CHEMICAL ENGINEERING TRANSACTIONS

VOL. 99, 2023

A publication of
AIDIC
The Italian Association
of Chemical Engineering
Online at www.cetjournal.it

Guest Editors: Sauro Pierucci, Flavio Manenti

ISBN 978-88-95608-98-3; **ISSN** 2283-9216

DOI: 10.3303/CET2399033

A Machine Learning Approach to Predict the Materials' Susceptibility to Hydrogen Embrittlement

Alessandro Campari[a,*], Maryam Darabi[a], Antonio Alvaro[b], Federico Ustolin[a], Nicola Paltrinieri[a]

[a] Department of Mechanical and Industrial Engineering, NTNU, Trondheim, Norway
[b] SINTEF Industry, SINTEF, Trondheim, Norway
alessandro.campari@ntnu.no

Hydrogen is widely considered a promising energy carrier capable of mitigating human environmental impact. Nevertheless, safety aspects represent one of the major bottlenecks for the widespread utilization of hydrogen technologies. Industrial equipment operating in hydrogen environments is prone to hydrogen-induced damages, which may manifest through a reduction of mechanical properties, fracture toughness, and fatigue performance. They may cause component failures at stress levels significantly below the design level, therefore determining loss of containment. The occurrence of hydrogen embrittlement (HE) relies on the synergy of several factors, such as hydrogen concentration, operating conditions, level of internal and applied stress, microstructure and chemical composition of the material. However, the interlinked dependence of these factors makes a direct and clear evaluation challenging, subsequently creating serious difficulties in planning inspection and maintenance activities. In this study, a comprehensive review of the experimental data of tensile tests carried out in hydrogen was performed and analyzed through an advanced machine learning approach. This study can provide critical insights into the susceptibility to hydrogen embrittlement for several materials operating under different environmental conditions. In particular, the Embrittlement Index was estimated and used as determining parameter to predict the likelihood of component failures. The model demonstrated accurate and reliable predicting capabilities. The outcome of this study can increase the understanding of hydrogen-induced material damages and facilitate decision-making processes in planning the inspection and maintenance of hydrogen technologies.

1. Introduction

The European Commission indicated hydrogen as a promising energy carrier to mitigate the issue of global warming and make the countries energetically self-sufficient in the long term (Campari et al., 2022). It can be produced from water electrolysis or methane steam reforming coupled with carbon capture and storage with reduced environmental impact (Ustolin et al., 2022). In addition, hydrogen can be efficiently used in fuel cell systems or directly burned in thermal engines with near-zero pollutant emissions. Nevertheless, hydrogen production, transport, and storage entail some safety concerns related to the extreme combustion properties of this substance and its capability of permeating and embrittling most metallic materials. Hydrogen embrittlement is a long-known phenomenon which often manifests as a degradation in materials' mechanical properties and cracks resistance properties under monotonic or cyclic loading. HE can particularly reduce the ductility of metals and this can lead to component failures under stress levels below the nominal tensile strength. Storage tanks, fasteners, process reactors, pipelines, fuel cell vehicles, and aircraft components are likely to be affected by HE; this can result in sudden equipment failures with potentially severe consequences and significant financial losses (Campari et al., 2023). For this reason, several codes and standards have been developed over the years to facilitate the selection of metallic materials for hydrogen service. For instance, the Technical Report ISO/TR 15916 provides a list of metals suitable for hydrogen applications based on their low HE susceptibility (ISO, 2015). The European Industrial Gases Association (EIGA) has produced Codes of Practice on material

Paper Received: 7 December 2022; Revised: 19 March 2023; Accepted: 2 May 2023
Please cite this article as: Campari A., Darabi M., Alvaro A., Ustolin F., Paltrinieri N., 2023, A Machine Learning Approach to Predict the Materials' Susceptibility to Hydrogen Embrittlement, Chemical Engineering Transactions, 99, 193-198 DOI:10.3303/CET2399033

testing and selection, which divide materials for H_2 service into different classes based on compositional and strength ranges (EIGA, 2011, 2014). Nevertheless, the material classes and the classification methods need to be updated based on the most recent data describing their performances under realistic operating conditions. Given this background, this study aims to develop a machine-learning approach capable of identifying the best-suited materials for hydrogen service and predicting their performances under different operating conditions. As an initial user case, the model has been trained with the results of experimental tests carried out in compressed gaseous hydrogen environments in compliance with the testing methodology prescribed in ISO 11114-4 (ISO, 2017) and published in international scientific journals and technical reports.

2. Hydrogen embrittlement

Hydrogen environment embrittlement is a material degradation resulting from the combined effect of stress (residual or applied), susceptible metallurgy, and hydrogen environment. Hydrogen molecules are dissociated into atoms and absorbed into the metal lattice; then, atoms diffuse in regions with high triaxial stress, locally affecting the mechanical properties of the material (tensile and fracture properties and fatigue performance). Figure 1 shows a summary of the main parameters influencing HE and a simplified schematic of the processes involved in the actualization of HE degradation.

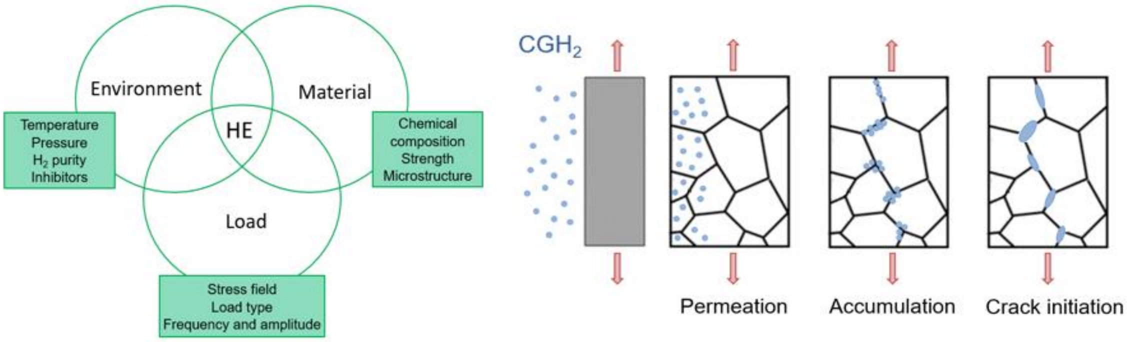

Figure 1: Hydrogen embrittlement resulting from the combination of several factors

When a component operating in a compressed gaseous hydrogen (CGH_2) environment undergoes tension, the stress concentration gradient generated in the proximity of heterogeneities can cause the migration of absorbed hydrogen in defects and inclusions. It provokes local embrittlement and eventually facilitates the crack initiation from the metal surface or stress concentrators (e.g., notches). The crack progressively grows due to a further hydrogen migration to the crack tip, until a critical load is reached, and fracture occurs. The applied load can be roughly divided into two categories: quasi-static and cyclic. The former refers to a stress level which is constant or varies slowly enough to maintain a steady-state hydrogen distribution into the metal lattice, while the latter is associated with dynamic components or undesired vibrations in static equipment. The material resistance against quasi-static loads is mainly defined by its tensile properties and its fracture toughness. The ductility loss, tested through slow strain rate tests (SSRT) or linearly increasing stress tests (LIST), is a common effect of HE and may represent a serious concern for static pressurized components (Hagen and Alvaro, 2020). A common parameter used to quantify the degree of susceptibility against HE for a given material is the Embrittlement Index (EI), which is expressed as the relative change of reduced areas obtained when a tensile test is performed in a hydrogenated environment and in a reference environment (e.g., in air):

$$\mathrm{EI} = \frac{\mathrm{RA_{air}} - \mathrm{RA_{H_2}}}{\mathrm{RA_{air}}} \cdot 100 = \frac{[(\mathrm{A_i} - \mathrm{A_f})/\mathrm{A_f}]_{\mathrm{air}} - [(\mathrm{A_i} - \mathrm{A_f})/\mathrm{A_f}]_{\mathrm{H_2}}}{[(\mathrm{A_i} - \mathrm{A_f})/\mathrm{A_f}]_{\mathrm{air}}} \cdot 100 \qquad (1)$$

where RA_{air} and RA_{H2} represent the reduced area at fracture in air and hydrogen atmospheres, respectively, while A_i and A_f represent the initial and final fracture areas, respectively. The higher the EI, the higher will be the difference in the reduced area in air and in an H_2 atmosphere, and consequently the HE susceptibility of the material. In most ferritic steels, HE causes a loss of RA ranging from 20% to 50% for smooth specimens and up to 80% for notched specimens (NASA, 2016). The triaxial stress region ahead of the notch causes a more significant hydrogen accumulation, which enhances the HE effects.

As a rule of thumb, the hydrogen concentration in the metal lattice is proportional to the hydrogen partial pressure (according to Sievert's law), and the same goes for the H_2-induced detriment in tensile properties. A saturation level can be observed so that further pressure increases above a certain level (the specific value depends on the material) does not further affect ductility. On the other hand, the temperature effect is more complex since it affects the kinetic of surface reactions, hydrogen solubility, diffusivity, and trapping contemporarily. At high temperatures, atoms' mobility is high and de-trapping is favored, while, at low temperatures, the diffusion through the metal lattice is slow and longer exposure times are required to obtain the critical combination of local hydrogen concentration and local stress for embrittlement to be displayed. As a result, HE impact reaches a maximum at near-ambient temperature for a wide variety of ferritic steels (Gangloff and Somerday, 2012). Finally, materials metallurgy is also considered an overriding factor for the occurrence of hydrogen embrittlement. High-strength steels are normally more prone to HE, and certain microstructures (e.g., martensite) are more susceptible than others (e.g., austenite) even at the same strength level. Moreover, the chemical composition represents an important parameter: aluminum and copper alloys can always be used, along with austenitic steels with Ni content above 7% (San Marchi and Somerday, 2012). Finally, the testing conditions influence the HE severity. In the case of tensile tests, the nominal strain rate is the most important factor. It should be low enough to allow hydrogen to diffuse through the metal lattice and accumulate near defects and inclusions.

3. Methodology

The methodology adopted for classifying the materials based on their hydrogen embrittlement susceptibility can be divided into three main steps: database creation and pre-processing, target definition, and machine learning simulation. It is described in detail in the following sections.

3.1 Database creation and pre-processing

The database contains information about tensile test results, where the specimens were subjected to stress in combination with exposure to high-pressure hydrogen gas. The features include environmental parameters (i.e., testing environmental conditions), material properties (i.e., metallurgy and characterization of the specimen), and loading conditions (i.e., nominal strain rate and presence of stress concentrators). All these data have been retrieved from the literature. Among the others, the "Technical Reference for Hydrogen Compatibility of Materials" (San Marchi and Somerday, 2012) represents the primary source of information. The structure of the original database is summarized in Table 1. Considering that most tests have been performed under standard hydrogen purity level (i.e., 99.99%), the correspondent feature has not been deemed relevant for the analyses and has been discarded. Moreover, the features' values have been converted to have a unique unit for each physical quantity. In this case, the normalization of the numerical values was not deemed useful since no scale effects were observed. The pre-processed database included 200 tests (i.e., rows) and 24 features (i.e., columns).

Table 1: Structure of the database (EB: Electron Beam; ERW: Electric Resistance Weld; GMA: Gas Metal Arc; GTA: Gas Tungsten Arc; HAZ: Heat-Affected Zone; SA: Submerged Arc; SMA: Shielded Metal Arc)

Factor	Feature	Feature type
Environment	Pressure (MPa)	Numerical
	Temperature (°C)	Numerical
	Hydrogen purity (%)	Numerical
Material	Material: X80, X100, A515, A516, etc.	Categorical
	Composition: Fe, Cr, Ni, Mn, Mo, Nb, Ti, V, Al, Cu, Si, C, N, B, S, P, Co (%)	Numerical
	Microstructure: ferritic, pearlitic, bainitic, martensitic, austenitic	Categorical
	Type of weld or HAZ: EB, ERW, GMA, GTA, SA, SMA	Categorical
	Yield strength (MPa)	Numerical
	Ultimate tensile strength (MPa)	Numerical
Load	Stress intensity factor	Numerical
	Nominal strain rate (s^{-1})	Numerical

3.2 Target identification

The susceptibility to HE can be determined through the effects of hydrogen on material tensile properties. These effects can be quantified by comparing the elongation or the reduced area in the air with those in H_2, thus obtaining the elongation reduction or the EI. The latter parameter was chosen as target. Hence, all the tests which did not provide the information to calculate the EI were excluded. Two susceptibility classes were defined based on the EI value, according to the classification provided in the technical report NASA/TM-2016-218602 (NASA, 2016). These classes are reported in Table 2, together with recommendations regarding the use of materials for hydrogen applications. The susceptibility classes "Small", "Medium", and "High" have been merged since the small range of testing conditions does not allow a narrower classification.

Table 2: Hydrogen embrittlement susceptibility classes

Susceptibility	Embrittlement index	Material screening notes
Small / Medium / High (SMH)	$EI < 50\%$	Recommended for hydrogen service at specific operating conditions and with tailored inspection and maintenance activities
Extreme (E)	$EI \geq 50\%$	Not recommended for hydrogen applications in any temperature and pressure range

3.3 Machine learning simulation

A Gradient Boosting Machine (GBM) classification model has been trained and evaluated on the materials' susceptibility database. The purpose of the algorithm is to classify the materials into two categories: "Small / Medium / High" (SMH), and "Extreme" (E) hydrogen embrittlement susceptibility. In other words, the model should predict if a certain type of material will be affected by hydrogen-induced ductility loss under specific environmental and loading conditions to an extent that makes its utilization inherently unsafe. This supervised learning method was employed for a classification task rather than a regression because the punctual value of the EI depends upon a variety of factors (e.g., the presence of internal defects and inclusions, or the grain refinement and orientation), which can neither be included in a database nor considered to infer general recommendations regarding the materials' susceptibility to HE. The flow diagram to set up and perform the GBM classification is illustrated in Figure 2.

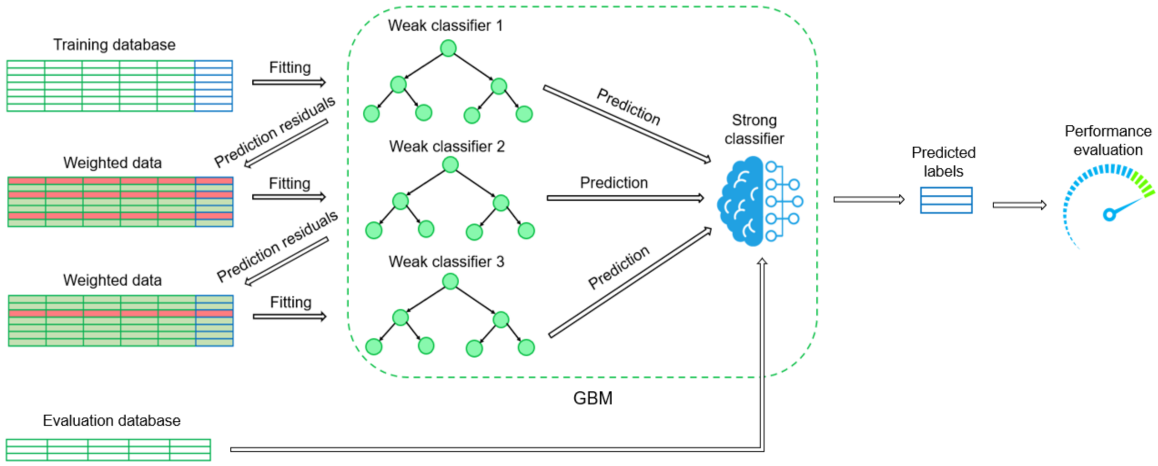

Figure 2: Training and evaluation processes of the Gradient Boosting classifier

The classification process is based on two steps: training and evaluation. The data are divided into two sub-databases, namely the training and the evaluation database, in a ratio of 70:30. The former is fed to the Gradient Boosting Machine. This model is composed of a weak learner, a loss function, and an additive model. The weak learner is a decision tree where leaves represent class labels, and branches represent the correlations between features that lead to those class labels; it is normally characterized by a high error rate (Bramer, 2020). The loss function is a differentiable function which should estimate how good the model is in making predictions with the given data. A logarithmic loss has been selected for this binary classification problem. The additive model is the iterative approach of adding one tree after each iteration, thus progressively reducing the value of the loss function. After each iteration, the labels correctly predicted are kept, while the others are stripped out. A new decision tree is created and tested on the dataset incorrectly classified, and then only the examples successfully labeled are kept. Every time a new weak learner is added, the weights of the pre-existing trees are left unchanged. The loss is calculated, and gradient descent is performed to reduce its value. Then, the tree parameters are modified to reduce the residual loss. In other words, at each iteration, GBM fits the new decision tree to the residual errors made by the previous weak learners instead of fitting it on the data. This process is repeated until the loss is reduced below a certain threshold or a certain number of trees is reached (Mason et al., 1999).

4. Results and discussion

After the model evaluation, three metrics have been calculated to assess the performance of the trained Gradient Boosting Machine: accuracy, precision, and recall. The former expresses the fraction of correct predictions and can be calculated through Eq(2). Precision, estimated through Eq(3), indicates the fraction of true positive predictions, while recall, calculated through Eq(4), refers to the fraction of positive labels correctly predicted (Seliya et al., 2009).

$$\mathrm{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} = 0.886 \tag{2}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} = 0.860 \tag{3}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} = 0.889 \tag{4}$$

where TP indicates true positive results (i.e., both real and predicted labels are SMH), TN the true negatives (i.e., both real and predicted labels are E), FP the false positives (i.e., the real label is E, and the predicted label is SMH), and FN the false negatives (i.e., the real label is SMH, and the predicted label is E). These metrics have been calculated using a 0.5 default probability decision threshold. Nevertheless, this might not be the value leading to the best metrics. The precision-recall curve can be obtained by varying from 0 to 1 the probability threshold associated with the GBM binary classifier. The area under the curve (AUC) represents another important performance metric, ranging from 0 to 1. The higher the AUC value, the better the classifier's performance. The confusion matrix in Figure 3 shows the percentages of TP, TN, FP, and FN predictions, while the precision-recall curve shows the probability threshold set at 0.5.

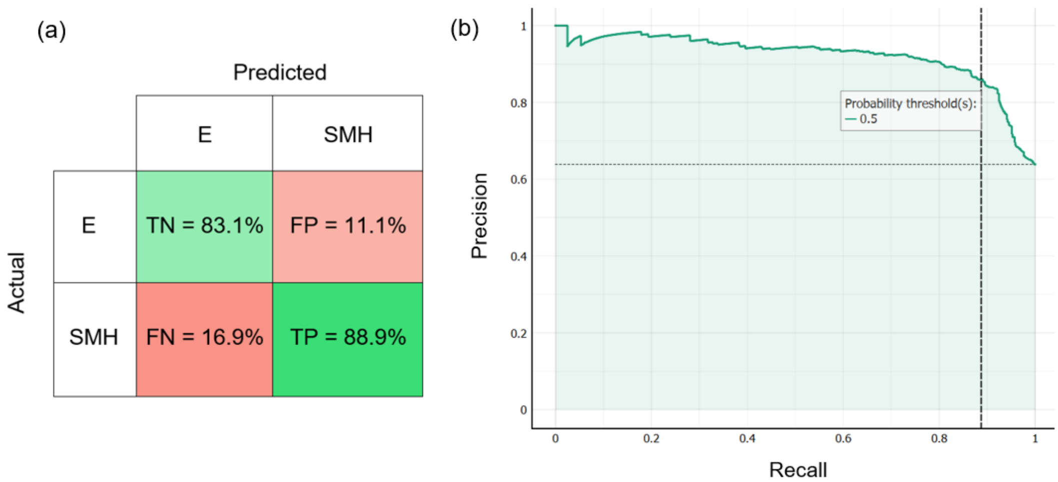

Figure 3: (a) Confusion matrix (SMH: Small/Medium/High susceptibility; E: Extreme susceptibility) and (b) Precision-recall curve for the GBM classifier

The confusion matrix in Figure 3(a) shows how most of the tests are correctly labeled, and the percentages of true positives and true negatives are comparable. Despite this, the model still tends to mislabel a significant number of tests (16.9% of false negatives and 11.1% of false positives). It should be noted that mislabeling an "Extreme" susceptibility as a "Small / Medium / High" is more critical than vice versa. This wrong prediction may cause an underestimation of the HE susceptibility of metal, thus leading to an improper material selection for specific working conditions and a higher risk of component failure. In this perspective, the number of false positive results should be minimized; thus, precision is the most meaningful metric to consider in this specific task. It is possible to maximize precision at the expense of recall by increasing the decision threshold. Nevertheless, Figure 3(b) shows how this modification has only minor benefits on the precision optimization, but it has high costs in terms of a decrease in recall. The high value of AUC (i.e., 0.922) and the nearly flat precision-recall curve confirm this observation. When evaluating this classification model's overall performance, the dataset's dimensions should be considered. The limited amount of data is mainly due to the selective inclusion criteria. The HE effects are taken into account through the Embrittlement Index only, thus excluding all the tests expressed in terms of elongation reduction. In addition, only tensile tests in compressed gaseous hydrogen, without pre-charging, were considered. On the one hand, this restrictive selection significantly reduces the size of the dataset with which the ML model is trained, but on the other limits the risk of troubling the waters with unmeaningful information or inhomogeneous data. Analyzing in depth the predictions made by the model, the mislabeled targets did not appear to be preferentially associated with one material rather than with another. Hence, it is reasonable to expect that more accurate predictions will be achieved by increasing the size of the training dataset and including measurements obtained under various testing conditions (for instance, only a few tests were performed at temperatures other than 22 °C, or in the presence of gaseous impurities and inhibitors).

5. Conclusions

A machine learning approach has been used to classify metallic materials for hydrogen applications into two macro-categories: extremely susceptible and potentially usable. The HE effect on materials exposed to compressed gaseous hydrogen is measured through the Embrittlement Index, hence only the hydrogen-induced reduction in tensile properties is considered. The model, based on the Gradient Boosting algorithm, can predict the HE severity with 88.6% accuracy. The synergistic interplay of environmental conditions, material properties, and loading conditions is inherently complex and can occasionally result in inaccurate predictions. Nevertheless, it will be possible to obtain better performance by increasing the size of the database. In the future, a complete assessment of the reliability of steel exposed to compressed hydrogen gas will consider the material performance under dynamic loads and the hydrogen-enhanced fatigue crack growth rate.

Acknowledgments

This work was undertaken as a part of the research projects H_2CoopStorage and SH_2IFT-2, and the authors would like to acknowledge the financial support of the Nordic Energy Research within the program ERA-Net Smart Energy System (Grant No. 106291) and of the Research Council of Norway (Grant No. 327009).

References

Bramer M., 2020, Principles of Data Mining: 4th Edition, Springer Science, London, UK.

Campari A., Pio G., Ustolin F., Paltrinieri N., Salzano E., 2022, Design and Optimization of an Emergency Auto-thermal Burner for Liquid Hydrogen, Chem. Eng. Trans., 91, 325-330, DOI: 10.3303/CET2291055.

Campari A., Nakhal A.J., Ustolin F., Alvaro A., Ledda A., Agnello P., Moretto P., Patriarca R., Paltrinieri N., 2023, Lessons Learned from HIAD 2.0: Inspection and Maintenance to Avoid Hydrogen-Induced Material Failures, Comput. Chem. Eng., 173, DOI: 10.1016/j.compchemeng.2023.108199.

EIGA, 2011, IGC Doc 100/11/E – Hydrogen cylinders and transport vessels.

EIGA, 2014, IGC Doc 121/14 – Hydrogen pipeline systems.

Gangloff R., Somerday B., 2012, Gaseous Hydrogen Embrittlement of Materials in Energy Technologies: The Problem, its Characterisation and Effects on Particular Alloy Classes, Woodhead Publishing Series, Philadelphia, USA.

Hagen A.B., Alvaro A., 2020, Hydrogen Influence on Mechanical Properties in Pipeline Steel, Norway.

ISO, 2015, ISO/TR 15916 – Basic consideration for the safety of hydrogen systems.

ISO, 2017, ISO 11114 – Test methods for selecting steels resistant to hydrogen embrittlement.

Mason L., Baxter J., Bartlett P., Frean M., 1999, Boosting Algorithms as Gradient Descent, Advances in Neural Information Processing Systems (NIPS 1999), Denver, USA.

NASA, 2016, NASA/TM-2016–218602 – Hydrogen Embrittlement, Huntsville, USA.

San Marchi C., Somerday B.P., 2012, Technical Reference for Hydrogen Compatibility of Materials, USA.

Seliya N., Khoshgoftaar T.M., Van Hulse J., 2009, A Study on the Relationships of Classifier Performance Metrics, Proceedings of the 21st IEEE ICTAI, Newark, USA, 59–66, DOI: 10.1109/ICTAI.2009.25.

Ustolin F., Campari A., Taccani R., 2022, An Extensive Review of Liquid Hydrogen in Transportation with Focus on the Maritime Sector, J. Mar. Sci. Eng., 10, 1222, DOI: 10.3390/jmse10091222.