



Application of Machine Learning to Regression Analysis of a Large SMA Weld Metal Database

A cluster analysis of a coalesced Fe-C-Mn high-strength steel dataset revealed that the ultimate tensile strength of weld metal can be related to austenite-to-ferrite transformation temperature in at least four ways

BY R. VARADARAJAN AND K. SAMPATH

Abstract

A machine learning approach was used to perform a regression analysis of Evans's shielded metal arc (SMA) weld metal (WM) database involving several groups of Fe-C-Mn high-strength steels. The objective of this investigation was to develop an expression for austenite-to-ferrite (A_{r3}) transformation temperature that also included the effects of principal and minor alloy elements (in wt-%) and weld cooling rate (in °C/s) and relate this expression with WM ultimate tensile strength (UTS). The A_{r3} data from 257 records obtained from several selected sources were combined with A_{r3} projections at extreme end points in Evans's WM database.

Subsequently, a cluster analysis was performed. The data in Evans's database was filtered with the carbon equivalent number limited to 0.3 maximum, carbon content limited to 0.1 wt-% maximum, nitrogen content limited to 99 ppm (0.0099 wt-%) maximum, preassigned A_{r3} values limited to 680°C minimum, and WM UTS limited to 710 MPa maximum. The results provided a good approximation to the expression for A_{r3} transformation temperature in terms of elemental compositions and cooling rate. This allowed the A_{r3} to correlate with WM UTS of Fe-C-Mn in at least four ways depending on the sign of correlation of the data clusters.

The elemental combinations in the cluster with the highest negative correlation revealed highly predictable WM UTS. In particular, the new A_{r3}

expression helped to predict decreases observed in certain A_{r3} experimental data on WMs with balanced Ti, B, Al, N, and O additions reported among 13 records with additional dilatometry results.

This correlation between the new expression for the A_{r3} temperature and UTS of Fe-C-Mn WM is expected to complement the Japan Welding Engineering Society artificial neural network model currently available to predict Charpy V-notch test temperature for 28 J absorbed energy based on WM chemical composition. It will thereby provide a pair of effective tools for efficient development and/or evaluation of high-performance welding electrodes based on an Fe-C-Mn system for demand-critical applications.

Keywords

- Fe-C-Mn High-Strength Steels
- Shielded Metal Arc Welding (SMAW)
- Database
- Machine Learning
- Cluster Analysis
- Transformation Temperature

Introduction

In May 2017, Dr. Glyn M. Evans (formerly with Oerlikon, Switzerland) posted a large shielded metal arc (SMA) weld metal (WM) database on ResearchGate (Ref. 1). This database contains more than 900 WM compositions based on an Fe-C-Mn system. These WM compositions belong to 74 types or groups of Fe-C-Mn alloy systems and were derived from

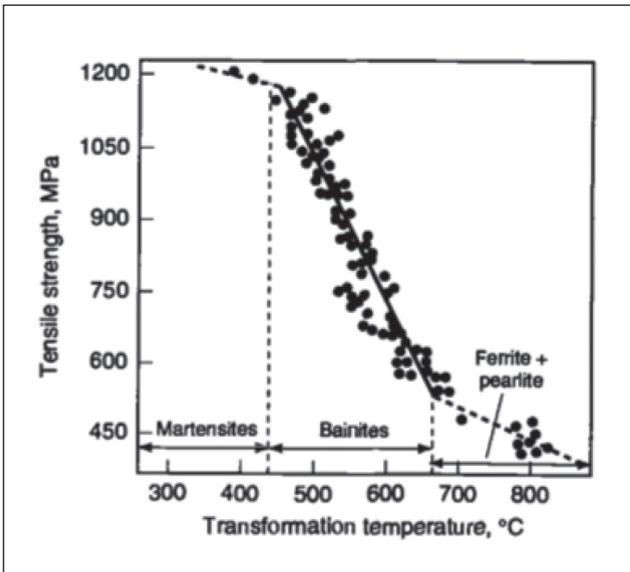


Fig. 1 — Relationship between UTS (in MPa) and transformation temperature (in °C) of ferrite-pearlite, bainitic, and martensitic steels (Refs. 5, 6).

the book *Metallurgy of Basic Weld Metal* by G. M. Evans and N. Bailey (Ref. 2). Each WM composition includes individual ranges of 16 (C, Si, Mn, P, S, Cu, Ni, Cr, Mo, Nb (Cb), V, Ti, B, Al, N, and O) alloy element additions to Fe, along with the respective values for six WM properties that include yield strength (YS), ultimate tensile strength (UTS), reduction of area (%RA), elongation (%EL), and the test temperature required to achieve 100 J ($T_{100J}/^{\circ}\text{C}$) and 28 J ($T_{28J}/^{\circ}\text{C}$) absorbed energy during Charpy V-notch (CVN) impact testing.

A recent research effort (Ref. 3) utilized a constraints-based model (CBM) as a simple and effective framework for organizing and analyzing the Fe-C-Mn SMA WM database (see aws.org/2021.100.036-database) to gain valuable insights. The CBM was built on the metallurgical principle that lowering relevant solid-state phase transformation (i.e., austenite decomposition) temperatures is beneficial in improving WM strength and fracture toughness while simultaneously reducing carbon content. The carbon equivalent number (CEN) developed by Yurioka et al. (Ref. 4) was also beneficial in improving the weldability of high-strength steels.

Figure 1 (Refs. 5, 6) illustrates the relationship between the transformation temperature and UTS of ferritic-pearlitic, bainitic, and martensitic steels. The UTS of the ferritic-pearlitic steels appeared to range between 400 and 550 MPa, while the corresponding transformation temperature appeared to range between 650° and 900°C. Interestingly, these ranges allow one to limit data selection when performing regression analysis of a large database such as Evans's SMA WM database.

To gain valuable insights into Evans's database, the CBM used various statistical (regression) equations and obtained several calculated metallurgical characteristics (CMCs). The CMCs related the chemical composition of high-strength steel WM through Yurioka et al.'s CEN and selected solid-state phase transformation-start (T_s) temperatures, such as Ouchi et al.'s ferrite-start (A_{r3}) temperature (Ref. 7) and Steven and

Haynes's bainite-start (B_s) and martensite-start (M_s) temperatures (Ref. 8), through respective constitutive equations. The individual CMCs allowed classification and/or ranking of the WMs in the database.

The CEN regression equation developed by Yurioka et al. (Ref. 4) is commonly used to evaluate the hydrogen cracking sensitivity of various types of modern structural, pipeline, and pressure vessel steels:

$$CEN = C + \{A(C) \times EMU\} \quad (1)$$

where $A(C)$ refers to the accommodation factor that is a function of C content, while EMU refers to a set of elemental multiplication units involving Si, Mn, Cu, Ni, Cr, Mo, V, Nb (Cb), and B.

$$A(C) = 0.75 + 0.25 \tanh [20 \times (C - 0.12)] \quad (2)$$

$$EMU = \{Si/24 + Mn/6 + Cu/15 + Ni/20 + (Cr + Mo + V + Nb)/5 + 5 \times B\} \quad (3)$$

Yurioka et al.'s aforementioned equation includes microalloy additions such as V, Nb, and B in addition to various principal alloy elements such as C, Mn, Cr, Ni, Mo, and Cu.

Equations 4–6 indicate that all principal alloy elements decrease austenite decomposition temperatures with C affecting to a maximum extent, particularly when C content exceeds 0.12 wt-%. These equations were developed several decades ago and can be used to calculate or estimate the transformation temperatures when the cooling rate supposedly remains constant.

$$A_{r3} (^{\circ}\text{C}) = 910 - 310(C) - 80(Mn) - 20(Cu) - 55(Ni) - 15(Cr) - 80(Mo) \quad (4)$$

$$B_s (^{\circ}\text{C}) = 830 - 270(C) - 90(Mn) - 37(Ni) - 70(Cr) - 83(Mo) \quad (5)$$

$$M_s (^{\circ}\text{C}) = 561 - 474(C) - 33(Mn) - 17(Ni) - 17(Cr) - 21(Mo) \quad (6)$$

The classification and/or ranking of all WMs in Evans's database using various CMCs obtained using the CBM approach (Ref. 3) reaffirmed that controlling the C content, CEN value, and calculated solid-state phase transformation temperatures, particularly the difference between the calculated B_s and calculated M_s temperatures, is critical to developing and identifying high-performance, high-strength steel welding electrodes. A dual approach that manipulated the contents of principal alloy elements, such as C, Mn, Cu, Ni, Cr, and Mo, based on Equations 4–6 and added balanced amounts of Ti, B, Al, N, and O appeared to offer the best means to lower

relevant solid-state T_s temperatures to produce WMs with high strength and exceptional fracture toughness.

A part of Evans's large SMA WM database used an independent scheme to build a total of 24 SMA welds, based on a TiBAlN series. These 24 welds included three subsets of eight welds, each at three levels of nitrogen content: normal (below 85 ppm or 0.0085 wt-%), intermediate (120 to 164 ppm or 0.012 to 0.0164 wt-%), and high (217 to 249 ppm or 0.0217 to 0.0249 wt-%). The primary intent of these three subsets of WMs was to identify and correlate the effects of Ti-B-Al-N-O microalloy additions on WM tensile strength, CVN impact toughness, and microstructure development in the fusion zone and reheated WM.

Table 1 shows the chemical composition of 13 of 24 TiBAlN series of SMA WMs. The C content of these 13 WMs varied between 0.066 and 0.078 wt-%, while the Mn content varied between 1.4 and 1.66 wt-%. The Si content varied between 0.25 and 0.63 wt-%. Other principal alloy elements remained constant: Cr content at 0.03 wt-%, Ni content at 0.03 wt-%, Mo content at 0.005 wt-%, and Cu content at 0.03 wt-%. Microalloy additions V and Nb were held constant at 0.0005 wt-%. Compared to the aforementioned principal alloy elements and V and Nb microalloy additions, the Ti content varied between 0.0001 and 0.054 wt-%, B content ranged between 0.0001 and 0.0167 wt-%, Al content varied between 0.0001 and 0.058 wt-%, N content varied between 0.0041 and 0.0249 wt-%, and O content varied between 0.0282 and 0.0475 wt-%.

Following weld mechanical testing that included all-weld metal tensile testing at ambient temperature and CVN impact testing over a wider temperature range from -120° to $+40^\circ\text{C}$, these 13 WMs were subjected to dilatometric evaluation (Refs. 9–11). These 13 WMs were selected to allow dilatometric evaluation of specific alloy additions relative to a range of Ti, B, Al, N, and O contents, particularly the effect of N content at three levels on both T_s and transformation-finish (T_f) temperatures.

The dilatometric evaluation studied the austenite-to-ferrite transformation during continuous cooling. Test specimens were machined to form hollow cylinders with the following dimensions: 10 mm long by 5 mm O.D. with 1 mm wall thickness. The axis of the test specimen was maintained parallel to the original welding direction. The specimens were subjected to the following controlled thermal cycle: austenitization at 1250°C for 2 min, followed by continuous cooling at a typical (weld) cooling rate of 13°C/s from 800° to 500°C (also known as $\Delta t_{8/5}$; 13°C/s corresponds to about 23 s to cool from 800° to 500°C).

The study determined the T_s , 50% transformation (T_{50}), peak rate transformation (T_{PRTT}), and T_f temperatures of the 13 (O, O2, X, X2, Y, Y2, Z, Z2, U, U2, V, V1, and V2) WMs. Table 2 shows the numerical values for the various transformation temperatures besides ($T_s - T_f$) values, along with UTS and CVN test temperatures for 100 and 28 J absorbed energy of the 13 WMs (Refs. 9–11).

The results found that weld V had the lowest T_s temperature at 680°C and correspondingly the highest UTS at 732 MPa. A progressive increase in T_s temperature occurred for welds U, Y2, Z2, Y, Z, O2, V2, X, X2, V1, O, and U2. Correspondingly, a progressive decrease in UTS from 644 to 528 MPa occurred

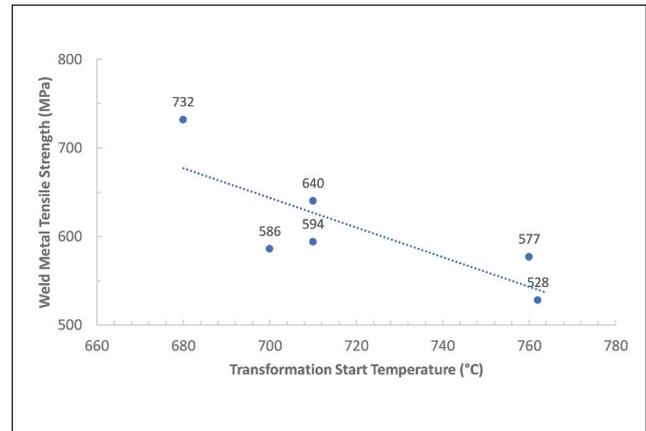


Fig. 2 — Effect of T_s temperature on WM tensile strength.

in welds V1, Z, X2, O2, Y2, Y, V2, U, Z2, X, U2, and O. When WM N content was below 80 ppm (0.008 wt-%), the overall trend between T_s temperature and UTS among all six WMs was found to be highly correlated, as shown in Fig. 2. Two of these six welds, welds Z (with 640 MPa UTS) and Y (with 594 MPa UTS), appeared closer and on either side of the trend line, indicating that their Ti, B, Al, N, and O additions were well or adequately balanced. The trendline equation showed the following:

$$\text{Weld Metal UTS (in MPa)} = 1814 - (1.6722 \times T_s \text{ in } ^\circ\text{C}), \text{ with } R^2 = 0.6368 \quad (7)$$

thereby clearly confirming the metallurgical principle that lowering the transformation temperature aided to increase WM UTS. The trendline also indicated that a T_s temperature greater than 680°C achieved a WM UTS less than 700 MPa (100 ksi).

A recent review (Ref. 12) of the dilatometric results of the 13 Fe-C-Mn high-strength steel SMA WMs (see aws.org/2022.101.010-database) revealed that balanced Ti, B, Al, N, and O additions in welds Z and Y reduced the T_s temperature. For example, weld Z containing microalloy additions as listed in Table 1 showed a total Ti, B, Al, N, and O content at 0.1133 wt-% that appeared to ensure effective deoxidation, formed complex inclusions, and distributed inclusions to enable development of highly fracture-resistant refined WM microstructures. Depending on nominal WM chemical composition and actual effects during welding, these Ti, B, Al, N, O additions further lowered the actual T_s temperatures, thereby promoting a cloudburst of austenite-to-ferrite phase transformation over a narrow ($T_s - T_f$) temperature range. It may be wiser to avoid the rich and lean ends for these microalloy additions, except N, which should be held at the lean end, preferably much below 80 ppm (0.008 wt-%).

As shown in Table 2, at a N content below 0.01 wt-%, the total Ti, B, Al, N, O additions were at 0.0833 wt-% in weld Y and 0.1133 wt-% in weld Z. These two welds offered nearly a 100°C shift in lowering CVN test temperature for either 28 or 100 J absorbed energy. Dilatometric evaluations of reheated WMs showed the following: 1) The balanced total

Table 1 – Chemical Composition of 13 TiBAIN Series Weld Metals (Ref. 1)

Weld ID	C (wt-%)	Si (wt-%)	Mn (wt-%)	P (wt-%)	S (wt-%)	Cu (wt-%)	Ni (wt-%)	Cr (wt-%)
O	0.074	0.25	1.4	0.007	0.008	0.03	0.03	0.03
O2	0.073	0.27	1.66	0.008	0.009	0.03	0.03	0.03
X	0.069	0.45	1.47	0.006	0.005	0.03	0.03	0.03
X2	0.068	0.47	1.46	0.006	0.007	0.03	0.03	0.03
Y	0.07	0.45	1.57	0.01	0.006	0.03	0.03	0.03
Y2	0.069	0.36	1.51	0.008	0.007	0.03	0.03	0.03
Z	0.072	0.49	1.56	0.01	0.007	0.03	0.03	0.03
Z2	0.068	0.5	1.45	0.011	0.006	0.03	0.03	0.03
U	0.073	0.4	1.52	0.011	0.006	0.03	0.03	0.03
U2	0.066	0.36	1.4	0.012	0.007	0.03	0.03	0.03
V	0.078	0.6	1.44	0.007	0.006	0.03	0.03	0.03
V1	0.067	0.63	1.44	0.01	0.005	0.03	0.03	0.03
V2	0.069	0.6	1.42	0.012	0.006	0.03	0.03	0.03
Low	0.066	0.25	1.4	0.006	0.005	0.03	0.03	0.03
High	0.078	0.63	1.66	0.012	0.009	0.03	0.03	0.03
Range	0.012	0.38	0.26	0.006	0.004	0	0	0

Table 1 – (continued)

Mo (wt-%)	Nb (wt-%)	V (wt-%)	Ti (wt-%)	B (wt-%)	Al (wt-%)	N (wt-%)	O (wt-%)
0.005	0.0005	0.0005	0.0001	0.0001	0.0006	0.0079	0.0475
0.005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0235	0.0399
0.005	0.0005	0.0005	0.041	0.0002	0.0001	0.0077	0.0282
0.005	0.0005	0.0005	0.045	0.0002	0.0005	0.0249	0.0297
0.005	0.0005	0.0005	0.039	0.0039	0.0013	0.0083	0.0308
0.005	0.0005	0.0005	0.041	0.0044	0.0005	0.0232	0.0292
0.005	0.0005	0.0005	0.042	0.0048	0.016	0.0067	0.0438
0.005	0.0005	0.0005	0.047	0.0045	0.018	0.023	0.044
0.005	0.0005	0.0005	0.039	0.0158	0.0005	0.0084	0.029
0.005	0.0005	0.0005	0.039	0.0167	0.0005	0.0217	0.0297
0.005	0.0005	0.0005	0.054	0.0056	0.058	0.0041	0.044
0.005	0.0005	0.0005	0.048	0.0044	0.056	0.012	0.0473
0.005	0.0005	0.0005	0.043	0.0035	0.056	0.0235	0.047
0.005	0.0005	0.0005	0.0001	0.0001	0.0001	0.0041	0.0282
0.005	0.0005	0.0005	0.054	0.0167	0.058	0.0249	0.0475
0	0	0	0.0539	0.0166	0.0579	0.0208	0.0193

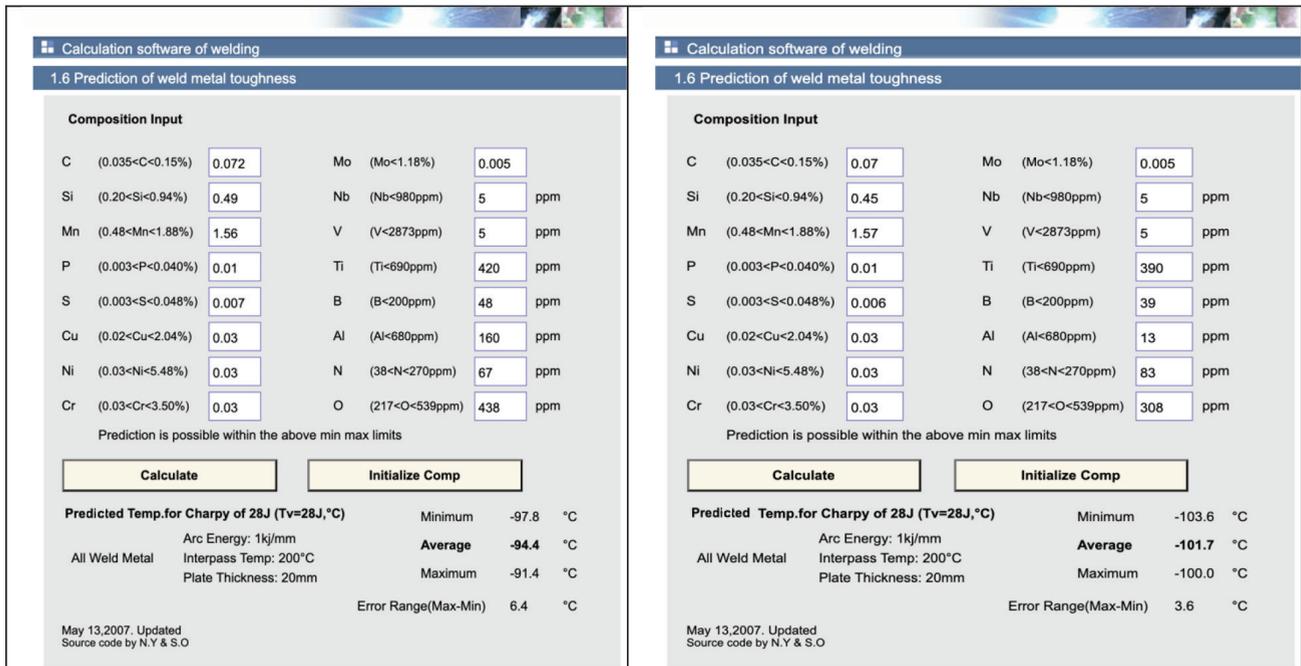


Fig. 3 – The JWES neural network–predicted temperature for 28 J absorbed energy. Weld Z is on the left, and weld Y is on the right.

Ti, B, Al, N, O additions lowered the actual T_s temperature by about 60°C compared to the calculated A_{r3} transformation temperature obtained from Ouchi et al.’s constitutional relationship (Equation 4); 2) N more than 100 ppm (0.010 wt-%) effectively nullified the beneficial effects of Ti, B, and Al additions in lowering the transformation temperature; and 3) at N content much below 80 ppm (0.008 wt-%), both a lower T_s temperature and a narrow start-to-finish ($T_s - T_F$) temperature range helped in achieving exceptional WM CVN impact toughness.

The Japan Welding Engineering Society (JWES) offers a website (www-it.jwes.or.jp/weld_simulator/en/cal6.jsp) wherein one can calculate and predict the temperature for 28 J CVN impact toughness or absorbed energy of Fe-C-Mn WMs based on their chemical composition with certain minimum and maximum limits for all 16 (C, Si, Mn, P, S, Cu, Ni, Cr, Mo, Nb (Cb), V, Ti, B, Al, N, and O) alloy additions to Fe. The prediction is performed using artificial neural network (ANN) analysis by a software developed by D. J. C. MacKay at the University of Cambridge. The prediction is possible within the minimum and maximum limits set for the 16 alloy elements and uses Evans’s SMA WM database on low-alloy, high-strength steel WM as a basis.

Evans’s database (Refs. 1, 2) has been made available to the University of Cambridge (phase-trans.msm.cam.ac.uk/map/data/materials/). The database contains properties of all-weld metals obtained under the constant welding conditions of 1 kJ/mm arc energy and 200°C interpass temperature on 20-mm-thick plates. The ANN prediction gives maximum, minimum, and average values of transition temperature for 28 J CVN impact toughness or absorbed energy along with the degree of prediction error. When the difference between the maximum and minimum predicted values is more than 30°C, the prediction is considered unreliable.

Figure 3 shows the predicted temperature for 28 J absorbed energy for welds Z (left) and Y (right). Table 3 shows a comparison of the actual test values and predicted results of CVN temperature for 28 J absorbed energy for all the 13 original welds. The predicted values for the 13 original welds were quite consistent with actual test results in both values and trend, and the error values associated with predictions were much less than 30°C for each of these welds. Furthermore, when the Ti, B, Al, N, and O contents of all 12 WMs were modified to the same values as in weld Z, the predicted temperature for 28 J absorbed energy for the welds decreased (Table 3) in all cases except weld Y with a total (Ti+B+Al+N+O) content at 0.0833 wt-% and weld V1 with a total (Ti+B+Al+N+O) content at 0.1677 wt-%.

These findings clearly demonstrated that the UTS of Fe-C-Mn ferritic WMs increased with the decreasing T_s , and a superior WM CVN toughness could be achieved by balancing Ti, B, Al, N, and O additions.

As shown in Table 2, the balancing of Ti, B, Al, and O additions may be related to T_s , with a decreasing T_s requiring a lower amount of Ti, B, Al, and O additions. As revealed by weld V, when T_s is at its low end, excessive amounts of Ti, B, Al, and O additions likely raised the CVN test temperature for 28 J absorbed energy, indicating the possibility to form numerous inclusions, which resulted in a so-called dirty weld. By contrast, as revealed by weld O, when T_s is at its high end, disproportionate amounts of Ti, B, Al, and O additions likely raised the CVN test temperature for 28 J absorbed energy, indicating the possibility of free oxygen in solution.

Interestingly, the JWES ANN template allows one to identify balanced Ti, B, Al, N, and O contents of WMs by manipulating their contents within the specified ranges mentioned in the JWES ANN template and achieve a CVN test temperature for 28 J absorbed energy colder than -60°C while ensuring

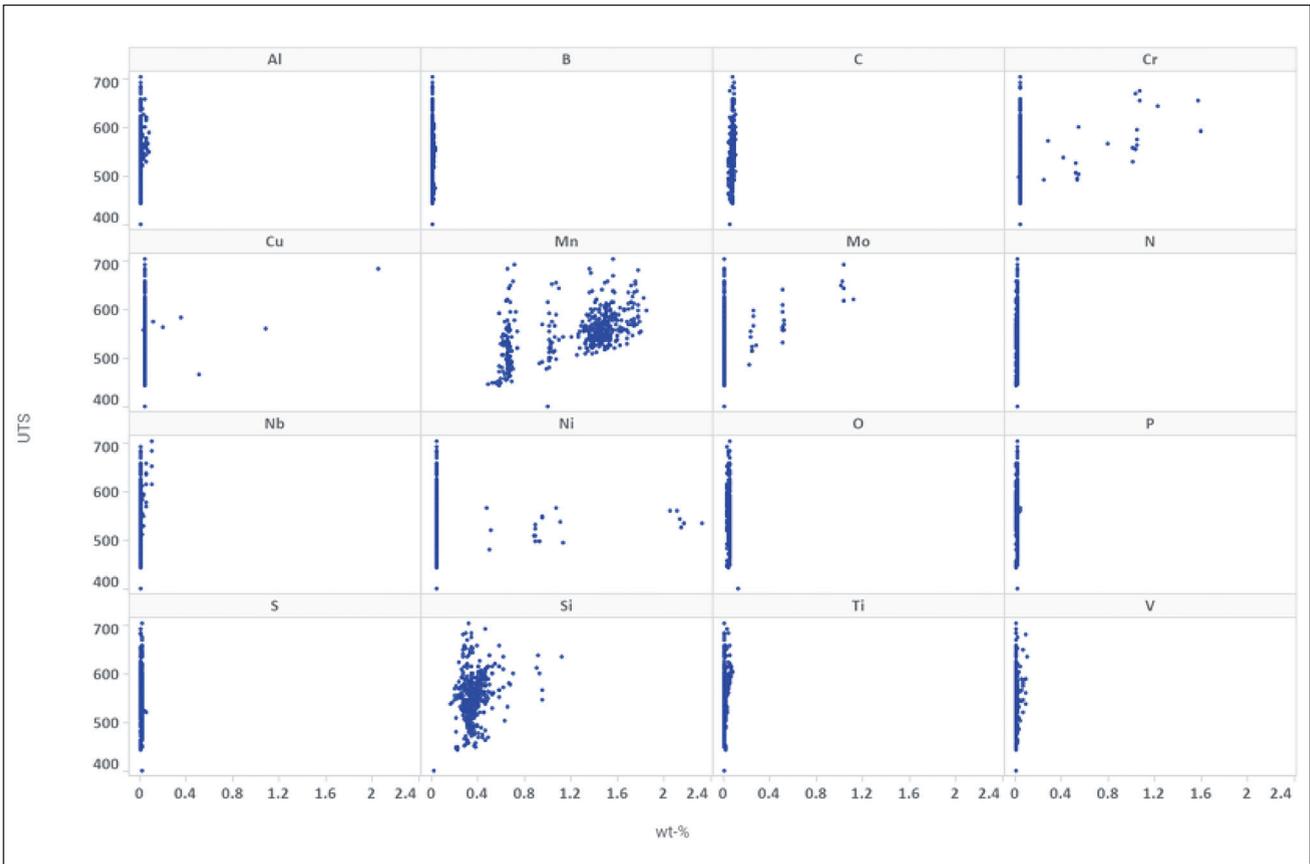


Fig. 4 – Trellis plot of element amounts (in wt-%) against WM UTS in Evans's dataset.

that the error values associated with predictions were much less than 30°C for each of these welds. In other words, one could use the -60°C CVN test temperature for 28 J absorbed energy as a benchmark to distinguish welds with balanced Ti, B, Al, N, and O contents.

While the JWES ANN template is available to predict CVN test temperature for 28 J absorbed energy based on WM chemical composition, a complementary relationship or ANN template involving WM UTS and WM chemical composition is currently unavailable.

Modeling of WMs

In recent years, there has been a growing interest to develop computer-based models on WM mechanical properties, particularly WM tensile and impact or fracture toughness properties based on WM chemical composition. For the most part, these modeling activities primarily involved ANNs, which are emerging as powerful tools with the capacity to reconstruct a database on weld properties through data selection and augmentation.

These modeling activities implicitly recognize that suppressing austenite decomposition or lowering solid-state phase transformation temperatures induces greater nucleation rates and refines the resultant microstructural constituents (Ref. 13), thereby enhancing weld mechanical properties. However, there had been no explicit modeling efforts to correlate WM tensile and CVN impact toughness

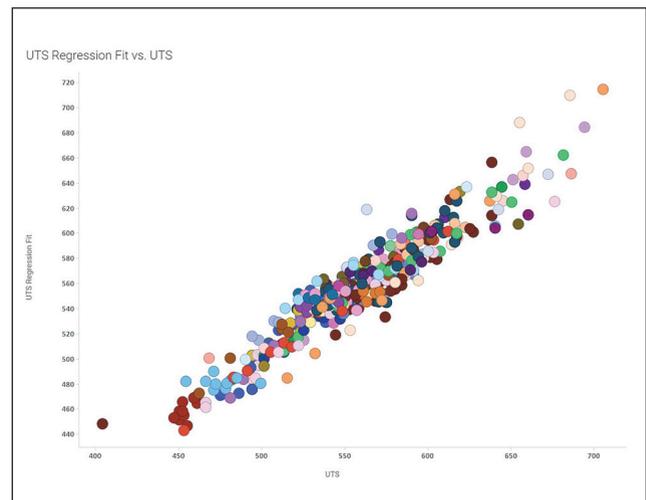


Fig. 5 – The regression fit of UTS against the values reported in Evans's database.

or fracture appearance transition temperature (FATT) with WM chemical composition through austenite (decomposition) transformation temperatures. Availability of such a correlation or relationship between WM UTS and WM chemical composition could complement the JWES ANN currently available to predict CVN test temperature for 28 J absorbed energy based on WM chemical composition, thus providing a pair of effective tools for efficient development

Table 2 — Test Results of Transformation Temperature of 13 TiBAlN Weld Metals at 13°C/s Cooling Rate (Refs. 9–11)

Weld ID	Ti (wt-%)	B (wt-%)	Al (wt-%)	N (wt-%)	O (wt-%)
O	0.0001	0.0001	0.0006	0.0079	0.0475
O2	0.0005	0.0005	0.0005	0.0235	0.0399
X	0.041	0.0002	0.0001	0.0077	0.0282
X2	0.045	0.0002	0.0005	0.0249	0.0297
Y	0.039	0.0039	0.0013	0.0083	0.0308
Y2	0.041	0.0044	0.0005	0.0232	0.0292
Z	0.042	0.0048	0.016	0.0067	0.0438
Z2	0.047	0.0045	0.018	0.023	0.044
U	0.039	0.0158	0.0005	0.0084	0.029
U2	0.039	0.0167	0.0005	0.0217	0.0297
V	0.054	0.0056	0.058	0.0041	0.044
V1*	0.048	0.0044	0.056	0.012	0.0473
V2	0.043	0.0035	0.056	0.0235	0.047
Low	0.0001	0.0001	0.0001	0.0041	0.0282
High	0.054	0.0167	0.058	0.0249	0.0475
Range	0.0539	0.0166	0.0579	0.0208	0.0193

*The transformation temperature for weld V1 was obtained from interpolation of graphical data reported in Ref. 11.

Table 2 – (continued)

(Ti + B + Al + N + O) (wt-%)	UTS (MPa)	CVN Test Temperature (°C)		Transformation Temperature (°C)				
		@ 100 J	@ 28 J	T _S	T ₅₀	T _{PRTT}	T _F	(T _S -T _F)
0.0562	528	-14	-42	762	658	650	554	208
0.0649	607	20	-16	754	630	606	534	220
0.0772	577	-61	-77	760	660	630	568	192
0.1003	631	-30	-58	760	650	638	572	188
0.0833	594	-82	-98	710	625	618	560	150
0.0983	605	-24	-56	703	612	606	510	193
0.1133	640	-83	-100	710	645	642	550	160
0.1365	583	13	-18	703	635	632	560	143
0.0927	586	-53	-80	700	626	620	531	169
0.1076	541	-52	-81	765	662	650	574	191
0.1657	732	-12	-46	680	598	596	507	173
0.1677	644	-64	-93	760	680	640	555	205
0.173	591	-45	-70	754	644	642	588	166
0.0562	528	-83	-100	680	598	596	507	143
0.173	732	20	-16	765	680	650	588	220
0.1168	204	103	84	85	82	54	81	77

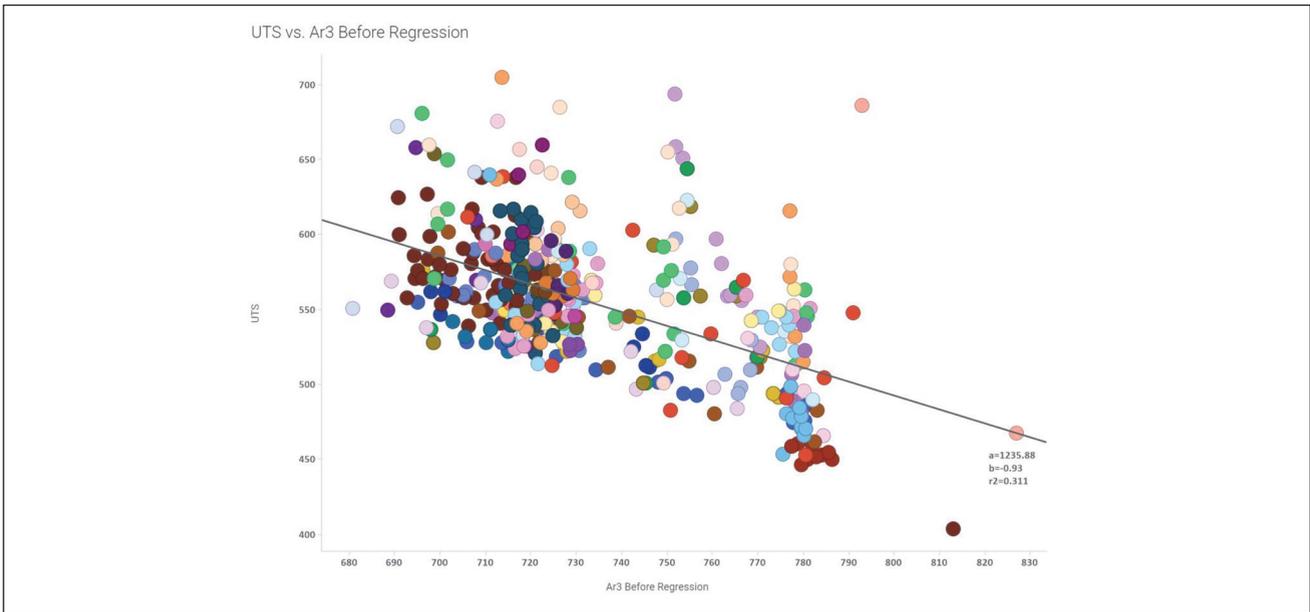


Fig. 6 – A regression of coalesced A_{r3} values (Salganik et al., Trzaska, and Ouchi et al.) against WM UTS.

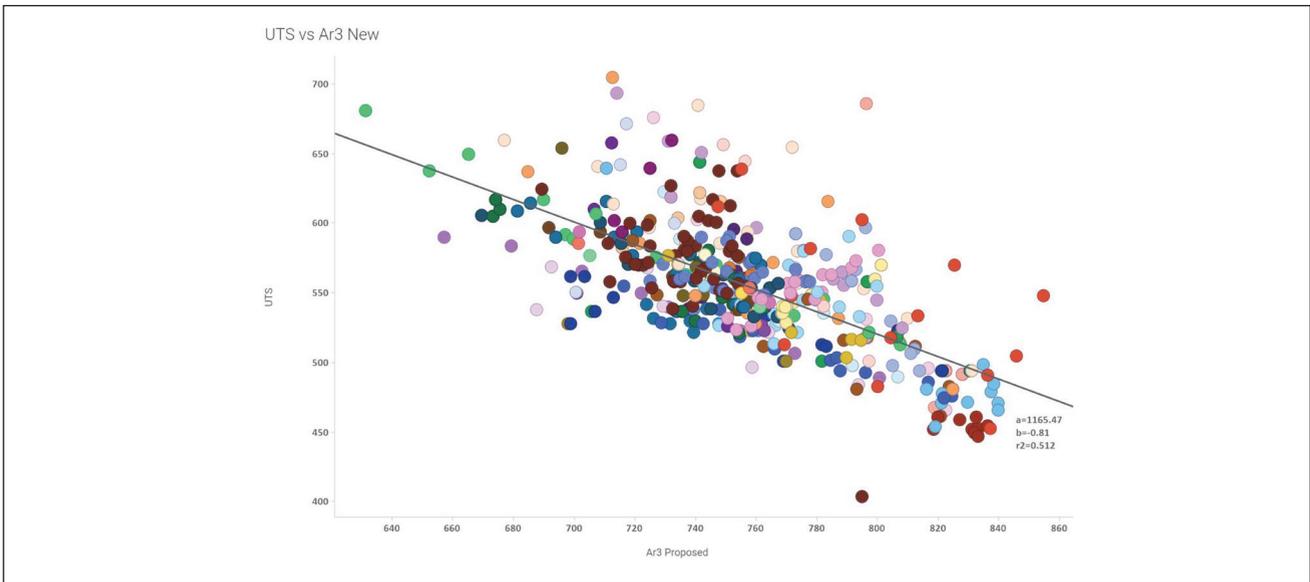


Fig. 7 – A plot of new A_{r3} values against WM UTS.

of welding electrodes based on an Fe-C-Mn system for high-performance applications.

Fujii and Ichikawa (Ref. 14) developed an ANN that could predict weld properties, including WM strength, FATT, and hardness. In addition, as one of the characteristics of their prediction, they identified that unreliability in estimated values can be displayed by the magnitude of the error bar. According to this system, the magnitude of the error bar was dependent on input conditions (test conditions) at the time; for example, where the data dispersion was large and reliability was low, the error bar was displayed as large and the computer itself was equipped with a function that could display the reliability of its prediction. The prediction of this error bar substantially extended the application scope of

conventional neural networks and allowed the possibility of their application to the reconstruction of databases related to various properties.

The University of Cambridge performed an ANN analysis of a vast and fairly general database assembled from publications on WM properties involving YS, UTS, elongation, and CVN impact toughness of ferritic steel WMs expressed as functions of chemical composition, heat input during welding, and postweld heat treatment (Refs. 15–18). This effort also used Evans's SMA WM database on Fe-C-Mn WMs (Refs. 1–2). The outputs of the model were assessed in a variety of ways, including specific studies of model predictions for the Fe-C-Mn and Fe-2.25Cr-1Mo systems. Comparisons were also made with corresponding methods that used simple

Table 3 — Predicted CVN Test Temperature of 13 TiBAIN Series Weld Metals with Balanced TiBAINO Additions

Weld ID	Measured CVN Test Temperature (°C) @ 28 J	CVN Predicted Test Temperature (°C) @ 28 J	
		Original Weld	Balanced Weld
0	-42	-49	-82
O2	-16	-15	-88
X	-77	-80	-90
X2	-58	-59	-89
Y	-98	-102	-94
Y2	-56	-54	-89
Z	-100	-94	—
Z2	-18	-28	-89
U	-80	-79	-92
U2	-81	-78	-84
V	-46	-47	-95
V1	-93	-93	-90
V2	-70	-71	-90

physical metallurgical principles. The models appeared to capture vital metallurgical trends and emulate expectations from current physical metallurgy principles yet required much more systematic experimental data to improve the accuracy of their predictions.

The U.S. Navy used an ANN technique developed by MacKay with a Bayesian framework wherein the probability of occurrence is interpreted as a reasonable expectation of a state of current knowledge but allows estimation of error bars like the ones introduced by Fujii and Ichikawa (Ref. 14). It also warns the user when data is sparse or locally noisy. This ANN was

trained and tested on a set of data obtained from WMs of various steel types used for shipbuilding (Ref. 19). The input variables for the network used WM chemical composition and weld cooling rate. The output consisted of YS and UTS, elongation, and reduction of area. This effort created many models using different network configurations and initial conditions. The method revealed significant trends describing the dependence of WM mechanical properties on WM chemical composition and cooling rate.

The U.S. Navy also used a similar ANN approach to model WM toughness characterized by CVN and dynamic tear tests

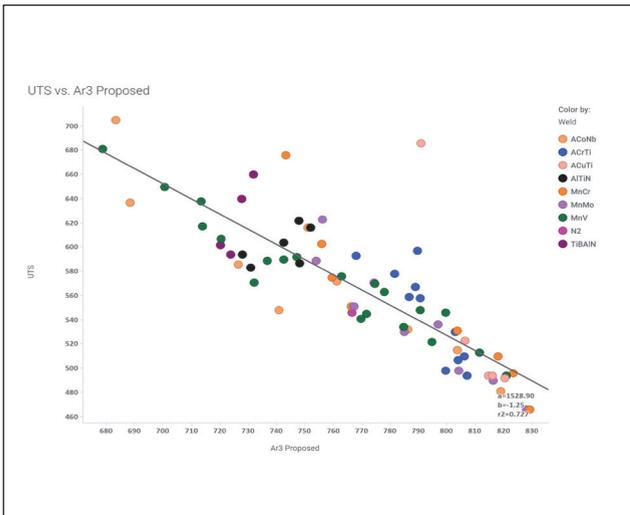


Fig. 8 — A plot of UTS against new A_{r3} values (Equation 17) in clusters with a strong negative correlation.

of the same types of steels used for shipbuilding (Refs. 20, 21). The level of noise in the experimental data was perceived to be high, but it nevertheless allowed one to recognize reasonable trends and uncertainties when making predictions. For example, the WM toughness showed a nonlinear deterioration as the WM oxygen concentration increased, yet this behavior could be assessed quantitatively.

Recently, Kim et al. (Ref. 22) performed an ANN analysis of Evans's SMA WM database and investigated the effect of WM chemical composition on WM mechanical properties, including YS and UTS, and test temperature on CVN impact toughness testing to provide 100 J absorbed energy. Based on the data collected from previously performed experiments, Kim et al. developed correlations between related variables, analyzed the results, and offered predictive models. They prepared sufficient datasets using data augmentation techniques to overcome problems caused by insufficient data and enable better predictions. Finally, they developed closed-form equations based on the predictive models to evaluate WM mechanical properties according to WM chemical composition. Each ANN model developed in this study considered changes in the content of only two elements. The study is mainly useful to predict the relative increase or decrease according to the change in the content of any two elements.

A recent research effort by Xiong et al. (Ref. 23) applied machine learning to predict mechanical properties of steels. The investigators selected 360 data on four mechanical properties (fatigue strength, tensile strength, fracture strength, and hardness) of both carbon steels and low-alloy steels from the National Institute for Materials Science (NIMS) database. They applied five machine learning algorithms on the 360 datasets to predict mechanical properties and determined that random forest regression provided the best correlation among the four most important features (tempering temperature and alloying elements of C, Cr, and Mo) for the mechanical properties of steels. They also used symbolic regression to generate mathematical expressions that explic-

itly predicted how each of the four mechanical properties varied quantitatively with the four most important features.

Objectives

The objective of the current effort was to use machine learning (Refs. 24, 25) for the following:

- 1) Determine a new expression for A_{r3} temperature applicable to Evans's SMA WM database that includes the effects of 16 principal and minor alloy elements (in wt-%) and weld cooling rate (in °C/s),
- 2) Use the new expression to develop a relationship with WM UTS, and
- 3) Perform a cluster analysis to gain additional insights.

It is well known that the austenite-to-ferrite (A_{r3}) transformation temperature is a function of alloy additions, prior austenite grain size, cooling rate, and possibly prior thermomechanical processing history. However, Evans's database didn't have information on prior austenite grain size and weld cooling rate. Only the 13 WMs (Refs. 9–11) that were subjected to dilatometry evaluation recorded the weld cooling rate. Consequently, development of a new A_{r3} expression based on Evans's database is likely to have several limitations. However, as most material specifications and welding electrode specifications require the UTS to exceed a certain minimum value, one could still derive exceptional benefit from a new A_{r3} expression that would be helpful in correlating A_{r3} with WM UTS.

Procedure

A review of the metallurgical literature revealed numerous formulae on A_{r3} expressions (Refs. 26–31). For example, Gorni (Ref. 32) in the *Steel Forming and Heat-Treating Handbook* documented a large collection of formulae that included regression equations for A_{r3} temperatures of several types of steels. A few of these equations include various microalloy additions along with cooling rates or cooling times from 800° to 500°C, or $\Delta t_{8/5}$. However, many times, the A_{r3} formulae reported in the literature are applicable only over limited ranges of element compositions and other predictors such as grain size, strain rate, etc. In support of the current research effort, some A_{r3} equations, including Ouchi et al.'s expression (Equation 4), seemed to be suitable starting points. Still, they are limited in chemical composition range and predictive capabilities because they do not include all the 16 alloying elements reported in Evans's database.

In addition to Ouchi et al.'s A_{r3} regression expression, two other A_{r3} regression equations, mentioned below, were considered for use in the current research effort.

The equation by Salganik et al. is as follows (Ref. 26):

$$A_{r3}(^{\circ}\text{C}) = 735.6 + 180.1(C + Cr) + 1206.9(S + P) - 10.9(Si + Mn + Ni + Cu + Mo) + 755.3(Al + N) - 328.8(V + Nb + Ti) \quad (8)$$

The equation by Trzaska is as follows (Ref. 27):

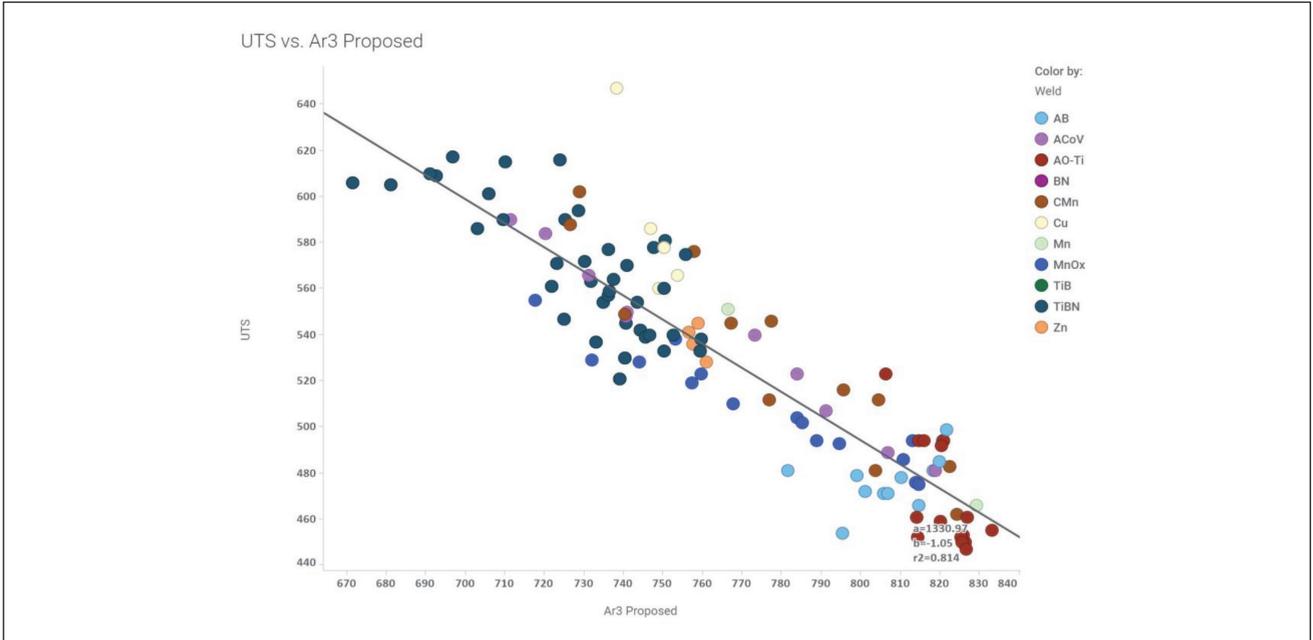


Fig. 9 – A plot of UTS against new A_{r3} values (Equation 18) in clusters with a moderate negative correlation.

$$A_{r3}(\text{°C}) = 857 - 257(C) + 23(Si) - 69(Mn) + 26(Cu) - 38(Ni) - 20(Cr) - 20(Mo) + 34(V) + 0.07(T_A) - 17CR^{0.25} \quad (9)$$

where T_A is the austenitizing temperature in °C, and the CR is the cooling rate in °C/min.

The sign of Si in Salganik et al.'s A_{r3} expression seemed to interfere with the correct evolution of the regression coefficient for Si in the current work. It was set to +10.9, and the A_{r3} expression by Salganik et al. was modified as follows:

$$A_{r3}(\text{°C}) = 735.6 + 180.1(C + Cr) + 1206.9(S + P) + 10.9Si - 10.9(Mn + Ni + Cu + Mo) + 755.3(Al + N) - 328.8(V + Nb + Ti) \quad (10)$$

Similarly, the Cu term in Trzaska's A_{r3} expression seemed to interfere with the correct evolution of the Cu coefficient sign in the current work. So, it was changed to $-10(Cu)$, and Trzaska's A_{r3} expression was modified as follows.

$$A_{r3}(\text{°C}) = 857 - 257(C) + 23(Si) - 69(Mn) - 10(Cu) - 38(Ni) - 20(Cr) - 20(Mo) + 34(V) + 0.07(T_A) - 17CR^{0.25} \quad (11)$$

The data ranges for the applicability of these equations are given in Gorni's handbook (Ref. 32). Trzaska's equation for A_{r3} seemed to be applicable over almost the entire range of Evans's database, although it did not include all the alloying

elements. Salganik et al.'s expression for A_{r3} was applicable only to a limited number of records in Evans's database because of the narrow data range of this expression, though it included several minor alloying elements. The above expressions were used in this work as discussed below.

The initial data preparation involved an examination of Evans's database and addition of columns to include various CMCs, including a CEN (Ref. 4) and selected austenite decomposition temperatures, such as T_s or A_{r3} (Ref. 7) or B_s and M_s (Ref. 8), based on WM compositions using corresponding constitutive equations.

Following the aforementioned computation using Microsoft Excel, certain records that did not obey the relationship on the required ordering of these temperatures (i.e., $A_{r3} > B_s > M_s$) were eliminated. Also, some records containing no information in some columns were excluded. This first phase filtering yielded 858 records. Then the records were limited to the high wt-% of elements indicated by Trzaska's A_{r3} data range. This filtering yielded 809 records. Then two indicators were set on the data to indicate if the records were in Salganik et al.'s data range or in Trzaska's data range. Then A_{r3} values from the modified Salganik et al. (Equation 10) and modified Trzaska (Equation 11) equations were added to the records as appropriate. A coalesced A_{r3} column was also added to the data. Salganik et al.'s modified A_{r3} , Trzaska's modified A_{r3} , and Ouchi et al.'s A_{r3} (Equation 4) were coalesced in this order.

The above steps were taken to identify and select valuable records for regression analysis. These A_{r3} values provided tentative values that appeared in certain records that were added as high-end guideposts to regression. Subsequently, the data was further restricted to certain WM compositions wherein the CEN was limited to 0.3 maximum, carbon content was limited to 0.1 wt-% maximum, and nitrogen content was limited to 99 ppm (0.0099 wt-%) maximum. Also, records with coalesced A_{r3} values less than 680°C that appeared as strong outliers in an UTS vs. A_{r3} regression were excluded.

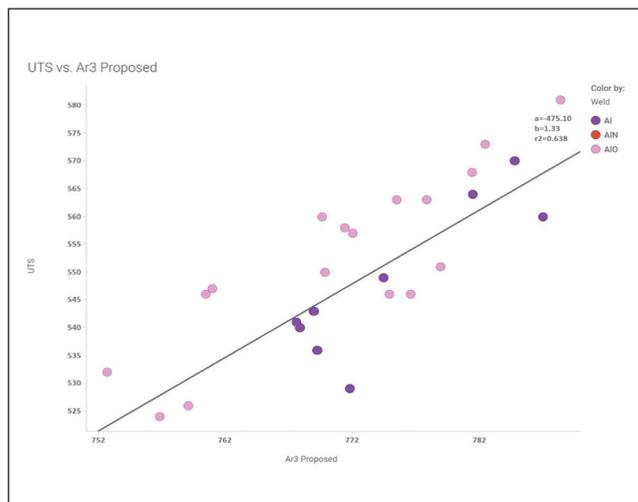


Fig. 10 — A plot of UTS against new A_{r3} values (Equation 19) in clusters with a strong positive correlation.

Data records with UTS values greater than 710 MPa were excluded. All these data filters reduced the number of records to 595 out of more than 900 records in Evans’s database.

To study potential regression equations for A_{r3} , the 13 WM records in the TiBAlN series that had experimental A_{r3} (i.e., T_p) values (Refs. 9–11) were extracted and combined with A_{r3} predictions for 21 other records that were within the range of Salganik et al.’s (Ref. 26) modified regression equation and 75 records found in the Mintz et al. dataset (Ref. 28). To make the new expression for A_{r3} applicable over the entire range in Evans’s dataset, 23 records with highest wt-% of elements and fractions of the highest wt-% in Evans’s dataset were also added with A_{r3} values predicted using the A_{r3} coalescing logic described earlier. Then additional data on A_{r3} and respective element compositions and cooling rates found in Refs. 29–43 were collected. These experimental datasets included a wide variety of low-alloy steel compositions and cooling rates and respective A_{r3} values. Also, 20 sets of numerical records on various steels with random element composition values at low range were included to improve the regression intercept and accuracy. A composite dataset created in this manner was used to derive a new regression equation for A_{r3} , including the effects of major and minor alloy elements (in wt-%) and weld cooling rate (in °C/s). Appendix I (aws.org/2023.102.004.appendix) provides the experimental datasets extracted from selected sources and used in this investigation.

A standard machine learning approach implemented in R (Refs. 24, 25) using linear model (`lm()`) and model identification using `regsubsets()` was used in the investigation, and the final formula was hand tuned as well. A data frame of linear and nonlinear predictors with the experimental or coalesced temperatures was prepared, and the model was generated as follows:

```
model = lm( $A_{r3} \sim .$ , data = df) or
model = regsubsets( $A_{r3} \sim .$ , data = df, method = ‘backward’)
```

where `df` is the data frame. Several functions were invoked on the model output to print out various properties of the model.

The function `summary(mdl)` is typically used to indicate the fit coefficients in the case of `lm()`, and the `summary()` function provides a high-level summary of various models generated when `regsubsets()` is used. Additional functions were invoked to probe the contents of the models further.

The output of `regsubsets()` contained various models tried, their R^2 , adjusted R^2 , C_p (a variant of Akaike information criterion or AIC, developed by Colin Mallows), and BIC (Bayesian information criterion) values for all models. C_p , BIC, and adjusted R^2 are the model metrics typically used to select an appropriate model.

Many linear and nonlinear predictors were provided to the models by adding respective columns in the data frame. In the beginning, many square and square root terms were added to the data frame for this purpose along with the main linear terms. Many models suggested by model identification were examined and modified. Usually, a model having the least Mallows’s C_p or BIC merits selection.

The current investigation also evaluated if the model was good in reproducing the decrease in A_{r3} temperatures for a few selected experimental records with the balance Ti, B, Al, N, and O content (Refs. 9–11). As a result, a few models recommended by automated model selection having the lowest Mallows’s C_p were examined, and a manual selection was made based on a few required model properties — such as negative coefficients for primary alloying elements, a positive coefficient for Si, and a negative coefficient for cooling rate — and the model satisfactorily predicting the observed A_{r3} decreases in the five experimental records related to the presence of balanced Ti, B, Al, N, and O content.

Overall, 257 data records on WM or low-alloy steels were combined in an R data frame to determine a new expression for A_{r3} . The rest of the analysis used 595 data records obtained after applying various filters to Evans’s database.

Results and Discussion

Various data properties and the results of correlations with the new regression formulae for the A_{r3} temperature are described below.

Initially, the elemental levels in Evans’s database were plotted as shown in Fig. 4. These Trellis plots revealed that carbon content in the data was largely centered around 0.075 wt-% while manganese content seemed to form two major clusters. Cr, N, Mo, Cu, Nb, Al, and B contents largely clustered close to 0 wt-% though much higher numbers were found in some data records. Si, Ti, N, O, S, and P contents seemed to form largely single clusters with few outliers. V content seemed to cluster largely around two values. These clusters and outliers would create similar clusters or spreads in dependent variables such as UTS, YS, and CVN at 28 J. CVN test temperatures at 28 and 100 J vs. UTS showed wide scatter.

A formal analysis following data preparation allowed an assessment of the effect of certain predictors on WM UTS and A_{r3} temperature.

Table 4 – Minimum and Maximum Limits for Elemental Composition and Cooling Rate

Element	Minimum (wt-%)	Maximum (wt-%)
C	0.024	0.792
Si	0	2.04
Mn	0	2.52
P	0	0.11
S	0	0.046
Cu	0	2.04
Ni	0	3.49
Cr	0	2.8
Mo	0	1.11
Nb	0	0.098
V	0	0.099
Ti	0	0.069
B	0	0.02
Al	0	1.55
N	0	0.0270
O	0	0.118
Cooling Rate (°C/s)	0.001	30

UTS vs. Predictors

Initially, a regression of UTS using multiple linear terms was obtained as follows:

$$\begin{aligned}
 UTS(MPa) = & 294 + 844(C) + 112.8(Si) + 79.2(Mn) + \\
 & 72.3(Cu) + 32.4(Ni) + 83.3(Cr) + 151.5(Mo) + \quad (12) \\
 & 1773.3(Nb) + 995.1(V) + 855.3(Ti) + \\
 & 289(Al) + 3638.2(N)
 \end{aligned}$$

Figure 5 shows the UTS fit given by Equation 12 plotted against reported WM UTS values in Evans's database. It is interesting to note a near linear relationship with less scat-

ter in UTS between the 450 and 600 MPa range, which is consistent with the data on ferritic-pearlitic steels shown in Fig. 1. This UTS regression was performed over 595 records in Evans's database obtained after applying several filtering conditions as described earlier.

Simple regression that included most of the elements except P and S showed an adjusted R^2 of 0.9118. This is a pretty good fit despite a few key unknowns about the WM, such as (average) prior-austenite grain size and cooling rate. The p-values of all the coefficients were well less than 0.05. The coefficients of B, O, and S were not significant. The coef-

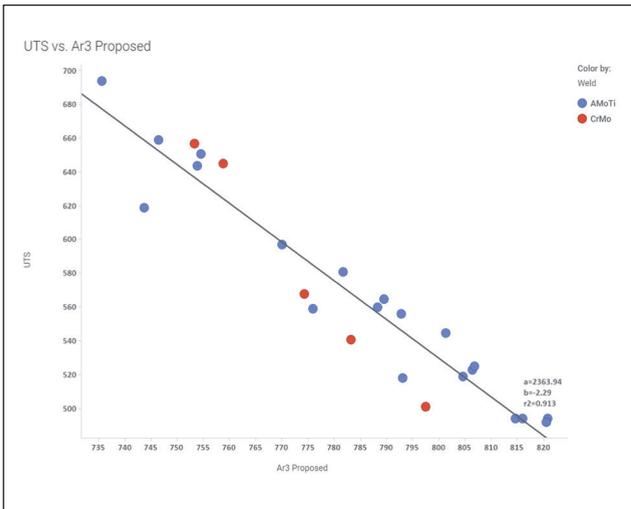


Fig. 11 – A plot of UTS against new A_{r3} values (Equation 20) in another cluster with a very strong negative correlation.

ficient for P was barely significant, and its effect was small; so, it was excluded from the equation.

The regression shown in Equation 12 for WM UTS appeared similar to the following regression equation by Mesplont (Ref. 29), which is suitable for high-strength bainitic steels with C content below 0.8 wt-% and the following range of compositions: Mn < 2 wt-%, Si < 1.8 wt-%, Cr < 2 wt-%, Mo < 0.8 wt-%, Cu < 1.6 wt-%, Ti < 500 ppm, P < 700 ppm, Nb < 800 ppm, and B < 30 ppm.

$$UTS(MPa) = 288 + 803(C) + 178(Si) + 83(Mn) + 1326(P) + 60(Cu) + 122(Cr) + 320(Mo) + 2500(Nb) + 180(Ti) + 36000(B) \quad (13)$$

The adjusted R^2 of the fit improved when the grain size was approximately calculated and added to the model. This can be done using the equation for YS that contains grain size (Ref. 5) and further making an assumption about dissolved nitrogen.

UTS is also found to be linearly correlated with YS. Regression expressions for CVN at 28 and 100 J could not be identified.

A_{r3} vs. UTS

Subsequently, the coalesced A_{r3} values assigned to 595 records in the analysis set were correlated with experimental values for WM UTS reported in Evans's database. A regression of computed A_{r3} against UTS showed a poor fit with an adjusted R^2 of only 0.304, as shown in Fig. 6. This poor fit is understandable as Ouchi et al. and Trzaska's equations for A_{r3} do not include all principal and minor alloy additions, cooling rate, or (average) prior-austenite grain size.

During the filtering process, records having a computed A_{r3} value lower than 680°C were excluded. If these records were included, they would appear as strong outliers in Fig. 6. The correlation factor between the initial computed A_{r3} and WM UTS was -0.551 .

As mentioned earlier, Evans's WM dataset did not provide a direct correlation between computed values of A_{r3} and experimental values of WM UTS. Consequently, the proposed new expression for A_{r3} was formulated as discussed below.

A_{r3} Regression Using Ilman's Experimental Data

The 13 experimental A_{r3} (i.e., T_s) values provided separately by Ilman et al. (Refs. 9–11) are well explained by cooling rate and material compositions, although certain elemental weight percentages remain constant. To mitigate the adverse effects of multicollinearity issues, the data was augmented using a few records reported in Lolla et al. (Ref. 35), Vega et al. (Ref. 36), and Deva et al. (Ref. 38).

A good A_{r3} regression model for this small dataset was as follows:

$$A_{r3}(^{\circ}C) = 1185.95 - 6.55(CR) - 534.02(C) + 394.05(Si) - 330.45(Mn) + 1586.63(Ti) + 64477.78(B) + 3093.8(Al) + 4606.92(N) + 7827.45(O) - 24156.63(Ti^2) - 842.71(\sqrt{Al}) - 1502.42(\sqrt{O}) - 1653160.12(Ti \times B) - 79158.05(Ti \times N) - 91474.94(Al \times N) \quad (14)$$

where CR is the cooling rate in $^{\circ}C/s$. The cooling rate for Ilman et al.'s experimental weld data is set at 13 $^{\circ}C/s$. The adjusted R^2 of the above regression is 0.95. This regression expression indicated that cooling rate and a few of the major elements (in wt-%) are good predictors of A_{r3} . The decreases observed in some experimentally determined values of T_s were especially well predicted by this model. The prediction error is less than 1% for all the data included in the model.

It was mentioned earlier that a balanced addition of Ti, B, Al, N, and O significantly decreased the measured T_s vs. the calculated A_{r3} values. This can be observed in the regression model without the inclusion of the cooling rate as a predictor. As shown in Table 2, the sum of Ti, B, Al, N, and O minor alloy additions decreased as the experimental T_s value increased in this small dataset. This seemed to be a good indicator. But the adjusted R^2 of the model, including only the alloying elements, was not high even though it offered good predictive capabilities. Once the cooling rate was added as a predictor and the effects of the multicollinearity issue were mitigated, the adjusted R^2 of the model improved significantly.

A_{r3} Regression Using Composite Data

It is quite apparent that more experimental data is needed to refine the A_{r3} regression from Ilman et al.'s data to achieve higher reliability and to make the formula applicable over a much wider data range. The 75 experimental records provided by Mintz et al. (Ref. 28) were included first. Additionally, 21 records in Evans's dataset that were in Salganik et al.'s (Ref. 26) data range were selected and assigned A_{r3} values using Salganik et al.'s modified A_{r3} expression (Equation 10). Another 23 records from Evans's dataset with coalesced A_{r3} values were added to constrain the regression at applicable data boundaries and interior points and to guide the regres-

sion process. Additional experimental data reported in Refs. 29, 30, and 33–43 were also included. Appendix I provides details of the respective datasets. Additionally, 20 numerical steel records with random element concentrations at a very low range around 0.01 for primary elements and around 0.0001 for minor elements were added using Ouchi et al.'s expression (Equation 4). Overall, 257 records were collected to obtain a new A_{r3} expression using a multiple linear regression. The experimental data along with coalesced A_{r3} values at selected extreme and interior points in Evans's dataset served as guideposts to regression in lieu of full experimental data on WM transition temperatures in Evans's database.

The regression expression for A_{r3} from this composite dataset was obtained as follows:

$$A_{r3}(\text{°C}) = 906.49 - 2.78(CR) - 439.3(C) + 34.17(Si) - 36.7(Mn) - 8.5(Cu) - 51.2(Ni) - 27.08(Cr) - 63.48(Mo) - 1765.95(Nb) - 520.29(V) - 2401.12(Ti) - 1784.44(B) + 21.89(Al) + 5300.15(N) - 420.96(O) + 297.07(C^2) - 16.4(Mn^2) + 11668.54(Nb^2) + 458.21(\sqrt{Ti}) - 1142.45(\sqrt{N}) + 298.91(\sqrt{O}) \quad (15)$$

where CR is the cooling rate in °C/s. The adjusted R^2 of this fit was 0.9087. The standard error of the residuals was 24.89. The intercept value of this new A_{r3} expression was close to 910°C, the A_{e3} (equilibrium austenite-ferrite transformation) temperature for pure iron.

Several nonlinear terms were included to cover a wide range of 14 elemental compositions (except P and S) in Evans's dataset. Most of the p-values of the intercept and the coefficients of C, Mn, Ni, Cr, Mo, Si, Ti, Nb, N, C^2 , Mn^2 , Nb^2 , Al, V, \sqrt{Ti} , \sqrt{O} , and \sqrt{N} were very significant or marginally significant. The coefficients of Cu and O also did not have p-values below 0.05, but they were left as indicated by regression because their values were in the expected range. The new A_{r3} regression equation predicted the decreases observed in measured values of T_s reported in the experimental data in Table 2 within a 2 to 3% error range in most cases. This equation is also likely to predict A_{r3} values over the entire range of Evans's dataset reasonably well.

The above new A_{r3} regression equation for WM UTS appears similar to Miettinen's (Ref. 44) A_{r3} regression equation, which also includes several nonlinear terms. It is also partly similar to the complex expressions for transition temperatures reported by Kasatkin et al. (Ref. 31).

Other nonlinear terms and cross terms are recommended in the context of critical temperature and CCT diagram models by Miettinen et al. (Ref. 45). The regression equation by Miettinen et al. (Ref. 45) also refers to the cross term $Cu \times B$, among others. The model from the current investigation also predicted the significance of this cross term. The quadratic term Nb^2 is also recommended by Yuan et al. (Ref. 46). For the sake of simplicity and to assess the impact of major and minor elements in Evans's database, the new regression expression for A_{r3} shown in Equation 15 was felt sufficient for the intended purpose.

The new A_{r3} regression formula is applicable over almost the entirety of Evans's database. The related data limits for various elements and weld cooling rate are shown in Table 4.

Figure 7 shows a plot of UTS against the proposed new A_{r3} . A linear regression of UTS against the proposed new A_{r3} is indicated by the straight line in Fig. 7. The line passes through the data better compared to the one indicated in Fig. 6. The adjusted R^2 of the fit of the new A_{r3} against UTS is 0.499 compared to 0.3058 in the previous fit in Fig. 6. This indicates that the new A_{r3} is likely to better predict UTS. The regression relation is the following:

$$UTS \text{ (MPa)} = 1242.93 - 0.91 A_{r3}(\text{°C}) \quad (16)$$

Cluster Analysis

The scatter in Fig. 7 is attributed to various inherent metallurgical characteristics of different clusters of experimental data in Evans's dataset. Many of these metallurgical characteristics include the amount of free nitrogen; size and composition of the inclusions and precipitates; effect of tempering caused by the deposition of over-lying runs; microphase morphology (e.g., the form of the carbides); relative grain sizes of the coarse-grained, fined-grained, and the intercritical regions; etc. (Ref. 47).

The individual experimental clusters were indicated by weld ID tags in the data with different colors in Fig. 7. The scatter can be better explained when each experimental cluster is examined and combined with similar individual experimental clusters. Evans's database contains 73 weld experimental clusters. The pruned dataset of 595 records contains 55 weld clusters. These individual weld clusters exhibit various trends in UTS vs. the new A_{r3} values. Some individual weld clusters have strong down trends, several of them have even up trends, and many of them do not have strong trends in UTS vs. A_{r3} . Some weld clusters have only one or two data points. The rest have three or more data points. Table 5 shows the statistics on selected individual weld clusters and the regression of UTS vs. the new A_{r3} of each weld cluster using the R^2 value, intercept and slope of the line, and trend indicator. Table 5 also shows the minimum and maximum WM UTS (in MPa) in the respective weld series and the corresponding number of data points in the weld series.

The above individual clusters can be grouped further. This was accomplished manually using TIBCO Spotfire® (Ref. 48). When the individual weld clusters are grouped, the information can be condensed in a few charts to explain the scatter in Fig. 7, and the impact of combining various microelements on UTS can also be understood more clearly. In general, the trend between UTS and A_{r3} temperature should be as indicated in Fig. 1. However, Evans's WM data contains at least four types of composite weld clusters: a cluster of clusters with a strong or moderate negative correlation between UTS and A_{r3} , a cluster of clusters with a positive correlation between UTS and A_{r3} , and a composite cluster of clusters with a very strong negative correlation. The respective data are illustrated in Figs. 8–11.

Table 5 — Regression and Summary Statistics of Selected Weld Series in Evans's Database

Weld ID	R ²	Intercept	Slope	Trend	Minimum UTS (MPa)	Maximum UTS (MPa)	Count
AB	0.251	81.238	0.493	Up	454	499	12
ACoNb	0.876	1600.655	-1.390	Down	481	705	9
ACoV	0.910	1355.749	-1.076	Down	481	590	10
ACrTi	0.766	1473.870	-1.276	Down	494	597	11
ACuTi	0.623	4176.206	-4.508	Down	492	686	6
Al	0.658	-629.981	1.520	Up	529	570	9
AlN	0.657	-627.983	1.518	Up	529	570	9
AlO	0.616	-517.523	1.375	Up	524	581	26
AlTi	0.341	1122.213	-0.747	Down	514	622	33
AlTiN	0.407	-1588.414	3.030	Up	583	622	6
AMoTi	0.840	2333.428	-2.248	Down	492	694	20
ANiTi	0.714	884.939	-0.475	Down	492	562	15
AoPlus	0.281	942.227	-0.531	Down	494	619	8
AO-Ti	0.382	1436.715	-1.177	Down	447	523	15
BN	0.196	1112.977	-0.786	Down	521	561	7
CB	0.453	1397.047	-1.135	Down	521	617	44
CMn	0.898	1610.519	-1.382	Down	462	602	12
CoPlus	0.740	1413.219	-1.177	Down	526	658	5
Cplus	0.473	1296.832	-0.965	Down	533	672	5
CrMo	0.899	1570.725	-1.357	Down	501	657	5

Table 5 – (continued)

Weld ID	R ²	Intercept	Slope	Trend	Minimum UTS (MPa)	Maximum UTS (MPa)	Count
MnCr	0.854	1554.366	-1.330	Down	466	676	8
MnMo	0.904	2190.718	-2.089	Down	466	623	9
MnNb	0.843	1506.371	-1.219	Down	494	685	17
MnNi	0.692	1156.997	-0.823	Down	466	569	10
MnOx	0.881	1086.432	-0.748	Down	475	555	15
MnSi	0.301	1362.136	-1.023	Down	453	639	14
MnTi	0.777	1215.296	-0.879	Down	492	577	17
MnV	0.847	1504.494	-1.210	Down	494	681	20
S	0.393	-2332.078	3.723	Up	523	546	5
Tab	0.205	1133.158	-0.756	Down	404	638	45
Ti	0.678	1822.660	-1.669	Down	537	654	9
TiB	0.474	1349.502	-1.066	Down	521	617	39
TiBAlN	0.554	4341.670	-5.254	Down	594	660	4
TiBN	0.474	1349.502	-1.066	Down	521	617	39
TiN	0.517	1335.580	-1.041	Down	528	597	10
TiOX	0.466	-560.604	1.473	Up	539	596	7
Zn	0.242	1442.055	-1.186	Down	528	545	4

In general, the experimental series indicated by weld IDs in the data can be joined to form a few major clusters. However, the conventional cluster analysis, such as K-means clustering (Ref. 24), cannot be performed for this dataset. The weld series can be combined when their UTS vs. A_{r3} regression slopes are nearly the same and if the distances between two

regression lines are small or their combined regression has nearly the same characteristics, such as adjusted R² before and after adding in the new weld series to the composite group. The trends of the weld series in the composite clusters were manually identified, and the weld series were grouped

using Spotfire. The underlying weld series are indicated in the legends in Figs. 8–11.

The respective simple regression equations of the four major clusters between UTS and A_{r3} are as follows:

Strong negative correlation

$$UTS (MPa) = 1528.9 - 1.25A_{r3}(\text{°C}) \quad (17)$$

Moderate negative correlation

$$UTS (MPa) = 1330.97 - 1.05A_{r3}(\text{°C}) \quad (18)$$

Strong positive correlation

$$UTS (MPa) = -475.1 + 1.33A_{r3}(\text{°C}) \quad (19)$$

Very strong negative correlation

$$UTS (MPa) = 2363.94 - 2.29A_{r3}(\text{°C}) \quad (20)$$

The R^2 values of the regression lines in Figs. 8–11 were 0.727, 0.814, 0.638, and 0.913, respectively.

Most of the remaining weld series can be combined into another composite cluster. The adjusted R^2 value of the regression of UTS vs. A_{r3} in this cluster was 0.55.

Figures 8, 9, and 11 indicate that UTS decreased with A_{r3} in general. Figure 10 indicates that UTS increased with A_{r3} for a small number of cases, but this increase strongly depended on WM elemental compositions. The element combinations in the first cluster (Fig. 8) are likely to yield quite predictable weld strengths. The element combinations in the second cluster (Fig. 9) with moderate negative correlation may yield stronger welds though UTS may be less predictable. A_{r3} values below 680°C seemed to provide UTS above 640 MPa. Figure 10 indicates that aluminum and nitrogen contents created an unexpected positive correlation between UTS and A_{r3} , and the respective UTS values were below 580 MPa. It can be noted that the charts indicate different ranges of UTS for different combinations of the underlying experimental clusters. Figures 8 and 11 show clusters that can yield a maximum UTS close to 700 MPa. Weld experimental data on certain groups such as TiBAlN also indicate the possibility of achieving UTS greater than 710 MPa. These high UTS values from the TiBAlN series were not included in the analysis. So, they do not appear in Figs. 8 or 9.

The above correlation or relationship between WM UTS and the new expression for A_{r3} temperature that includes the effects of principal and minor alloy elements (in wt-%) and weld cooling rate (in °C/s) are expected to complement the JWES neural network currently available to predict CVN test temperature for 28 J absorbed energy based on WM chemical composition, thus providing a pair of effective tools for efficient development of welding electrodes based on an Fe-C-Mn system for high-performance applications.

Conclusions

Several new relationships have been obtained using the selected SMA WM dataset in Evans's database. A new equation for A_{r3} temperature in low-alloy steels containing several minor elements was derived using a novel approach. The analysis was carried out using a machine learning approach for multiple linear regression in R.

Initially, all the records in this database were examined from various perspectives. The previously discussed charts and regression results indicate salient perspectives of this data and the new results reported.

The SMA WM database is quite large compared to typical datasets reported in metallurgical literature. Understanding the message conveyed by this large database is quite daunting. Machine learning approaches, including automated model finding using multiple regressions, improve our productivity in discovering the relationships embedded in the data.

The analysis reported here indicates that WM UTS can be correlated linearly with WM elemental composition. This regression equation improved notably after the inverse square root of approximate grain size was added as another predictor. The WM UTS had a linear correlation with YS. The test temperature for CVN at 28 or 100 J absorbed energy did not correlate fully well with elemental compositions, UTS, YS, and A_{r3} temperature. Only a poor-quality regression could be obtained even after including nonlinear terms in the regression equation of CVN at 28 J.

Evans's database is supported by only a small dataset of experimentally determined A_{r3} (i.e., T_s) values, and this database does not provide critical temperatures for all the reported WMs in the database. This was rectified in our effort. The 13 experimental values of records containing A_{r3} (or T_s) were combined with various experimental data indicated in Appendix I, a small number of records in Evans's dataset that obeyed data range for Salganik et al.'s expression for A_{r3} (Ref. 26), and 23 extreme high-end points in Evans's dataset. Twenty numerical records on various steels were also added at very low element concentrations, and their A_{r3} values were set using Ouchi et al.'s expression (Equation 4). Using this composite dataset, a new A_{r3} regression relation applicable to the entirety of Evans's database was derived. This expression includes several nonlinear predictors, and it was designed to include a large data range in Evans's dataset and to accommodate some nonlinearities expected in the fit and in the underlying physical phenomena. The expression also improves prediction accuracy for five records in question in Ilman's experimental data. This expression should also predict A_{r3} for all records in Evans's database to a very good approximation.

The relation between UTS and A_{r3} exhibits a significant scatter and low-regression quality. Closer examination using a cluster analysis revealed that individual weld series in Evans's database contributed to the scatter, and they can be combined into at least four clusters, namely, one cluster having a strong negative correlation between UTS and A_{r3} , another having a moderate negative correlation between UTS and A_{r3} , a third small cluster having a strong positive correlation between UTS and A_{r3} , and a fourth cluster having

a very strong negative correlation between UTS and A_{r3} . The second and fourth clusters seemed to provide clues to create the strongest welds with UTS approaching 700 MPa.

Future research may generate more experimental data to improve the regression result for the A_{r3} expression applicable to Evans's WM dataset. Various WM transition temperatures and cooling rates could be determined experimentally and documented for a few additional points in Evans's WM dataset, including those with elemental composition in the high wt-% range. A full factorial design does not appear to be necessary.

Acknowledgments

Krishna Sampath expresses his gratitude to Dr. Glyn M. Evans for his continued encouragement, technical guidance, and support. The authors also remain thankful to the AWS peer reviewers for their insightful comments that enabled the authors to further improve the organization and content of the manuscript.

References

1. Evans, G. M. 2015. Database — Weld metal composition and properties. DOI: 10.13140/RG.2.1.3628.1764
2. Evans, G. M., and Bailey, N. 1999. *Metallurgy of Basic Weld Metal*. Abington, UK: Abington Publishing.
3. Sampath, K. 2021. Analysis of a high-strength steel SMAW database. *Welding Journal* 100(12): 410-s to 420-s. DOI: 10.29391/2021.100.036
4. Yurioka, N., Suzuki H., Ohshita, S., and Saito, S. 1983. Determination of necessary preheating temperature in steel welding. *Welding Journal* 62(6): 147-s to 153-s.
5. Bramfitt, B. L. 1998. "Structure/Property Relationships in Irons and Steels." In *Metals Handbook Desk Edition, Second Edition*, edited by Davis, J. R., 153–173. Materials Park, Ohio: ASM International. DOI: 10.31399/asm.hb.mhde2.a0003090
6. Pickering, F. B. 1978. *Physical Metallurgy and the Design of Steels*. London, UK: Applied Science Publishers Ltd.
7. Ouchi, C., Sampei, T., and Kozasu, I. 1982. The effect of hot rolling condition and chemical composition on the onset temperature of γ - α transformation after hot rolling. *Transactions of the Iron and Steel Institute of Japan* 22(3): 214–222. DOI: 10.2355/isijinternational1966.22.214
8. Steven, W., and Haynes, A. G. 1956. The temperature of formation of martensite and bainite in low-alloy steels. *Journal of the Iron and Steel Institute* 183(8): 349–359.
9. Ilman, M. N., Cochrane, R. C., and Evans, G. M. 2012. Effect of nitrogen and boron on the development of acicular ferrite in reheated C-Mn-Ti steel weld metals. *Welding in the World* 56: 41–50. DOI: 10.1007/BF03321394
10. Ilman, M. N., Cochrane, R. C., and Evans, G. M. 2014. Effect of titanium and nitrogen on the transformation characteristics of acicular ferrite in reheated C-Mn steel weld metals. *Welding in the World* 58: 1–10. DOI: 10.1007/s40194-013-0091-x
11. Ilman, M. N., Cochrane, R. C., and Evans, G. M. 2015. The development of acicular ferrite in reheated Ti-B-Al-N-type steel weld metals containing various levels of aluminium and nitrogen. *Welding in the World* 59: 565–575. DOI: 10.1007/s40194-015-0231-6
12. Sampath, K. 2022. Metallurgical design rules for high-strength steel weld metals. *Welding Journal* 101(5): 123-s to 143-s. DOI: 10.29391/2022.101.010
13. Bhadeshia, H. K. D. H. 2007. Frontiers in the modelling of steel deposits. *Journal of the Japan Welding Society* 76(2): 24–30. DOI: 10.2207/jjws.76.102
14. Fujii, H., and Ichikawa, K. 2001. Estimation of weld properties by Bayesian neural network. *Welding International* 15(12): 935–939. DOI: 10.1080/09507110109549469
15. Lalam, S. H. 2000. Modelling of mechanical properties of ferritic weld metals. PhD diss., Cambridge University.
16. Lalam, S. H., Bhadeshia, H. K. D. H., Evans, G. M., and MacKay, D. J. C. 2000. Estimation of mechanical properties of C-Mn weld metals. Paper read at *IIV Subcommittee IIA Autumn Meeting*, Sandvik, Sweden.
17. Lalam, S. H., Bhadeshia, H. K. D. H., and MacKay, D. J. C. 2000. Estimation of mechanical properties of ferritic steel welds. Part 1: Yield and tensile strength. *Science and Technology of Welding and Joining* 5(3): 135–147. DOI: 10.1179/136217100101538137
18. Lalam, S. H., Bhadeshia, H. K. D. H., and MacKay, D. J. C. 2000. Estimation of mechanical properties of ferritic steel welds. Part 2: Elongation and Charpy toughness. *Science and Technology of Welding and Joining* 5(3): 149–160. DOI: 10.1179/136217100101538146
19. Metzbower, E. A., DeLoach, J. J., Lalam, S. H., and Bhadeshia, H. K. D. H. 2001. Neural network analysis of strength and ductility of welding alloys for high strength low alloy shipbuilding steels. *Science and Technology of Welding and Joining* 6(2): 116–124.
20. Metzbower, E. A., DeLoach, J. J., Lalam, S. H., and Bhadeshia, H. K. D. H. 2001. Analysis of toughness of welding alloys for high strength low alloy shipbuilding steels. *Science and Technology of Welding and Joining* 6(6): 368–374.
21. Metzbower, E. A., DeLoach, J. J., Lalam, S. H., and Bhadeshia, H. K. D. H. 2002. "Secondary Effects in Neural Network Analysis of the Mechanical Properties of Welding Alloys for HSLA Shipbuilding Steels." In *Mathematical Modelling of Weld Phenomena 6*, edited by Cerjak, H., and Bhadeshia, H. K. D. H., 231–242. London, UK: Maney Publishing.
22. Kim, J.-H., Jung, C.-J., Park, Y. I., and Shin, Y.-T. 2022. Development of closed-form equations for estimating mechanical properties of weld metals according to chemical composition. *Metals* 12(3): 528. DOI: 10.3390/met12030528
23. Xiong, J., Zhang, T., and Shi, S. 2020. Machine learning of mechanical properties of steels. *Science China Technological Sciences* 63(7): 1247–1255. DOI: 10.1007/s11431-020-1599-5
24. James, G., Witten, D., Hastie, T., and Tibshirani, R. 2013. *An Introduction to Statistical Learning: with Applications in R*. Berlin, Germany: Springer Nature.
25. Weisberg, S. 1980. *Applied Linear Regression*. Hoboken, N.J.: John Wiley & Sons Inc.
26. Salganik, V. M., Chikishev, D. N., Pozhidaeva, E. B., and Nabat-chikov, D. G. 2016. Analysis of structural and phase transformations in low-alloy steels based on dilatometric studies. *Metallurgist* 59(9): 766–773. DOI: 10.1007/s11015-016-0172-3
27. Trzaska, J. 2015. Empirical formulae for the calculation of austenite supercooled transformation temperatures. *Archives of Metallurgy and Metals* 60(1): 181–185. DOI: 10.1515/amm-2015-0029
28. Mintz, B., Banerjee, J. R., and Banks, K. M. 2011. Regression equation for A_{r3} temperature for coarse grained as cast steels. *Ironmaking & Steelmaking* 38(3): 197–203. DOI: 10.1179/030192310X12827375731429
29. Mesplont, C. 2002. Phase transformations and microstructure-mechanical properties relations in complex phase high strength steels. PhD diss., University of Lille. DOI: 1854/333
30. Lotter, U. 1988. Aufstellung von Regressionsgleichungen zur Beschreibung des Umwandlungsverhaltens beim thermomechanischen Walzen, 17–18. Technical report, Thyssen Stahl.

31. Kasatkin, O. G., Vinokur, B. B., and Pilyushenko, V. L. 1984. Calculation models for determining the critical points of steel. *Metal Science and Heat Treatment* 26(1): 27–31. DOI: 10.1007/BF00712859
32. Gorni, A. A. 2011. *Steel Forming and Heat-Treating Handbook*, 24. São Vicente, Brazil. DOI: 10.13140/RG.2.1.1695.9764
33. Schindler, I., Nemec, J., Kawulok, P., Kawulok, R., Sevcak, V., Opela, P., and Ruzs, S. 2018. The combined effect of chemical composition and cooling rate on transformation temperatures of hypoeutectoid steels. *Kovové Materiály* 56(3): 163–170. DOI: 10.4149/km_2018_3_163
34. Zhang, C., Yang, Z., Enomoto, M., Chen, H., Yang, Z., and Zhang, C. 2016. Prediction of A_{r3} during very slow cooling in low alloy steels. *ISIJ International* 56(4): 678–684. DOI: 10.2355/isijinternational. ISIJINT-2015-602
35. Lolla, T., Babu, S. S., Lalam, S., and Manohar, M. 2013. Comparison of simulated heat affected zone microstructures of niobium microalloyed steels subjected to multi-pass weld thermal cycles. *Proceedings of International Seminar on Welding of High Strength Pipeline Steels*, 281–297. Araxá, Brazil: CBMM, Pittsburgh, Pa.: TMS.
36. Vega, M. I., Medina, S. F., Chapa, M., and Quispe, A. 1999. Determination of critical temperatures (T_{nr} , A_{r3} , A_{r1}) in hot rolling structural steels with different Ti and N contents. *ISIJ International* 39(12): 1304–1310. DOI: 10.2355/isijinternational.39.1304
37. Rodrigues, K. F., and Lúcio de Faria, G. 2021. Characterization and prediction of continuous cooling transformations in rail steels. *Materials Research* 24(5). DOI: 10.1590/1980-5373-MR-2020-0519
38. Deva, A., De, S. K., Kumar, V., Deepa, M., and Jha, B. K. 2013. Influence of boron on the hardenability of unalloyed and low alloyed steel. *International Journal of Metallurgical Engineering* 2(1) 47–51. DOI: 10.5923/j.ijmee.20130201.07
39. Symmach, S. M. 2016. Mechanical properties of steels with a microstructure of bainite/martensite and austenite islands. Master's thesis, Colorado School of Mines. hdl.handle.net/11124/170090
40. Zhang, Z., and Farrar, R. A. 1995. *An Atlas of Continuous Cooling Transformation (CCT) Diagrams Applicable to Low Carbon Low Alloy Weld Metals*. London, UK: Institute of Materials. DOI: 10.1201/9781003070405
41. Naik, A. K., Roshan, R., Arora, K. S., Shajan, N., and Mishra, S. C. 2021. Continuous cooling transformation diagram and mechanical properties in weld coarse-grain heat-affected zone of API-X70 steel. *Sadhana* 46(2): 88. DOI: 10.1007/s12046-021-01623-2
42. El-Shennawy, M., Farahat, A. I., Masoud, M. I., and Abdel-Aziz, A. I. 2016. Effect of boron content on metallurgical and mechanical characteristics of low-carbon steel. *International Journal of Mechanical Engineering* 5(2): 1–14.
43. Zavdoveev, A., Poznyakov, V., Baudin, T., Rogante, M., Kim, H. S., Heaton, M., Demchenko, Y., Zhukov, V., and Skoryk, M. 2021. Effect of heat treatment on the mechanical properties and microstructure of HSLA steels processed by various technologies. *Materials Today Communications* 28: 102598. DOI: 10.1016/j.mtcomm.2021.102598
44. Miettinen, J. 1996. Simple semiempirical model for prediction of austenite decomposition and related heat release during cooling of low alloy steels. *Ironmaking & Steelmaking* 23(4): 346–356.
45. Miettinen, J., Koskenniska, S., Somani, M., Louhenkilpi, S., Pohjonen, A., Larkiola, J., and Komi, J. 2021. Optimization of CCT curves for steels containing Al, Cu and B. *Metallurgical and Materials Transactions B* 52: 1640–1662. DOI: 10.1007/s11663-021-02130-9
46. Yuan, X. Q., Liu, Z. Y., Jioa, S. H., Ma, L. Q., and Wang, J. D. 2006. The onset temperatures of γ to α -phase transformation in hot deformed and non-deformed Nb micro-alloyed steels. *ISIJ International* 46(4): 579–585. DOI: 10.2355/isijinternational.46.579
47. Evans, G. M. 2022. The effect of micro-alloying elements on the ultimate tensile stress of C-Mn multi-run welds. IIW Doc. II-2445-2022 (II-C-619-2022), Tokyo, Japan.
48. TIBCO Spotfire®. tibco.com/products/tibco-spotfire

RAJAN VARADARAJAN is with Norfolk Southern, Atlanta, Ga. **KRISHNA SAMPATH** (krishna.sampath@icloud.com) is with Creative Concepts, Marietta, Ga.