



This is an electronic reprint of the original article. This reprint may differ from the original in pagination and typographic detail.

Fangnon, Eric; Malitckii, Evgenii; Latypova, Renata; Vilaça, Pedro

# Prediction of hydrogen concentration responsible for hydrogen-induced mechanical failure in martensitic high-strength steels

Published in: International Journal of Hydrogen Energy

DOI: 10.1016/j.ijhydene.2022.11.151

Published: 15/02/2023

Document Version Publisher's PDF, also known as Version of record

Published under the following license: CC BY

Please cite the original version:

Fangnon, E., Malitčkii, E., Latypova, R., & Vilaça, P. (2023). Prediction of hydrogen concentration responsible for hydrogen-induced mechanical failure in martensitic high-strength steels. *International Journal of Hydrogen Energy*, *48*(14), 5718-5730. https://doi.org/10.1016/j.ijhydene.2022.11.151

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.



Available online at www.sciencedirect.com

# **ScienceDirect**



journal homepage: www.elsevier.com/locate/he

# Prediction of hydrogen concentration responsible for hydrogen-induced mechanical failure in martensitic high-strength steels



# Eric Fangnon <sup>a,\*</sup>, Eugenii Malitckii <sup>a</sup>, Renata Latypova <sup>b</sup>, Pedro Vilaça <sup>a</sup>

<sup>a</sup> Department of Mechanical Engineering, School of Engineering, Aalto University, P.O Box 11000, FI-00076, Espoo, Finland

<sup>b</sup> Faculty of Technology, Materials and Mechanical Engineering, Centre for Advanced Steels Research (CASR), University of Oulu, P.O Box 4200, FI-90014, Oulu, Finland

#### HIGHLIGHTS

- Tensile testing under continuous hydrogen-charging.
- Performance of martensitic ultrahigh-strength steels in hydrogenated conditions.
- Development of deep learning artificial neural network models.
- Prediction of hydrogen concentration responsible for hydrogeninduced steel failure.

# ARTICLE INFO

Article history: Received 27 September 2022 Received in revised form 8 November 2022 Accepted 14 November 2022 Available online 30 November 2022

Keywords: Steel Hydrogen embrittlement Hydrogen concentration Deep learning Artificial neural network

# GRAPHICAL ABSTRACT



# ABSTRACT

Hydrogen, at critical concentrations, responsible for hydrogen-induced mechanical property degradation cannot yet be estimated beforehand and can only be measured experimentally upon fracture with specific specimen sizes. In this work, we develop two deep learning artificial neural network (ANN) models with the ability to predict hydrogen concentration responsible for early mechanical failure in martensitic ultra-high-strength steels. This family of steels is represented by four different steels encompassing different chemical compositions and heat treatments. The mechanical properties of these steels with varying size and morphology of prior austenitic grains in as-supplied state and after hydrogen-induced failure together with their corresponding hydrogen charging conditions were used as inputs. The feed forward back propagation models with network topologies of 12-7-5-3-2-1 (I) and 14-7-5-3-2-1 (II) were validated and tested with unfamiliar data inputs. The models I and II show good hydrogen concentration prediction capabilities with mean absolute errors of 0.28, and 0.33 wt.ppm at test datasets, respectively. A linear correlation of 80% and 77%, between the experimentally measured and ANN predicted

\* Corresponding author.

E-mail address: eric.a.fangnon@aalto.fi (E. Fangnon).

https://doi.org/10.1016/j.ijhydene.2022.11.151

0360-3199/© 2022 The Author(s). Published by Elsevier Ltd on behalf of Hydrogen Energy Publications LLC. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

hydrogen concentrations, was obtained for Model I and II respectively. This shows that for this family of steels, the estimation of hydrogen concentration versus property degradation is a feasible approach for material safety analysis.

© 2022 The Author(s). Published by Elsevier Ltd on behalf of Hydrogen Energy Publications LLC. This is an open access article under the CC BY license (http://creativecommons.org/ licenses/by/4.0/).

# Introduction

The detrimental effect of hydrogen on the mechanical integrity of steels and alloys is a major industrial challenge. Atomic hydrogen enters metallic materials during manufacturing processes or in service through physisorption and chemisorption [1,2]. Being the smallest and lightest element, upon surface adsorption, the hydrogen is absorbed via diffusion within the lattice of steels and alloys and tends to be trapped at microstructural features like interstitials, grain boundaries, and interfaces [3–5]. During diffusion, hydrogen can be trapped reversibly or irreversibly by microstructural features and defects like dislocations, voids and precipitates depending on the local thermodynamic state [6–9]. Consequently, because of simultaneous action of high local stress state and hydrogen content above critical concentrations, early degradation of mechanical properties, leading to crack initiation and failure occurs [10-12]. This phenomenon is generally termed hydrogen embrittlement (HE). As yet, a real-time experimental observation of hydrogen-metal interaction leading to early failure is still difficult and the phenomenon is not fully understood. Nevertheless, several hydrogen-induced failure mechanisms have been proposed.

The mechanisms include hydrogen enhanced decohesison (HEDE) [13,14], hydrogen enhanced localised plasticity (HELP) [15], adsorption induced dislocation emission (AIDE) [16], hydrogen enhanced stress induced vacancies (HESIV) [17], stress-induced hydride formation and cleavage [18] and the newly proposed HEDE + HELP mechanism [19]. The underlying principle of HEDE mechanism is that the presence of hydrogen in the bulk material reduces any interatomic bond strength leading to decohesison and consequential failure at relatively lower stresses [13]. In reheated-quenched (RQ) and quenched and tempered (QT) martensitic steels, the HEDE mechanism is mostly responsible for failure where the fracture mode is intergranular along the prior austenitic grain (PAG) boundaries [20-22]. The fundament of the HELP mechanism is that hydrogen enhances the movement of dislocations resulting in dislocation pile-up. The movement of dislocations in turn facilitates hydrogen transport and increases its concentration at pile-ups. This mechanism is found to be responsible for the observed plastic strain localisation in hydrogen-induced brittle fractures and fracture modes transitioning from ductile to transgranular, quasi-cleavage and intergranular in martensitic steels [15,23]. Generally, for DQ and QT steels, the synergic activity of HEDE + HELP as emphasized by Djukic et al. [19] is seen in action when hydrogen, transported by dislocations reaches a critical concentration at pile-ups caused by dislocation barriers such as carbide and inclusion-matrix interfaces. At this critical concentration, hydrogen at interfaces instigates the HEDE mechanism leading to eventual early fracture [19].

Principally, the magnitude of steel' susceptibility to HE is proportional to their strength and hardness [24]. Thus, DQ and QT martensitic ultra-high strength are at higher risk of HE. In addition to alloying, the volume fraction and shape of retained austinite (RA) as well as the morphology of PAG [25,26] are parameters that influence martensitic ultra-high steels' susceptibility to hydrogen. Studies have shown that HE associated with RA is not only dependent on the volume fraction but also the shape. Zhu et al. [27] reported that massive hydrogeninduced cracks are found to initiate at the interface of blocky RA. On the other hand, steels with filmy RA are relatively more resistant to hydrogen-induced cracks. As for PAGs which serve as crack initiation points and preferred propagation paths, studies have shown that smaller PAGs reduce the sensitivity to hydrogen in DQ steels [8,28]. This can be associated not only with the reduced trapping ability at the grain boundaries but also with the prevention of hydrogen concentration from attaining critical concentrations sooner due to improved diffusion over a larger area [29].

Hydrogen embrittlement research has benefited from several investigative techniques over the years, including experimental methods and atomistic simulations [17,30-33]. These, help investigate the HE phenomenon from macroscopic to nanoscopic levels in variety of steels with distinmicrostructural compositions and guished service applications. Contemporary investigation of hydrogenassisted degradation on macroscopic levels is carried out by mechanical testing after hydrogen charging or under continuous hydrogen charging. The mechanical tests include constant extension rate tests (CERT), constant load tests (CLT) [11], impact toughness tests, four-point bend tests [33] and the recently proposed tuning-fork test [34]. The commonly used hydrogen charging techniques include the classical electrochemical charging [35], plasma charging [36] and hydrogen gas charging [37]. Worth to note here is that although the results obtained via these methods provide some scientific insight for better understanding of the HE phenomenon, they are often used only to evaluate the performance of the steel with the attempt to mimic service conditions. In other words, evaluating the influence of hydrogen on macroscopic mechanical properties. For distinguished steels, the extent of mechanical properties degradation during testing is not only dependent on the hydrogen charging conditions, but more importantly the hydrogen concentration in the steels [31,33]. Thus far, there are no swift means available for obtaining hydrogen

Table 1 – Chemical composition and mechanical properties of the steels used in this study.									
Chemical composition [wt. %]			Mechanical properties						
	DQ1	DQ2		DQ1	DQ2	A860	A960		
С	0.370	0.250	Ultimate tensile strength, (UTS) [MPa]	2200	1570	1734	1710		
Mn	0.299	0.250	Yield strength (YS) at 2% offset [MPa]	1800	1350	1250	1230		
S	0.001	0.002	Measured hardness [HRC] $\pm$ STD	57 ± 0.5	$51 \pm 0.5$	48 ± 1	$52 \pm 1$		
Al	0.430	0.095	Elongation at fracture [%]	13	12	13	14		

concentration values in steels after hydrogen-assisted failure except for complex experimental measurement using mercury, gas chromatography and thermal desorption spectroscopy (TDS) techniques. Even with these methods, the measurements need to be performed immediately after fracture with specimens of specific shapes and sizes which is a challenge for steels that fail in service. So far, the ability to predict hydrogen contents based on service conditions is poor. Meanwhile, this information could assist designers in estimating the probabilities and magnitude of hydrogen assisted failures.

In recent times artificial neural networks (ANN) find their growing application in several fields including research in engineering and material sciences [38-42]. Among many functions, ANNs mimicking the neural networks that constitute human brains, possess the ability to learn, identify trends and build mathematical relationships between the feature variables of input data to predict specific target outputs [43]. In many disciplines, the immense reposited data available from experimental results over the years are used to train ANN models to predict target outputs with little or no further experimental requirements. Hydrogen embrittlement research has benefited from ANN as well. Titus et al. [38] developed ANN model to predict the degradation of mechanical properties of aluminium alloys in the presence of hydrogen while considering the influence of different elemental compositions. In other works, ANN models were trained to use hydrogen thermal desorption spectroscopy data believed to contain microstructural information to predict HE index of ferritic, austenitic, and ferritic-martensitic steels with up 98% accuracy [44]. Despite the available works using ANN to predict mechanical performance of steels and alloys in the presence of hydrogen, there are no available models to predict the actual hydrogen concentrations resulting from a combination of hydrogen charging conditions and loading which are simultaneously responsible for performance degradation of the steel.

In this paper, we develop, and evaluate the reliability of two ANN models to predict total hydrogen concentration responsible for performance degradation in DQ and DQ + RQ steels by using data curated from a combination of mechanical testing in as-supplied state, and under continuous hydrogen charging. The hydrogen charging parameters are also considered as input. The success and implementation of this approach will assist designers in estimating the levels of hydrogen required to cause hydrogenassisted failure of hot-rolled martensitic ultra-highstrength steels, hence influencing the selection of material and service conditions.

# Materials and methods

# Materials

Two direct-quenched martensitic steels with distinct chemical compositions were used in this study. The original microstructures were achieved by hot rolling in the nonrecrystallization region followed by direct quenching (DQ1 and DQ2). In addition to the base materials, two additional microstructures were obtained by re-heating and quenching (RQ) of DQ2 at austenisation temperatures of 860 °C and 960 °C with 25 min holding time, followed by quenching in water-oil emulsion bath (DQ2+RQ at 860  $^\circ C$   $\equiv$  A860 and DQ2+RQ at 960 °C  $\equiv$  A960, respectively). The re-austenisation led to equiaxed PAG structures with different grain sizes. Table 1 shows the elemental composition of DQ1, DQ2 with their corresponding mechanical properties as well as those of A860 and A960. The different chemical compositions and different PAG structures are intended to evaluate the sensitivity to hydrogen uptake of the mechanical property degradation. The encompassment of these steels and different heat treatment conditions will provide a wider application range of the proposed model for hydrogen concentration prediction at failure, for materials of this same class.

The steels used in the present study are predominantly martensitic in microstructure. Fig. 1 shows the similarity in the lath-martensite microstructure and carbide morphology of the steels. On the contrary, the steels possess rather dissimilar PAG structures: DQ1 and DQ2 have elongated PAG structures (Fig. 2(a) and (b)) resulting from hot rolling below non-recrystallization temperature followed by directquenching; while A860 and A960 have equiaxed PAG structures (Fig. 2(c) and (d)) as a result of additional reaustenisation and quenching. The average PAG sizes for all test materials, calculated from all directions with linear intercept method, were 11.2  $\mu$ m for DQ1, 9.6  $\mu$ m for DQ2, 9.4  $\mu m$  and 42.6  $\mu m$  for A860 and A960 respectively. Besides the shape, emphasis should also be made on the significant difference in average PAG size between the A860 and A960, versus the original DQ2 condition. X-ray diffraction results revealed that the steels contain negligible volume fractions of RA (<1%).

#### Hydrogen charging and mechanical testing

Electrochemical hydrogen charging was performed with a three-electrode electrochemical cell coupled with a Gamry potentiostat framework. Calomel and platinum electrodes



Fig. 1 – Dominant martensitic microstructure observed in test steels via scanning electron microscopy (SEM) of (a) DQ1, (b) DQ2, (c) A860 and (d) A960.



Fig. 2 – Distinguished PAG morphology via optical microscopy (OM) of (a) DQ1, (b) DQ2, (c) A860, and (d) A960.

were used as reference and counter electrodes, respectively. The steel specimen was the working electrode. 30 g/l of sodium chloride (NaCl) and 2–6 g/l of ammonium thiocyanate (NH<sub>4</sub>SCN) were utilised as main electrolyte and atomic hydrogen recombination inhibitor, respectively. Prior to tensile tests under continuous hydrogen charging, the steel specimens were pre-charged for 2 h, which was experimentally determined to provide sufficient and homogeneous distribution of hydrogen across the thickness of the specimen's gauge section. Electrochemical potentials of  $-0.8 V_{SCE}$  to -1.4



Fig. 3 – Depiction of the experimental features with (a) schematic and dimensions of MT specimens (b) general setup CERT under continuous hydrogen charging.

 $V_{SCE}$  were applied for hydrogen charging. A combination of varied applied electrochemical potential and concentration of hydrogen recombination inhibitor was used to provide a wide range of hydrogen concentration in the studied steels. The pH of the electrolyte was measured to be 5.5 and 4.5 for the electrolyte solution containing 2 g/l and 3 g/l NH<sub>4</sub>SCN, respectively. During hydrogen charging, the electrolyte was kept under constant stirring at 40 rev/min and deaeration by nitrogen gas flow. Hydrogen charging was performed at room temperature.

Mechanical testing (MT) specimens with size of 5 mm  $\times$  10 mm  $\times$  300 mm and gauge part size of 1.0 mm  $\times$  5.0 mm  $\times$  20 mm, shown in Fig. 3(a) were wire cut with EDM and polished mechanically, finishing with emery paper No. 1200. The specimens were Teflon-taped exposing only the gauge part to the hydrogen charging environment. CERTs were conducted under continuous hydrogen charging on a 30 kN MTS benchtop tensile test machine at the strain rate of  $10^{-4}$  s<sup>-1</sup> until fracture. Fig. 3(b) shows a general view of mechanical testing setup with continuous hydrogen charging.

After fracture the gauge part of the specimen is cut to the characteristic size of 1 mm  $\times$  5 mm  $\times$  10 mm for hydrogen concentration measurement.

#### Hydrogen concentration measurement

Hydrogen concentration was measured with TDS technique. The extracted specimens after fracture were cleaned with distilled water followed by drying in helium gas flow to prevent the formation of moisture on the specimen surface. The measurement of partial pressure of hydrogen occurs in ultrahigh-vacuum (UHV) chamber (ultimate pressure is 10<sup>-9</sup> mbar) coupled with a mass spectrometer (SRS residual gas analyser RGA100). To keep the UHV chamber at the required pressure and reduce pumping time before measurement, the sample is first placed in an airlock compartment and pumped to an intermediate pressure of  $10^{-6}$  mbar. After which the specimen is transported to the UHV chamber for hydrogen partial pressure measurement (see Fig. 4). The total time from the specimen preparation to TDS measurement did not exceeded 10 min with effective dwelling time in airlock as 5 min (up to  $10^{-6}$  mbar) ensuring negligible amounts of hydrogen loss. All TDS measurements were performed at a heating rate of 10 °C/min. The total hydrogen concentration was calculated by integrating the area below the desorption rate versus temperature curve.

## Data processing

The data utilised as input for the developed ANN models consists of the inherent mechanical properties of the test steels, parameters for hydrogen charging, results from mechanical tensile testing in air and under continuous hydrogen charging and their corresponding measured hydrogen concentrations. Table 2 presents the features of the input data utilised for training, validation and testing of the developed ANN models. The mechanical properties were obtained from classical hardness and tensile tests of as-received specimens. The tensile toughness was obtained by integrating the area under the nominal stress-strain curve, up to the UTS, resulting from tensile tests. Hydrogen sensitivity parameter (HSP) is used to quantify the level of mechanical property degradation as a result of increased hydrogen concentration from mechanical loading under hydrogen charging [32]. HSP was calculated for the test steels in this study considering tensile



1-specimen transport system 2-primary vacuum chamber 3-specimen 4-UHV chamber 5-furnace 6-turbo-molecular pumps 7-data acquisition setup

Fig. 4 – Schematic view of TDS apparatus used for experimental hydrogen concentration measurement.

Table 2 — Features of the input data utilised in this study.					
Mechanical properties	1. Hardness (HRC)				
	2. Ultimate tensile strength (UTS)				
	3. Elongation to fracture (EF)				
	4. Tensile toughness (TT)				
Hydrogen charging parameters	5. Hydrogen charging time T <sub>ch</sub>	$\Rightarrow$ System energy (E <sub>s</sub> )			
	6. Applied electrochemical potential (E <sub>a</sub> )	(Equation (2))			
	7. Initial current (I <sub>i</sub> )				
	8. Concentration of NH <sub>4</sub> SCN (C <sub>p</sub> )				
Resulting variables after mechanical testing and hydrogen	9. Fracture stress (FS)				
concentration measurement	10. Elongation at fracture (EF <sub>h</sub> )				
	11. Tensile toughness (TT <sub>h</sub> )				
	12. Measured hydrogen concentration (H <sub>c</sub> )				
	13. Hydrogen sensitivity parameter (HSP)				

strength, elongation to fracture and tensile toughness using Equation (1).

$$HSP = \frac{M_{p_{air}} - M_{p_{H}}}{M_{p_{air}}} 100\%$$
(1)

Where:  $M_{p_{air}}$  and  $M_{p_H}$  are the mechanical properties of the test steels in as-received condition and after electrochemical hydrogen charging respectively, excluding the hardness. It is important to mention that only the absolute values of applied potentials and their corresponding currents were used.

The collected experimental data was split into training, validation, and testing datasets according to the 60-20-20 rule. To ensure the generalisation the model, a data augmentation method was applied to the training and validation datasets using the *stats.truncnorm.rvs* function from the statistics package in spicy, an open-source library built on python with NumPy extension [45]. With this, probabilities within a normal distribution are defined for each input parameter as may naturally occur in an experimental setting or in service. Thus, random variates of desired size are generated within a predetermined standard deviation considering experimental data point as a mean [46]. Table 3 shows the set standard deviation values attributed to the feature samples of experimental data utilised for the augmentation. For each sample within the training and validation data structure, a 10-fold random

Table 3 – Standard deviations of attributed to feature used for data augmentation.						
Input features	Standard deviation					
Hardness [HRC]	0.35					
Ultimate tensile strength [MPa]	5					
Toughness [J/m <sup>3</sup> ]	5					
Elongation to fracture [%]	$5 \times 10^{-3}$					
Fracture stress [MPa]	5					
Toughness(H) [J/m³]	5					
elongation to fracture (H) [%]	$5 \times 10^{-5}$					
HSP(Stress) [%]	$5 \times 10^{-5}$					
HSP (elongation) [%]	$1 \times 10^{-4}$					
HSP (toughness) [%]	$1 \times 10^{-4}$					
System energy [J]	0.5					
Voltage [V]	0.1					
Current [A]	$1 \times 10^{-2}$					
Charging time [s]	60					
Concentration of NH <sub>4</sub> SCN [g/l]	0.1					
Measured hydrogen concentration [wt.ppm]	$5 \times 10^{-2}$					

variates were generated respectively with a controlled random *seed*. The resulting input data was 649 samples. No augmentation was carried out on the test dataset. In fact, the test data was not at any point exposed to the ANN during training and validation. The test data was strictly used to test the predictive error of the network.

In the preparation of ANN input data, normalisation is typically carried out to ensure the unification of input data, thus ensuring that all input features have equal importance on the prediction [38,47]. In this study, normalisation of the input data was performed using *preprocessing.normalize* from scikit-*learn* library. The normalisation model was saved and called at any time to ensure the reproducibility of the normalisation process.

## Neural network architecture

The multilayer feed-forward artificial neural network (ANN) models, with backpropagation, were developed using *Keras*, an open-source library built on *TensorFlow* software (version 2.3.0) for machine learning applications with *python* programming language [48]. All codes in this study were written and executed using the *Jupiter* platform.

The ANN is intended to receive as input data containing the features presented in Table 2 and predict a target output of total hydrogen concentration responsible of mechanical properties degradation and failure for the given data sample. Two ANN models of identical architecture were developed: Model I and Model II. The topology of the models however differs in terms of input feature sizes. In Model I, the charging parameters i.e. applied potentials ( $E_a$ ), hydrogen charging times ( $T_{ch}$ ) and initial current ( $I_i$ ) were transformed into system energy ( $E_s$ ) using Equation 2

$$E_{\rm S} = E_a \times I_i \times T_{ch} \tag{2}$$

Where:  $E_a$  is the absolute value of the applied electrochemical potential in volts,  $I_i$  the initial corresponding current in ampere and  $T_{ch}$  is the charging time in seconds. The standard deviation selected for data augmentation of energy system feature was 0.5 J, that can be introduced by the following parameters amplitude:  $E_a$ ,  $T_{ch}$ ,  $I_i$ , which are physically applicable, and at the same time significant in representing a wide range of laboratorial testing, or real service conditions. The ANN models were developed with an input layer with 12 nodes for Model I and 14 nodes for Model II corresponding to



Fig. 5 - Characteristic function of a unit node within an ANN hidden layer.

their respective input features number. The input layer is then followed by 4 hidden densely connected layers with respective nodes and summed up with a single node output layer. The topology of the networks can be summarised as 12-7-5-3-2-1 and 14-7-5-3-2-1 for Model I and Model II, respectively. The networks operate sequentially, where every node in a layer receives an input from nodes in preceding layer and transmits an output to nodes in subsequent layer within the framework of supervised leaning. Fig. 5 illustrates the characteristic function of a unit node within a layer.

Hyperparameters, especially hidden layers and nodes play the most important role in the perfomance of an ANN [49]. Despite the proposals in literature with equations governing the selection of hyperparameters [40,41,50], there is not thumb rule or robust methodology for the selection of hidden layers and nodes. Some proposed methods are suitable only for specific applications [38]. In this study, the network topology 12-7-5-3-2 was selected by trial and error while monitoring the perfomance of the models. The oupout from all layers in the model are governed by the rectified linear units (ReLU) activation function as it ensures better gradient propagation [51]. In order to efficiently control the adjustment of added weight and maximise the loss function which helps in monitoring the model performance, the adaptive estimation of first and second-order moments (adam) algorithm was used as an optimizer. This is based on a stochastic gradient descent method. Several studies testify of the computational efficiency and little memory requirement of adam, as well as its suitability for problems that are large in terms of input features [52,53]. In this research adam was used with its default arguments including default values for learning rate and step size. The learning perfomance evaluation of the developed ANNs is monitered by the mean square error (MSE) loss function [54]. MSE loss score during training of the model is determined by Equation (3).

$$MSE = \frac{1}{N} \sum_{i=0}^{N} (y - Y)^{2}$$
(3)

Where: N is the number of samples of input data, y is the experimental output data and Y is the predicted output by the ANN.

# **Results and discussion**

Hydrogen charging and mechanical testing were performed on four steels to highlight the deleterious effect of hydrogen on their mechanical perfomance. In addition, curate and process the obtained data for the development of ANN models that can predict the hydrogen concentration responsible for mechanical properties degradation and eventual failure. The influence of hydrogen on fracture strength, one of the input features of the test is presented in Fig. 6 (a), (b), (c) and (d) for DQ1, DQ2, A860 and A960, respectively.

For DQ1, the fracture stress of the steel reduces from 2200 MPa for small hydrogen concentrations of about 1 wt.ppm, to a clear lower plateau at about 1100 MPa for higher concentrations up to 4.8 wt.ppm. The observed lower plateau can be considered as the maximum hydrogen effect on the UTS of the studied steel under the employed experimental conditions. Similar phenomenon of the embrittlement curve has been reported for several steels grades with varying strengths, hardness, and microstructures [31,33]. Contrary to DQ1 which manifests sudden strength degradation, DQ2 rather shows a resistance to hydrogen presence in terms of strength degradation at hydrogen concentrations up to 1.5 wt.ppm. Further analysis of the mechanical data revealed nevertheless a reduction in ductility in this regime that did not manifest strength degradation. At hydrogen concentrations of 2 wt.ppm under load, DQ2 becomes highly sensitive to hydrogen, manifesting significant strength degradation up to 1150 MPa at concentrations of 2.7 wt.ppm. Further attempts to induce more hydrogen into the material after 2.7 wt.ppm by increasing the electrolyte poison and potential was not successful, DQ2 however continued to exhibit further strength degradation without increase in hydrogen concentration (right grey area in Fig. 6 (b)). This may have resulted from aggressive hydrogen charging conditions with high fugacity that may have caused substantial defects on the specimen surface. Nevertheless, further investigation is required to clarify the magnitude of damage to the surface caused by hydrogen at high-fugacity and its contribution to early failure despite the relatively low hydrogen concentration.

A860 and A960 steel conditions do not manifest localized strength degradation below the yield strength as DQ1 and DQ2. Fracture stress as a function of hydrogen concentration for A860 and A960, tends to follow linear trends up to stresses of about 1300 MPa at hydrogen concentrations of 5 wt.ppm and 3 wt.ppm for A860 and A960, respectively. It is noticeable that A960 manifests significant scatter as shown Fig. 6 (b). The same scatter was observed for hydrogen effect on ductility for A960 leading to the inference that the manifested scatter maybe associated with an inherent property of the steel. The scattering phenomenon observed for A960 has also been observed in other works for crack initiation time and time to fracture for steels with similar microstructure [34]. The disparate responses of the distinctive test steels to hydrogen can be mainly attributed to the difference in their



Fig. 6 - Fracture stress as a function of hydrogen concentration for (a) DQ1, (b) DQ2, (c) A860, and (d) A960.

microstructure especially the average size and shape of PAG as shown in Fig. 2. In addition, it is important to mention that microstructural defects resulting from the deformation during tensile testing may influence the hydrogen concentration measurements as the different steels deform and retain hydrogen differently. In future studies, this can be evaluated by analysing the obtained spectra from TDS hydrogen concentration measurements.

It is worth noting that the deleterious effect of hydrogen on the test steels shown in Fig. 6 represents effect on one feature, i.e. mechanical strength, meanwhile the effect of hydrogen on the steels in terms mechanical properties is multi-dimensional (including effect on ductility, steel surface quality, size, etc.). Nonetheless the results shown in Fig. 6 are clearly representative of the effect of hydrogen on the respective steels. What is important to underline from these results is that the hydrogen influence on the mechanical integrity of the respective steels is dissimilar. Despite that the effect of hydrogen on the steels are commonly deleterious, their nature and magnitudes are different. The input data consisting of different steels is important for the ANN models in terms of reliability and their suitability for a wider scope of applications.

A preliminary statistical correlation analyses was performed on the input data features. A correlation matrix (heatmap of Pearson's corelation) was used to investigate the

degree of correlation between the input feature pairs and target output. The utilised topology in terms of hidden layers and neurons was found by trial and error to be superlative for Model I. Attempts to add or subtract any hidden layers results in the model overfitting or underfitting quickly. It is worth to note that the current features have been retained after a preliminary feature selection procedure [55]. Fig. 7 (a) and (b) depicts the heatmap correlation between the features of the ANN input data for Model I and Model II, respectively. As shown in map of Fig. 7 (a) for Model I, less than 20% of the feature pairs have moderate to strong positive or negative correlation (|R| > 0.65). The strongest correlation is mostly between original M<sub>p</sub> and HSP which indubitably consist of original M<sub>p</sub> as A component per Equation (1). The same map further reveals a low correlation (|R| < 0.5) between more that 75% of the feature pairs, of which more than 95% are statistically insignificant at P.01 (highlighted by asterisks). For Model II, despites the increase in feature pairs, the map shows only about 25% of the feature pairs have relatively strong positive or negative correlation (|R| > 0.65) which is also attributed to inherent correlation between M<sub>p</sub> and HSP and charging time. Low to moderate Correlations with |R| < 0.5, exist between more that 65% of the feature pairs, of which more than 80% are statistically insignificant at P.01 (highlighted by asterisks). In other words, for both models, majority of the input data feature pairs



Fig. 7 – Heatmap of Pearson's linear correlation between the feature of ANN input data (a) Model I (b) Model II [where asterisks represent correlation that are insignificant at P.01].

have a weak correlation. Although other non-linear, multivariate forms of correlation may exist, classical linear regression solutions do not suffice to predict the hydrogen concentration with the available data. Nevertheless, more powerful methods like ANN can establish any existing nonlinear or multi-dimensional relationships between the input features and target output. This is of great value, when there is no analytical validated formulation for this multi-physical problem.

The learning performance of the ANN models, i.e., training history, prediction accuracy at training, validation and testing are evaluated using the mean absolute error (MAE) metric [56]. MAE is the measure of errors between pairs of true and predicted values over the training epochs, where each epoch is a full dataset training iteration. MAE is calculated with Equation (4).

$$MAE = \frac{\sum_{i=1}^{n} abs(T_i - P_i)}{n}$$
(4)

Where:  $T_i$  are the experimental true values,  $P_i$  are the predicted values, and *n* the number of epochs.

Fig. 8 (a) shows the final training history of Model I (12-7-5-3-2-1). The utilised topology in terms of hidden layers and neurons was found by trial and error to be superlative for Model I. Attempts to add or subtract any hidden layers results in the model overfitting or underfitting quickly and getting stuck in a local minimum, evidenced by straight line in MAE for both training and validation history [57]. Considering the training history with this topology, MAE for training consistently reduces up to 10<sup>5</sup> epochs. MAE for validation data originally declined but starts to overfits at about 10<sup>4</sup> epochs followed with a momentary spike. MAE nevertheless continues to decline until  $6 \times 10^4$  when the overfitting gap between training and validation starts to widen remarkably. Consequently, Model I was re-run and training was stopped at  $6 \times 10^4$  epochs and saved. The model was then called to make predictions on training, validation, and test datasets.



Fig. 8 – The learning performance of the ANN Model I, in terms of (a) training history over epochs for and (b) prediction of hydrogen concentration with training data set.



Fig. 9 — Correlation between experimental and predicted hydrogen concentration from ANN Model I, using (a) validation dataset, (b) test dataset.

Model I makes good predictions of hydrogen concentration on training dataset with a MAE of 0.04 wt.ppm and a linear correlation of 99% (R = 0.99) as shown in Fig. 8(b). With the validation dataset, Model I make predictions with a MAE of 0.18 wt.ppm and R = 0.84 as shown in Fig. 9 (a). It is worth noting that the training data was fully exposed to model during training and high accuracy predictions using the same data does not translate to a good predictive performance of the model. The same applies to predictions with validation datasets because of possible data leakage events that may occur during model training. Normalisation of the entire data to ensure data unification before splitting in to various sets and cross-validation procedure during training are mostly responsible for leakage [58,59]. The predictive performance of any ANN at a 'never seen before' test dataset is considered the most reliable indication of the model's performance [60]. With the test data, Model I predicts hydrogen concentration with good accuracy as shown in Fig. 9 (b), (MAE = 0.28 wt.ppm and R = 0.8). The results reflect a good correlation between the predicted and experimentally measured hydrogen concentration responsible for hydrogen assisted mechanical failure. It is worth stating that the model shows a good predictive performance that is not sensitive to respective steels with disparate morphology of PAG as the test dataset was carefully

selected to include all fours DQ test steels. However, further data and investigation is required ascertain the predictive performance of the model for the same steels under different mechanical test conditions (e.g. varying strain rates).

To evaluate the effect of increased feature size on the training and prediction performance of the developed ANN, Model II, was developed keeping the individual hydrogen charging parameters as input features resulting in a 14-7-5-3-2-1 network topology. This was performed to test the general assumption that an increase feature size may cause quicker overfitting or allow a good fit for training data but result in a bad prediction performance on test data due to lack of generalization [61,62]. Fig. 10 (a) shows the training history of Model II up to 10<sup>5</sup> epochs. Contrary to Model I, the Model II presents several spikes in MAE increase for training and validation are noticeable. These could be attributed to disturbances during the training caused by the presence of additional features. The trade-off between decreasing MAE for training data and increasing MAE for validation data set was made at  $6.5 \times 10^4$  epochs when the misfit in MAE between the training and validation datasets manifest significant increase. Model II was called to make hydrogen concentration predictions with train, validation, and test datasets as inputs. With the training dataset, Model II makes hydrogen



Fig. 10 — The learning performance of the ANN Model II, in terms of (a) training history over epochs, and (b) prediction of hydrogen concentration with training data set.



Fig. 11 – Correlation between experimental and predicted hydrogen concentration from ANN Model II, using (a) validation dataset, (b) test dataset.

concentrations predictions with MAE of 0.15 wt.ppm, and 96% corelation with experimental data (see Fig. 10 (b) compared to 0.02 wt Ppm and 99% for model I.

With validation and test datasets, as depicted in Fig. 11 (a) and (b), Model II makes predictions with MAE of 0.37 wt.ppm and 0.3 wt.ppm, respectively. It is important to remember that the information contained by a single feature Es (Equation (2)) in Model I, it is decomposed in the respective three physical features in Model II. In other words, both models contain the same information on hydrogen charging parameters and conditions. Despite that, it is evident that the increase in feature size has some effect on the learning and predictive performance of Model II, resulting in predictions with relatively higher errors irrespective of longer training time (5  $\times$  10<sup>3</sup> epochs more). In future works, a thorough feature extraction, selection and principal component analysis is strongly recommended to ascertain the specific influence of respective features on the predictive performance of the developed models. Nevertheless, this study demonstrates that steels' mechanical properties in air and after hydrogen induced failure in combination with hydrogen charging parameters (or hydrogen-rich service conditions) as input data, possess the requisite information to predict the hydrogen concentrations responsible for the respective damage via ANN techniques.

# Conclusion

Two artificial neural network (ANN) models were developed to predict the hydrogen concentration responsible for mechanical properties degradation DQ and DQ + RQ martensitic ultra-high-strength steels. The inherent mechanical properties of the steels in as supplied state, after hydrogen-induced failure and hydrogen charging parameters for four steels encompassing different chemical composition and thermal treatment, resulting in similar martensitic microstructure but different PAG size and morphology, were used as inputs to represent this family of steels and wide range of hydrogen rich service conditions. The following conclusions can be drawn.

- The developed ANN Model I (12-7-5-3-2-1) and Model II (14-7-5-3-2-1) show comparable hydrogen concentration prediction abilities with remarkable accuracy.
- The ANN models show good hydrogen concentration prediction for all considered steels irrespective of their different mechanical responses to varying levels of hydrogen concentration resulting from differences in prior austenitic grain (PAG) morphology.
- Increase in feature size in Model II, by considering the original three parameters controlling the hydrogen charging energy, requires longer training time to reach a good mean absolute error (MAE) trade-off for training and validation dataset, yet reduces prediction performance evaluated with test dataset compared to Model I.
- The developed models may not be suitable for hydrogen concentration at fracture for other family of DQ highstrength steels, namely with significant volume fraction of retained austinite, and steels with dominantly ferritic or austenitic microstructures as is expectable that they respond differently to hydrogen-induced fracture.
- In future works, the effect of PAG size and boundary surface area as input features on the prediction performance of the developed models can be considered for further studies.

# Funding

This researcher was supported by the Public Research Networked with Companies (Co-Innovation) program of Business Finland via the projects 7743/31/2018 (ISA Aalto-HydroSafeSteels) and 7537/31/2018 (ISA-Intelligent Steel Applications).

# **Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

The authors wish to acknowledge Dr. Saara Mehtonen, Dr. Esa Virolainen, Kim Widell and Alison Tshala for their committed support during this study.

#### REFERENCES

- Christmann K. Some general aspects of hydrogen chemisorption on metal surfaces. Prog Surf Sci 1995;48(1-4):15-26.
- [2] Panella B, Hirscher M. Hydrogen physisorption in metal-organic porous crystals. Adv Mater 2005;17(5):538–41.
- [3] Rudomilova D, et al. The effect of microstructure on hydrogen permeability of high strength steels. Mater Corros 2020;71(6):909-17.
- [4] Mallick D, et al. Study of diffusible behavior of hydrogen in first generation advanced high strength steels. Metals 2021;11(5):782.
- [5] Li H, et al. Hydrogen adsorption with micro-structure deformation in nanoporous carbon under ultra-high pressure. Int J Hydrogen Energy 2022;47(81):34555–69. https://doi.org/10.1016/j.ijhydene.2022.08.030.
- [6] Liu MA, et al. Microstructural influence on hydrogen permeation and trapping in steels. Mater Des 2019;167:107605.
- [7] Chen Y-S, et al. Observation of hydrogen trapping at dislocations, grain boundaries, and precipitates. Science 2020;367(6474):171–5.
- [8] Venezuela J, et al. The influence of microstructure on the hydrogen embrittlement susceptibility of martensitic advanced high strength steels. Mater Today Commun 2018;17:1–14.
- [9] Zhang Z, et al. Combined effects of stress and temperature on hydrogen diffusion in non-hydride forming alloys applied in gas turbines. Int J Hydrogen Energy 2022;47(71):30687–706.
- [10] Colla V, Valentini R. Assessment of critical hydrogen concentration in as-cast and hot-rolled billets in medium carbon steels. Steel Res Int 2020;91(9):2000126.
- [11] Fangnon E, et al. Determination of critical hydrogen concentration and its effect on mechanical performance of 2200 mpa and 600 hbw martensitic ultra-high-strength steel. Metals 2021;11(6):984.
- [12] Nowak C, et al. Atomistic simulations of hydrogen distribution in Fe–C steels. Int J Hydrogen Energy 2022;47(76):32732–40.
- [13] Oriani R, Josephic P. Equilibrium aspects of hydrogeninduced cracking of steels. Acta Metall 1974;22(9):1065–74.
- [14] Troiano AR. The role of hydrogen and other interstitials in the mechanical behavior of metals. trans. ASM 1960;52:54–80.
- [15] Sofronis P, Birnbaum HK. Mechanics of the hydrogendashdislocationdashimpurity interactions—I. Increasing shear modulus. J Mech Phys Solid 1995;43(1):49–90.
- [16] Clum JA. The role of hydrogen in dislocation generation in iron alloys. Scripta Metall 1975;9(1):51–8.
- [17] Nagumo M, Nakamura M, Takai K. Hydrogen thermal desorption relevant to delayed-fracture susceptibility of high-strength steels. Metall Mater Trans 2001;32(2):339–47.
- [18] Shih D, Robertson I, Birnbaum H. Hydrogen embrittlement of α titanium: in situ TEM studies. Acta Metall 1988;36(1):111–24.

- [19] Djukic MB, et al. The synergistic action and interplay of hydrogen embrittlement mechanisms in steels and iron: localized plasticity and decohesion. Eng Fract Mech 2019;216:106528.
- [20] Du Y, et al. Hydrogen diffusivity in different microstructural components in martensite matrix with retained austenite. Int J Hydrogen Energy 2021;46(11):8269–84.
- [21] Shibata A, et al. Hydrogen-related fracture behavior under constant loading tensile test in as-quenched low-carbon martensitic steel. Metals 2022;12(3):440.
- [22] Du Y, et al. Hydrogen embrittlement behavior of high strength low carbon medium manganese steel under different heat treatments. Int J Hydrogen Energy 2019;44(60):32292–306.
- [23] Robertson IM, et al. Hydrogen embrittlement understood. Metall Mater Trans 2015;46(6):2323-41.
- [24] Brahimi SV, Yue S, Sriraman KR. Alloy and composition dependence of hydrogen embrittlement susceptibility in high-strength steel fasteners. Phil Trans Math Phys Eng Sci 2017;375(2098):20160407.
- [25] Malitckii E, Yagodzinskyy Y, Vilaça P. Role of retained austenite in hydrogen trapping and hydrogen-assisted fatigue fracture of high-strength steels. Mater Sci Eng 2019;760:68–75.
- [26] Krauss G. Martensitic structure. In: Encyclopedia of iron, steel, and their alloys. CRC Press; 2016. p. 2182–7.
- [27] Zhu X, et al. Effect of retained austenite stability and morphology on the hydrogen embrittlement susceptibility in quenching and partitioning treated steels. Mater Sci Eng 2016;658:400–8.
- [28] Yang J, et al. Effect of retained austenite on the hydrogen embrittlement of a medium carbon quenching and partitioning steel with refined microstructure. Mater Sci Eng 2016;665:76–85.
- [29] Cho L, et al. Effects of hydrogen pressure and prior austenite grain size on the hydrogen embrittlement characteristics of a press-hardened martensitic steel. Int J Hydrogen Energy 2021;46(47):24425–39.
- [30] Tehranchi A, Curtin WA. The role of atomistic simulations in probing hydrogen effects on plasticity and embrittlement in metals. Eng Fract Mech 2019;216:106502.
- [31] Lovicu G, et al. Hydrogen embrittlement of automotive advanced high-strength steels. Metall Mater Trans 2012;43(11):4075–87.
- [32] Malitckii E, et al. Hydrogen effects on mechanical properties of 18% Cr ferritic stainless steel. Mater Sci Eng 2017;700:331–7.
- [33] Valentini R, et al. Investigation of mechanical tests for hydrogen embrittlement in automotive PHS steels. Metals 2019;9(9):934.
- [34] Latypova R, et al. Hydrogen-induced stress corrosion cracking studied by the novel tuning-fork test method. Mater Corros 2020;71(10):1629–36.
- [35] Claeys L, et al. Electrochemical hydrogen charging of duplex stainless steel. Corrosion 2019;75(8):880–7.
- [36] Malitckii E, Yagodzinskyy Y, Hänninen H. Hydrogen uptake from plasma and its effect on EUROFER 97 and ODS-EUROFER steels at elevated temperatures. Fusion Eng Des 2015;98:2025–9.
- [37] Yamabe J, Awane T, Matsuoka S. Investigation of hydrogen transport behavior of various low-alloy steels with highpressure hydrogen gas. Int J Hydrogen Energy 2015;40(34):11075–86.
- [38] Thankachan T, et al. Artificial neural network to predict the degraded mechanical properties of metallic materials due to the presence of hydrogen. Int J Hydrogen Energy 2017;42(47):28612–21.

- [39] Azimzadegan T, et al. An artificial neural-network model for impact properties in X70 pipeline steels. Neural Comput Appl 2013;23(5):1473–80.
- [40] Haque ME, Sudhakar K. Prediction of corrosion-fatigue behavior of DP steel through artificial neural network. Int J Fatig 2001;23(1):1-4.
- [41] Haque ME, Sudhakar K. ANN back-propagation prediction model for fracture toughness in microalloy steel. Int J Fatig 2002;24(9):1003–10.
- [42] Lu Z, et al. Predicting hydrogen storage capacity of V-Ti-Cr-Fe alloy via ensemble machine learning. Int J Hydrogen Energy 2022;47(81):34583-93. https://doi.org/ 10.1016/j.ijhydene.2022.08.050.
- [43] Plumb AP, et al. Optimisation of the predictive ability of artificial neural network (ANN) models: a comparison of three ANN programs and four classes of training algorithm. Eur J Pharmaceut Sci 2005;25(4–5):395–405.
- [44] Malitckii E, Fangnon E, Vilaça P. Study of correlation between the steels susceptibility to hydrogen embrittlement and hydrogen thermal desorption spectroscopy using artificial neural network. Neural Comput Appl 2020;32(18):14995–5006.
- [45] Huynh T, et al. General methods for quantitative interpretation of results of digital variable-volume assays. Analyst 2019;144(24):7209–19.
- [46] Wedel W, et al. Improving information gain from optimization problems using artificial neural networks. 2004.
- [47] Chollet F. Deep learning with Python. Simon and Schuster; 2021.
- [48] Moolayil J. An introduction to deep learning and keras. In: Learn Keras for deep neural networks. Springer; 2019.
   p. 1–16.
- [49] Uzair M, Jamil N. Effects of hidden layers on the efficiency of neural networks. In: 2020 IEEE 23rd international multitopic conference (INMIC). IEEE; 2020.
- [50] Xu L, et al. Artificial neural network prediction of retained austenite content and impact toughness of high-vanadium high-speed steel (HVHSS). Mater Sci Eng 2006;433(1–2):251–6.
- [51] Yang J. ReLU and softmax activation functions. Hämtat från6; 2017. https://github. com/Kulbear/deep-learningnano-foundation/wiki/ReLU-and-Softmax-Activation-Functions den.

- [52] Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980. 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings. 2015. Digital Bibliography & Library Project; 2015. https://arxiv.org/pdf/1412.6980.pdf?source=post\_page.
- [53] Liu Y, et al. Learning by turning: neural architecture aware optimisation. In: International conference on machine learning. PMLR; 2021.
- [54] Tieleman T, Hinton G. Lecture 6.5-rmsprop: divide the gradient by a running average of its recent magnitude. COURSERA: Neural networks for machine learning. 2012. p. 26–31.
- [55] Kaur A, Guleria K, Trivedi NK. Feature selection in machine learning: methods and comparison. In: 2021 international conference on advance computing and innovative technologies in engineering (ICACITE). IEEE; 2021.
- [56] Vastrad C. Performance analysis of neural network models for oxazolines and oxazoles derivatives descriptor dataset. arXiv preprint arXiv:1312.2853. 6th3. published International Journal of Information Sciences and Techniques (IJIST); 2013.. https://doi.org/10.5121/ijist.2013.3601.
- [57] De Villiers J, Barnard E. Backpropagation neural nets with one and two hidden layers. IEEE Trans Neural Netw 1993;4(1):136–41.
- [58] Elangovan A, He J, Verspoor K. Memorization vs. generalization: quantifying data leakage in NLP performance evaluation. arXiv preprint arXiv:2102.01818. Conference of the European Chapter of the Association for Computational Linguistics; 2021. https://doi.org/10.48550/arXiv.2102.01818.
- [59] Christensen R. Thoughts on prediction and cross-validation. Department of Mathematics and Statistics University of New Mexico. University of new mexico; 2015. https://stat.unm. edu/~fletcher/Prediction.pdf.
- [60] Wang H, et al. Machine learning basics. Deep learning. 2016. p. 98–164.
- [61] Ying X. An overview of overfitting and its solutions. In: Journal of physics: conference series. IOP Publishing; 2019.
- [62] Kernbach JM, Staartjes VE. Foundations of machine learningbased clinical prediction modeling: Part II—generalization and overfitting. Machine Learning in Clinical Neuroscience. 2022. p. 15–21.