solution is a critical environmental factor for metal corrosion and it directly determines whether localized corrosion will occur. In solutions without aggressive anions, predominantly chloride ions, localized corrosion does not occur, so additional care should be taken when interpreting the results obtained by electrochemical approaches. For instance, the rapid increase of current beyond the passive regime may be transpassive dissolution rather than pitting. Additionally, when the solution contains oxidizing agents, such as Fe^{3+}, the likelihood of localized corrosion will increase. Similarly, if corrosion inhibitors are present in the test solution, the results should not be directly combined with those inputs collected in an inhibitor-free environment.

- Test temp.: Test temperature (°C). Corrosion is fundamentally governed by both thermodynamics and kinetics, so any changes in temperature will influence the rate of corrosion. An elevated temperature not only helps to overcome the energy barrier required for a given corrosion process to occur, but also accelerates the corrosion rate by simply enhancing the mass transport. In addition, many CRAs do not undergo localized corrosion until a critical temperature range is exceeded, which is usually expressed as the critical crevice temperature or critical pitting temperature. Recent studies showed that these parameters are statistically distributed rather than single-valued[99]. Due to the critical role temperature plays during corrosion, this input must be taken into account for the prediction of the corrosion rate for any metals.

	Composition (wt. %)										
No.	Fe	Cr	Ni	Mo	Avg E$_{rp}$ (V$_{SCE}$)	Temp (°C)	Test Solution	[Cl—] (M)	pH	Material class	Ref.
1	80.69	18.00	0.00	0.00	−504.0	60	3.5% NaCl	0.6	7	Fe Alloy	[48]
2	78.26	18.40	0.00	1.96	−478.9	60	3.5% NaCl	0.6	7	Fe Alloy	[48]
3	76.45	22.20	0.01	0.03	−403.5	60	3.5% NaCl	0.6	7	Fe Alloy	[48]
4	74.38	22.40	0.00	1.91	−465.9	60	3.5% NaCl	0.6	7	Fe Alloy	[48]
48	19.00	17.00	19.00	0.00	49.0	25	1 N H2SO4	1E-6	0	HEA	[18]
49	18.00	17.00	19.00	0.00	25.0	25	1 N H2SO4	1E-6	0	HEA	[18]
50	18.00	17.00	19.00	0.00	−156.0	25	1 N H2SO4	1E-6	0	HEA	[18]
51	18.00	16.00	18.00	0.00	−344.0	25	1 N H2SO4	1E-6	0	HEA	[18]
80	0.00	13.31	81.45	5.24	−33.8	45	0.5 M NaCl	0.5	7	NiCrMo	[44]
81	0.00	20.53	74.55	4.92	81.7	45	0.5 M NaCl	0.5	7	NiCrMo	[44]
82	0.00	16.95	74.64	8.41	−3.9	45	0.5 M NaCl	0.5	7	NiCrMo	[44]
83	0.00	28.67	65.94	5.38	90.68	45	0.5 M NaCl	0.5	7	NiCrMo	[44]
150	0.00	0.00	0.00	0.00	−806	25	0.6 M NaCl	0.6	7	Al Alloy	[56]
151	0.00	0.00	0.00	0.00	−754	25	0.6 M NaCl	0.6	7	Al Alloy	[56]
152	0.00	0.00	15.76	0.00	−358	25	0.6 M NaCl	0.6	7	Al Alloy	[56]
153	7.99	0.00	0.00	0.00	−304	25	0.6 M NaCl	0.6	7	Al Alloy	[56]

Table 3. A section of the E_{rp} dataset. The full elemental composition includes the following elements: Fe, Cr, Ni, Mo, W, N, Nb, C, Si, Mn, Cu, P, S, Al, V, Ta, Re, Ce, Ti, Co, B, Mg, Y, and Gd.

- Test method: Type of electrochemical/chemical test performed, e.g. potentiodynamic, potentiostatic polarization, or immersion test. It is commonly observed that results acquired by different techniques can be slightly different, which should be considered if these inputs are combined.
- Scan rate (mV/s): The rate of potentiodynamic polarization scan in mV/s. During potentiodynamic polarization, the metals are polarized by incrementally increasing or decreasing the potential with a fixed step size, and the corresponding current density (i.e. corrosion rate) is measured. This electrochemical approach relies on the determination of surface reaction kinetics in a steady-state regime, otherwise the measured corrosion rate will deviate from the actual value, leading to inaccurate results that cannot be trusted. Therefore, the scan rate of potentiodynamic polarization must lie within a reasonable range for the results to be trusted, as has been reported elsewhere[100].
- Heat treatment: Description of reported heat processing steps. Heat treatment predominately influences the structure of the alloys, so this input should be considered if provided.
- Reference: Literature reference number or DOI.
- Comment: If given, this may include information such as the name of a commercial alloy used in the test or other experimental details.

The database is available online on both the Citrination website[101] as well as figshare[102]. On the Citrination website, the database can be downloaded as a collection of physical information files (PIFs)[103], downloaded as an Excel spreadsheet, or used with the Citrination platform to build ML models. On the figshare website, the database can be downloaded as an Excel spreadsheet, in which the six datasets are present as individual tabs. The Fe-based alloys are presented in the spreadsheet without color highlighting and with the "Material class" column as "Fe Alloy". The rest of the data include HEAs (orange color highlighting, "Material class": "HEA"), Ni-Cr-Mo ternary system of alloys (green color, "Material class": "NiCrMo Alloy")), Al-based alloys (blue color, "Material class": "Al Alloy")), and a category containing materials which did not fall into any of the four material classes (purple color, "Material class": "Other")). An example section of the E_{rp} dataset (without color highlighting) is shown in Table 3, and the other datasets follow the same or a very similar format. The reference number corresponds to the references in the final tab of the dataset. This data can be straightforwardly incorporated for use in a ML framework.

Technical Validation

The data were collected and verified as accurate by a team of scientists familiar with all of the electrochemical metrics of corrosion reported here. Furthermore, additional screening was performed by examining outliers in various statistical plots of the datasets. Outliers were investigated and corrected if a transcription error or misinterpretation of the literature value was identified.

This database contains compositions ranging from pure elements, multi-component high entropy alloys, and complex designed steels, with a total of 24 elements represented. To visualize the spread of the compositions in the data, Fig. 3 shows the configurational entropy, ΔS, for each of the six datasets. The ΔS of an alloy is given by[104]:

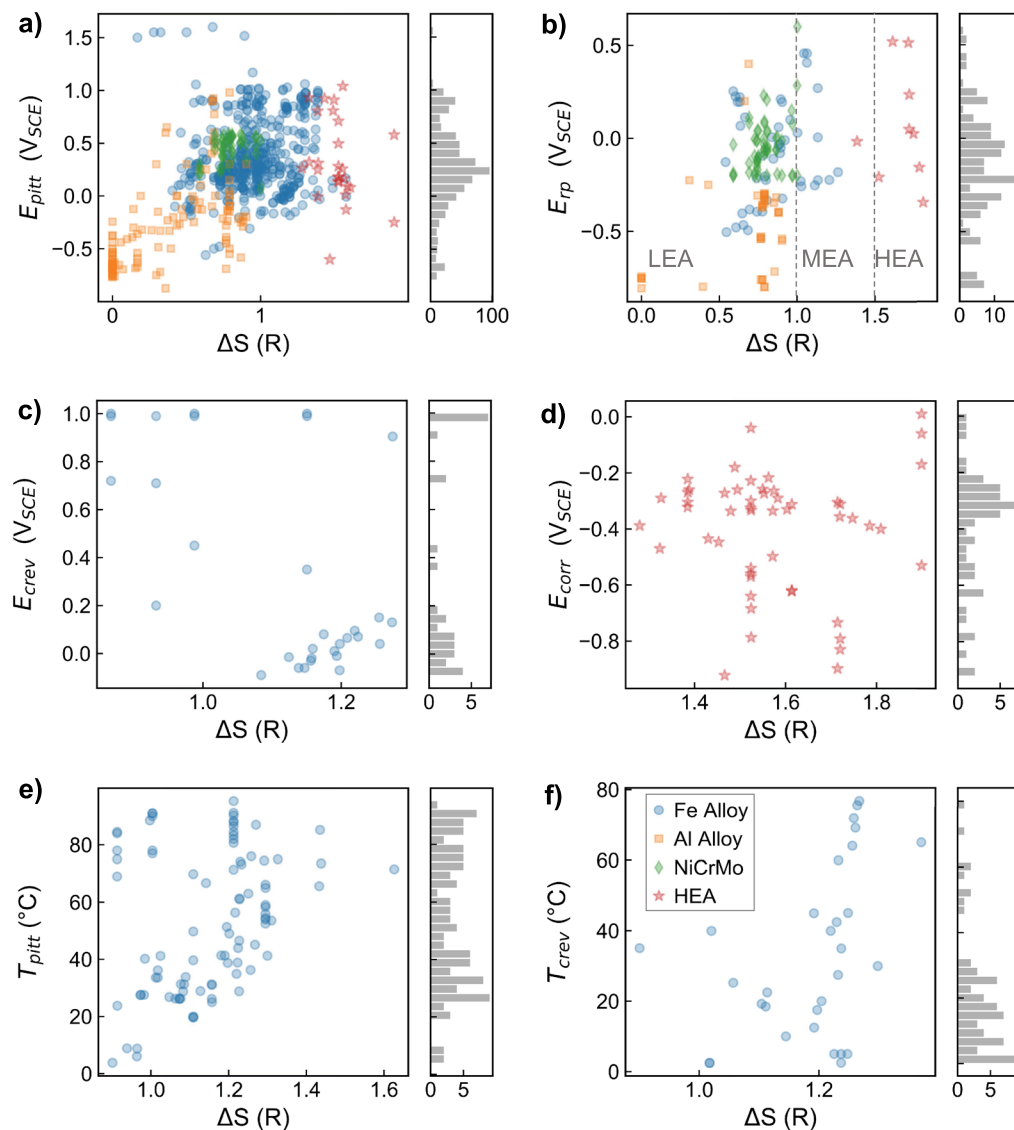$$\Delta S = -R\sum_{i=1}^{n}x_i ln x_i$$

(2)

Fig. 3 Configurational entropy of datasets. Here, we show the configurational entropy, ΔS, of the datasets reporting (**a**) E_{pit} in volts vs. saturated calomel electrode (V_{SCE}), (**b**) E_{rp} in V_{SCE}, (**c**) E_{crev} in V_{SCE}, (**d**) E_{corr} in V_{SCE}, (**e**) T_{pit} in degrees Celcius, and (**f**) $T_{crev,max}$ (maximum of measured T_{crev} value) in degrees Celcius. The dashed lines in b) show the guidelines for classifying a low entropy alloy (LEA), medium entropy alloy (MEA) or high entropy alloy (HEA) based on the ΔS value. The histogram to the right of each panel shows the distribution in measured values for each corrosion metric.

where n is the number of elements, x_i the concentration of the i-th element and R is the universal gas constant. Note that this quantity represents the maximum possible configurational entropy for a given alloy composition, where this maximum value would be reached in the case of an ideal solid solution[105]. Different regimes of configurational entropy include low entropy alloys (LEA) having $0 < \Delta S \leq 1.5R$, medium entropy alloys (MEA) in the range of $1 \leq \Delta S \leq 1.5R$ and HEAs have $\Delta S > 1.5R$[104]. We find that all of the classes of materials fall in a reasonable range, with the data points classified as HEAs in our database mostly above $\Delta S = 1.5R$, the Al-based alloys falling in the LEA range (with some measurements on pure Al present at $\Delta S = 0$), the Fe-based alloys spanning the LEA to MEA ranges, and the Ni-Cr-Mo ternary system of alloys falling in a cluster within the LEA range. These expected clusterings support the transcription accuracy of the elemental compositions collected in this database.

In Fig. 4, we show the temperature dependence of the database, linking data from sources that included measurements at multiple temperatures. Here, we clearly see that the majority of the data were measured at room temperature. Indeed, all measurements for Al-based alloys and HEAs were measured at room temperature. Still, among the measurements with multiple temperatures for a given composition, we see a general decrease in E_{pit}, E_{rp}, and E_{crev} with increasing temperature, which follows the expected trend[97], further supporting the validity of the dataset.
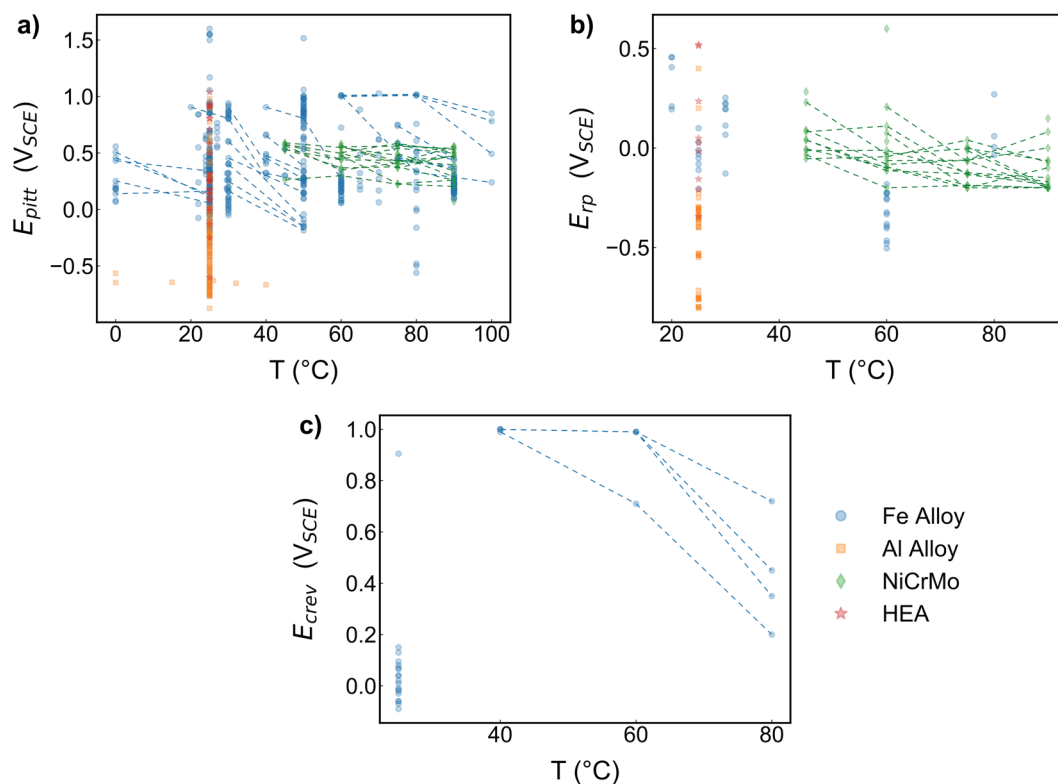
Fig. 4 Temperature distribution of datasets. Here, we show the temperature distribution for (**a**) E_{pit} in V_{SCE}, (**b**) E_{rp} in V_{SCE}, (**c**) E_{crev} in V_{SCE}. Dashed lines indicate measurements on the same material performed at varying temperatures.

Usage Notes

There are some cases in which the corrosion metric is given as a lower bound, rather than a real value. In particular, many of the entries in the Ni-Cr-Mo ternary dataset are reported as a lower bound for the E_{rp} potential because the corrosion resistance extended beyond the maximum potential of the electrochemical measurement. In order to use these data entries in a regression algorithm, these must be converted to a real number, with the method for conversion depending upon the intended application. Additionally, the content of some trace elements such as C, S, and P was often not provided by the original source. In those cases, the values for non-existing entries is reported as "NA". Since the localized corrosion of several alloys depends on these trace elements, such as the role of sulfide in stainless steels, these fields may need to be replaced with estimated values. In some rare cases, the concentrations of some major elements such as Cr, Ni, and Mo were not provided by the original source and were therefore also reported as NA. These entries should be treated differently compared to the trace elements and should not be replaced with zeros. Detailed information on the specimens, when provided by the original source, was included in the "comments" field. This information includes materials grade, samples preparation details, surface finish, cleanliness, metallurgy methods, test area, and test location. It should be noted that the data are not uniformly distributed over material classes, elemental composition, or other features such as chloride concentration. Due to the nature of experimental corrosion data found in literature, it is impossible to achieve a perfectly balanced set of data with respect to all features. Therefore, the imbalance in the database should be taken into account when incorporating the data into statistical models.

Additional composition-dependent features could be generated for use in ML algorithms. For example, the Magpie featurization library could be used to generate features based on elemental and ionic properties, stoichiometry, and electronic structure[106,107]. Additionally, physics-based simulations could be performed to generate additional composition-dependent features. For example, the properties of alloys are influenced by the thermodynamic activity of each component, which can be calculated and used as inputs to the model. Furthermore, the severity of metal corrosion strongly depends on the bonding strengths of metal-metal and metal-oxygen bonds[108]. Specifically, alloying elements with strong metal-metal bonds, such as Mo, Nb, Ta, and W, do not corrode easily so they can act as dissolution moderators or blockers. Elements with weak metal-metal but high metal-oxygen bonds, such as Al, Ti, and Cr, can facilitate the formation of the surface passivation film. Therefore, the metal-metal and metal-oxygen bond strengths can be calculated[6,7] as additional input features. Similarly, chloride ion adsorption susceptibility[10,11] can be calculated because different alloying elements have different affinity toward the chloride ions, which are the dominant aggressive species that can break down the alloy passivity.

Care should be taken when combining input variables acquired from different environments. For example, pitting potential could vary with the solution aeration condition, particularly for alloys with a lower pitting potential than the corrosion potential. In naturally aerated solutions, the pitting potential of aluminum alloy 2024,

for example, is close to its corrosion potential, E_{corr-1}. However, in deaerated conditions, the corrosion potential strongly decreases (E_{corr-2}) due to the absence of the oxygen reduction reaction, so the true pitting corrosion of this alloy can be revealed to be in the range of E_{corr-2} to E_{corr-1}. Therefore, the solution aeration condition may be considered as a separate input variable. Additionally, some corrosion data were acquired in synthetic sea water, which has a much more complicated chemistry compared to other solution environments. For simplicity, a predictive model could use only pH and chloride ion concentration to approximate the solution chemistry. However, other species existing in this environment could also play a role during the corrosion measurement. Furthermore, some corrosion studies were performed in solutions saturated with oxygen or hydrogen, which may influence the redox environment and thus affect the measured electrochemical metrics.

Code availability

No custom code was developed for the generation or processing of this dataset.

References

1. Koch, G., Brongers, M., Thompson, N., Virmani, Y. P. & Payer, J. *Corrosion costs and preventive strategies in the United States*. Report No. FHWA-RD-01-156 (Department of Transportation, 2002).
2. Thompson, N. G., Yunovich, M. & Dunmire, D. Cost of corrosion and corrosion maintenance strategies. *Corrosion Reviews* **25**, 247–261 (2007).
3. Taylor, C. D., Lu, P., Saal, J., Frankel, G. S. & Scully, J. R. Integrated computational materials engineering of corrosion resistant alloys. *npj Materials Degradation* **2** (2018).
4. Frankel, G. S. *et al*. A comparative review of the aqueous corrosion of glasses, crystalline ceramics, and metals. *npj Materials Degradation* **2** (2018).
5. Kelly, R. G., Scully, J. R., Shoesmith, D. & Buchheit, R. *Electrochemical Techniques in Corrosion Science and Engineering* (CRC Press, 2002).
6. Lorenz, K. & Medawar, G. Über das Korrosionsverhalten austenitischer Chrom-Nickel(Molybdän-)Stähle mit und ohne Stickstoffzusatz unter besonderer Berücksichtigung ihrer Beanspruchbarkeit in chloridhaltigen Lösungen. *Thyssen Forschung* **1**, 97–108 (1969).
7. Jargelius-Pettersson, R. F. Application of the Pitting Resistance Equivalent Concept to Some Highly Alloyed Austenitic Stainless Steels. *Corrosion* **54**, 162–168 (1998).
8. Ricker, R. CORR-DATA. https://doi.org/10.18434/M3TH4R (1997).
9. National Association of Corrosion Engineers. Corrosion survey database (COR SUR). https://search.library.wisc.edu/catalog/9910350218802121 (2002).
10. Horvath, J. & Uhlig, H. H. Critical Potentials for Pitting Corrosion of Ni, Cr-Ni, Cr-Fe, and Related Stainless Steels. *Journal of The Electrochemical Society* **115**, 791 (1968).
11. Bond, A. P. & Lizlovs, E. A. Anodic Polarization of Austenitic Stainless Steels in Chloride Media. *Journal of The Electrochemical Society* **115**, 1130 (1968).
12. Gerard, A. Y. *et al*. Aqueous passivation of multi-principal element alloy $Ni_{38}Fe_{20}Cr_{22}Mn_{10}Co_{10}$: Unexpected high Cr enrichment within the passive film. *Acta Materialia* **198**, 121–133 (2020).
13. Pessall, N. & Nurminen, J. I. Development of Ferritic Stainless Steels for Use in Desalination Plants. *Corrosion* **30**, 381–392 (1974).
14. Postlethwaite, J., Scoular, R. J. & Dobbin, M. H. Localized Corrosion of Molybdenum-Bearing Nickel Alloys in Chloride Solutions. *Corrosion* **44**, 199–203 (1988).
15. Hibner, E. L. Evaluation of test procedures for critical crevice temperature determination for Ni-base alloys in a ferric chloride environment. In *Corrosion/86*, 181 (NACE, Houston, TX, 1986).
16. Bond, A. P. Effects of Molybdenum on the Pitting Potentials of Ferritic Stainless Steels at Various Temperatures. *Journal of The Electrochemical Society* **120**, 603 (1973).
17. Böhni, H. & Uhlig, H. H. Environmental Factors Affecting the Critical Pitting Potential of Aluminum. *Journal of The Electrochemical Society* **116**, 906 (1969).
18. Guo, R. & Ives, M. B. Pitting susceptibility of stainless steels in bromide solutions at elevated temperatures. *Corrosion* **46**, 125–129 (1990).
19. Carroll, W. M. & Lynskey, E. E. A crevice-free electrode assembly for the determination of reproducible breakdown potentials for stainless steels in halide environments. *Corrosion Science* **36**, 1667–1678 (1994).
20. Richardson, W. H. & Guha, P. Improved ferritic-austenitic stainless steel. *British Corrosion Journal* **14**, 167–170 (1979).
21. Lizlovs, E. A. & Bond, A. P. Anodic Polarization of Some Ferritic Stainless Steels in Chloride Media. *Electrochem Soc-J* **116**, 574–579 (1969).
22. Tomashov, N. D., Chernova, G. P. & Marcova, O. N. Effect of Supplementary Alloying Elements On Pitting Corrosion Susceptibility Of 18Cr-14Ni Stainless Steel. *Corrosion* **20**, 166t–173t (1964).
23. Brigham, R. J. Temperature As a Crevice Corrosion Criterion. *Corrosion* **30**, 396–398 (1974).
24. Brigham, R. J. & Tozer, E. W. Temperature as a Pitting Criterion. *Corrosion* **29**, 33–36 (1973).
25. Brigham, R. J. & Tozer, E. W. Effect of alloying additions on the pitting resistance of 18% Cr austenitic stainless steels. *Corrosion* **30**, 161–166 (1974).
26. Azuma, S., Ogawa, K., Miyuki, H., Kudo, T. & Nishi, M. Corrosion Resistance of 28% Chromium Duplex Stainless Steel Tube in Hot Seawater Applications. *Stainless Steels'91* **1**, 133–138 (1991).
27. Qiu, Y., Gibson, M. A., Fraser, H. L. & Birbilis, N. Corrosion characteristics of high entropy alloys. *Materials Science and Technology (United Kingdom)* **31**, 1235–1243 (2015).
28. Lee, C. P., Chen, Y. Y., Hsu, C. Y., Yeh, J. W. & Shih, H. C. The Effect of Boron on the Corrosion Resistance of the High Entropy Alloys $Al_{0.5}CoCrCuFeNiBx$. *Journal of The Electrochemical Society* **154**, C424 (2007).
29. Hsu, Y. J., Chiang, W. C. & Wu, J. K. Corrosion behavior of $FeCoNiCrCu_x$ high-entropy alloys in 3.5% sodium chloride solution. *Materials Chemistry and Physics* **92**, 112–117 (2005).
30. Chen, Y. Y., Duval, T., Hung, U. D., Yeh, J. W. & Shih, H. C. Microstructure and electrochemical properties of high entropy alloys-a comparison with type-304 stainless steel. *Corrosion Science* **47**, 2257–2279 (2005).
31. Sugimoto, K. & Sawada, Y. Role of Alloyed Molybdenum in Austenitic Stainless Steels in the Inhibition of Pitting in Neutral Halide Solutions. *Corrosion* **32**, 347–352 (1976).
32. Lin, C. M. & Tsai, H. L. Evolution of microstructure, hardness, and corrosion properties of high-entropy $Al_{0.5}CoCrFeNi$ alloy. *Intermetallics* **19**, 288–294 (2011).

33. Kao, Y. F., Lee, T. D., Chen, S. K. & Chang, Y. S. Electrochemical passive properties of Al$_x$CoCrFeNi (x = 0, 0.25, 0.50, 1.00) alloys in sulfuric acids. *Corrosion Science* **52**, 1026–1034 (2010).
34. Chou, Y. L., Yeh, J. W. & Shih, H. C. The effect of molybdenum on the corrosion behaviour of the high-entropy alloys Co$_{1.5}$CrFeNi$_{1.5}$Ti0.5Mo$_x$ in aqueous environments. *Corrosion Science* **52**, 2571–2581 (2010).
35. Wang, H. L. *et al*. Microstructure, thermal properties, and corrosion behaviors of FeSiBAlNi alloy fabricated by mechanical alloying and spark plasma sintering. *International Journal of Minerals, Metallurgy and Materials* **23**, 77–82 (2016).
36. Lee, C. P., Chang, C. C., Chen, Y. Y., Yeh, J. W. & Shih, H. C. Effect of the aluminium content of AlxCrFe$_{1.5}$MnNi$_{0.5}$ high-entropy alloys on the corrosion behaviour in aqueous environments. *Corrosion Science* **50**, 2053–2060 (2008).
37. Shang, X. L. *et al*. The intrinsic mechanism of corrosion resistance for FCC high entropy alloys. *Science China Technological Sciences* **61**, 189–196 (2018).
38. Shi, Y. *et al*. Homogenization of Al$_x$CoCrFeNi high-entropy alloys with improved corrosion resistance. *Corrosion Science* **133**, 120–131 (2018).
39. Rodriguez, A. A. *et al*. Effect of Molybdenum on the Corrosion Behavior of High-Entropy Alloys CoCrFeNi$_2$ and CoCrFeNi$_2$ Mo 0.25 under Sodium Chloride Aqueous Conditions. *Advances in Materials Science and Engineering* 1–11 (2018).
40. Szklarska-Smialowska, Z. Z.-S. Pitting and crevice corrosion. In *NACE international Houston, TX* (2005).
41. Lucente, A. M. & Scully, J. R. Pitting of Al-based amorphous-nanocrystalline alloys with solute-lean nanocrystals. *Electrochemical and Solid-State Letters* **10**, 39–43 (2007).
42. Sadough Vanini, A., Audouard, J. P. & Marcus, P. The role of nitrogen in the passivity of austenitic stainless steels. *Corrosion Science* **36**, 1825–1834 (1994).
43. Lim, Y. S., Kim, J. S., Ahn, S. J., Kvvon, H. S. & Katada, Y. The influences of microstructure and nitrogen alloying on pitting corrosion of type 316 L and 20 wt.% Mn-substituted type 316 L stainless steels. *Corrosion Science* **43**, 53–68 (2001).
44. Olefjord, I. & Wegrelius, L. The influence of nitrogen on the passivation of stainless steels. *Corrosion Science* **38**, 1203–1220 (1996).
45. Roberge, R. Effect of the Nickel Content in the Pitting of Stainless Steels in Low Chloride and Thiosulfate Solutions. *Corrosion* **44**, 274–280 (1988).
46. Jargelius-Pettersson, R. F. Electrochemical investigation of the influence of nitrogen alloying on pitting corrosion of austenitic stainless steels. *Corrosion Science* **41**, 1639–1664 (1999).
47. Tsuge, H., Tarutani, Y. & Kudo, T. Effect of Nitrogen on the Localized Corrosion Resistance of Duplex Stainless Steel Simulated Weldments. *Corrosion* **44**, 305–314 (1988).
48. Galvele, J. R. & de De Micheli, S. M. Mechanism of intergranular corrosion of Al-Cu alloys. *Corrosion Science* **10**, 795–807 (1970).
49. Garfias-Mesias, L. F., Sykes, J. M. & Tuck, C. D. The effect of phase compositions on the pitting corrosion of 25 Cr duplex stainless steel in chloride solutions. *Corrosion Science* **38**, 1319–1330 (1996).
50. Pawel, S. J., Stansbury, E. E. & Lundin, C. D. Role of nitrogen in the pitting resistance of cast duplex CF-type stainless steels. *Corrosion* **45**, 125–133 (1989).
51. Jargelius, R. No Title. In *10th Scandinavian Corrosion Congress NKM* 10, 161–164 (Sweden, 1986).
52. Guillaumin, V. & Mankowski, G. Localized corrosion of 2024 T351 aluminium alloy in chloride media. *Corrosion Science* **41**, 421–438 (1998).
53. Palit, G. C., Kain, V. & Gadiyar, H. S. Electrochemical investigations of pitting corrosion in nitrogen-bearing type 316N stainless steel. *Corrosion* **49**, 979–991 (1993).
54. Wilde, B. E. & Williams, E. On the Correspondence between Electrochemical and Chemical Accelerated Pitting Corrosion Tests. *Journal of The Electrochemical Society* **117**, 775 (1970).
55. Hartline, A. G. Effect of nitrogen additions upton the pitting resistance of 18 PCT Cr, 18 PCT Mn stainless steel. *Metall Trans* **5**, 2271–2276 (1974).
56. Ujiro, T., Yoshioka, K. & Staehle, R. W. Differences in Corrosion Behavior of Ferritic and Austenitic Stainless Steels. *CORROSION* **50**, 953–962 (1994).
57. Malik, A. U., Siddiqi, N. A., Ahmad, S. & Andijani, I. N. The effect of dominant alloy additions on the corrosion behavior of some conventional and high alloy stainless steels in seawater. *Corrosion Science* **37**, 1521–1535 (1995).
58. Charles, J., Audouard, J. P. & Verneau, M. Metallic answers for F.G.D. systems. In *NACE - International Corrosion Conference Series*, vol. 1998-March (1998).
59. Dowling, N. J., Kim, Y. H., Ahn, S. K. & Lee, Y. D. Effect of alloying elements and residuals on corrosion resistance of type 444 stainless steel. *Corrosion* **55**, 187–199 (1999).
60. Herbsleb, G. & Westerfeld, K. J. Der Einfluß von Stickstoff auf die korrosionschemischen Eigenschaften lösungsgeglühter und angelassener austenitischer 18/10 Chrom-Nickel-und 18/12 Chrom-Nickel-Molybdän-Stähle III. Loch-und Spannungsrißkorrosionsverhalten in wässerigen Lösungen von NaCl. *Materials and Corrosion* **27**, 479–486 (1976).
61. Sweitzer, J. E., Shiflet, G. J. & Scully, J. R. Localized corrosion of Al$_{90}$Fe$_5$Gd$_5$ and Al$_{87}$Ni$_{8.7}$Y$_{43}$ alloys in the amorphous, nanocrystalline and crystalline states: Resistance to micrometer-scale pit formation. *Electrochimica Acta* **48**, 1223–1234 (2003).
62. Sugimoto, K. & Sawada, Y. The role of molybdenum additions to austenitic stainless steels in the inhibition of pitting in acid chloride solutions. *Corrosion Science* **17**, 425–445 (1977).
63. Mischler, S., Vogel, A., Mathieu, H. J. & Landolt, D. The chemical composition of the passive film on Fe$_{24}$Cr and Fe$_{24}$Cr$_{11}$Mo studied by AES, XPS and SIMS. *Corrosion Science* **32**, 925–944 (1991).
64. Landolt, D., Mischler, S., Vogel, A. & Mathieu, H. J. Chloride ion effects on passive films on FeCr and FeCrMo studied by AES, XPS and SIMS. *Corrosion Science* **31**, 431–440 (1990).
65. Sweitzer, J. E., Scully, J. R., Bley, R. A. & Hsu, J. W. Nanocrystalline Al$_{87}$Ni$_{8.7}$Y$_{4.3}$ and Al$_{90}$Fe$_5$Gd$_5$ alloys that retain the localized corrosion resistance of the amorphous state. *Electrochemical and Solid-State Letters* **2**, 267–270 (1999).
66. Wegrelius, L., Falkenberg, F. & Olefjord, I. Passivation of Stainless Steels in Hydrochloric Acid. *Journal of The Electrochemical Society* **146**, 1397–1406 (1999).
67. Moshier, W. C., Davis, G. D. & Cote, G. O. Surface Chemistry of Sputter-Deposited Al-Mo and Al-Cr Alloys Polarized in 0.1 N KCl. *Journal of The Electrochemical Society* **136**, 356–362 (1989).
68. Lu, Y. C., Clayton, C. R. & Brooks, A. R. A bipolar model of the passivity of stainless steels-II. The influence of aqueous molybdate. *Corrosion Science* **29**, 863–880 (1989).
69. Ives, M. B., Akano, U. G., Lu, Y. C., Ruijin, G. & Srivastava, S. C. Corrosion behaviour of molybdenum-implanted stainless steel. *Corrosion Science* **31**, 367–376 (1990).
70. Lizlovs, E. A. Anodic Polarization Behavior of High-Purity 13 and 18% Cr Stainless Steels. *Journal of The Electrochemical Society* **122**, 719 (1975).
71. Kim, J. G. & Buchanan, R. A. Pitting and crevice corrosion of iron aluminides in a mild acid-chloride solution. *Corrosion* **50**, 658–668 (1994).
72. Newman, R. C. The dissolution and passivation kinetics of stainless alloys containing molybdenum-II. Dissolution kinetics in artificial pits. *Corrosion Science* **25**, 341–350 (1985).
73. Moshier, W. C., Davis, G. D., Ahearn, J. S. & Hough, H. F. Corrosion Behavior of Aluminum-Molybdenum Alloys in Chloride Solutions. *Journal of The Electrochemical Society* **134**, 2677–2684 (1987).
74. Eckenrod, C., J. & Kovach. Properties of Austenitic Stainless Steels and Their Weld Metals (Influence of Slight Chemistry Variations). In *ASTM Special Technical Publication*, 679 (1979).

75. Truman, J. E., Coleman, M. J. & Pirt, K. R. Note on the influence of nitrogen content on the resistance to pitting corrosion of stainless steels. *British Corrosion Journal* **12**, 236–238 (1977).
76. Wilde, B. E. Influence of silicon on the pitting corrosion resistance of an 18Cr-8Ni stainless steel. *Corrosion* **42**, 147–151 (1986).
77. Bui, N., Irhzo, A., Dabosi, F. & Limouzin-Maire, Y. on the Mechanism for Improved Passivation By Additions of Tungsten To Austenitic Stainless Steels. *Corrosion* **39**, 491–496 (1983).
78. Ahn, M. K., Kwon, H. S. & Lee, H. M. Quantitative comparison of the influences of tungsten and molybdenum on the passivity of Fe-29Cr ferritic stainless steel. *Corrosion Science* **40**, 307–322 (1998).
79. McCafferty, E. The electrode kinetics of pit initiation on aluminum. *Corrosion Science* **37**, 481–492 (1995).
80. Goetz, R., Laurent, J. & Landolt, D. The influence of minor alloying elements on the passivation behaviour of iron-chromium alloys in HCl. *Corrosion Science* **25**, 1115–1126 (1985).
81. Kim, J. S. & Kwon, H. S. Effects of tungsten on corrosion and kinetics of sigma phase formation of 25% chromium duplex stainless steels. *Corrosion* **55**, 512–521 (1999).
82. Kesten, M. Dissolution of Nickel in the Presence of Hydrogen Sulfide and Its Consequences on the Mechanism of Pit Nucleation. *Corrosion* **32**, 94–98 (1976).
83. Frankel, G. S. On the Pitting Resistance of Sputter-Deposited Aluminum Alloys. *Journal of The Electrochemical Society* **140**, 2192 (1993).
84. Böhni, H. & Uhlig, H. H. Effect of alloyed Re on the critical pitting potentials of 18%Cr/10%Ni stainless steels. *Corrosion Science* **9**, 353–355 (1969).
85. Janik-Czachor, M., Lunarska, E. & Szklarska-Smialowska, Z. Effect of nitrogen content in a 18Cr-5Ni-10Mn stainless steel on the pitting susceptibility in chloride solutions. *Corrosion* **31**, 394–398 (1975).
86. Lunarska, E., Szklarska-Simialowska, Z. & Janik-Czachor, M. Susceptibility of Cr-Ni-Mn stainless steels to pitting in chloride solutions. *Corrosion* **31**, 231–234 (1975).
87. Goldman, M. E., Ünlü, N., Shiflet, G. J. & Scully, J. R. Selected corrosion properties of a novel amorphous Al-Co-Ce alloy system. *Electrochemical and Solid-State Letters* **8** (2005).
88. Bandy, R. & Van Rooyen, D. Pitting-Resistant Alloys in Highly Concentrated Chloride Media. *Corrosion* **39**, 227–236 (1983).
89. Wong, F. *The effect of alloy composition on the localized corrosion behavior of NiCrMo alloys*. Ph.D. thesis, The Ohio State University (2009).
90. Halada, G. P., Kim, D. & Clayton, C. R. Influence of Nitrogen on Electrochemical Passivation of High-Nickel Stainless Steels and Thin Molybdenum-Nickel Films. *Corrosion (Houston)* **52**, 36–46 (1996).
91. Perren, R. A. *et al.* Corrosion resistance of super duplex stainless steels in chloride ion containing environments: Investigations by means of a new microelectrochemical method. II. Influence of precipitates. *Corrosion Science* **43**, 727–745 (2001).
92. Sun, Y. T., Wang, J. M., Jiang, Y. M. & Li, J. A comparative study on potentiodynamic and potentiostatic critical pitting temperature of austenitic stainless steels. *Materials and Corrosion* **69**, 44–52 (2018).
93. Machin, R. & Guha, P. The use of potentiostatic techniques in the development of improved stainless steels for chemical plant. *Materials and Corrosion* **25**, 40–45 (1974).
94. Clark, C. A. & Guha, P. Improvements in corrosion resistance, mechanical properties, and weldability of duplex austenitic/ferritic steels. *Materials and Corrosion* **34**, 27–31 (1983).
95. Sedriks, A. J. *Corrosion of stainless steels* (Wiley, 1979).
96. A Rohatgi. WebPlotDigitizer 4.2 - Extract data from plots, images, and maps (2019).
97. Frankel, G. S. Pitting Corrosion of Metals: A Review of the Critical Factors. *Journal of The Electrochemical Society* **145**, 2186–2198 (1998).
98. Pourbaix, M. *Atlas of electrochemical equilibria in aqueous solutions* (National Association of Corrosion Engineers, 1974).
99. Li, T., Scully, J. R. & Frankel, G. S. Localized Corrosion: Passive Film Breakdown vs Pit Growth Stability: Part II. A Model for Critical Pitting Temperature. *Journal of The Electrochemical Society* **165**, C484–C491 (2018).
100. Zhang, X. L., Jiang, Z. H., Yao, Z. P., Song, Y. & Wu, Z. D. Effects of scan rate on the potentiodynamic polarization curve obtained to determine the Tafel slopes and corrosion current density. *Corrosion Science* **51**, 581–587 (2009).
101. Nyby, C. *et al*. Electrochemical Metrics for Corrosion Resistant Alloys. *Citrination* https://citrination.com/datasets/192297 (2020).
102. Nyby, C. *et al*. Electrochemical Metrics for Corrosion Resistant Alloys. *figshare* https://doi.org/10.6084/m9.figshare.13038257 (2020).
103. Michel, K. & Meredig, B. Beyond bulk single crystals: A data format for all materials structure-property-processing relationships. *MRS Bulletin* **41**, 617–622 (2016).
104. Yeh, J. W. Alloy design strategies and future trends in high-entropy alloys. *JOM* **65**, 1759–1771 (2013).
105. Miracle, D. B. High-Entropy Alloys: A Current Evaluation of Founding Ideas and Core Effects and Exploring "Nonlinear Alloys". *JOM* **69**, 2130–2136 (2017).
106. Ward, L., Agrawal, A., Choudhary, A. & Wolverton, C. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Computational Materials* (2016).
107. Ward, L. *et al*. Matminer: An open source toolkit for materials data mining. *Computational Materials Science* (2018).
108. Marcus, P. On some fundamental factors in the effect of alloying elements on passivation of alloys. *Corrosion Science* **36**, 2155–2158 (1994).

Acknowledgements

Author contributions

Clara Nyby co-wrote the manuscript. Xiaolei Guo oversaw the database construction and co-wrote the manuscript. James E. Saal aided in the writing of the manuscript. Szu-Chia Chien extracted and validated data from research articles. Angela Y. Gerard extracted and validated data from research articles. Huibin Ke extracted and validated data from research articles. Tianshu Li extracted and validated data from research articles. Pin Lu extracted and validated data from research articles. Christian Oberdorfer extracted and validated data from research articles. Sarita Sahu extracted and validated data from research articles. Sirui Li extracted and validated data from research articles. Christopher D. Taylor oversaw the database construction. Wolfgang Windl oversaw the database construction. John R. Scully provided technical expertise and reviewed the manuscript. Gerald S. Frankel oversaw the construction of the database and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.E.S. or G.S.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.