Article:

Nowosad, K.K., Stillman, E. and Cogswell, D. (2026) Influence of individual specimens, estimation accuracy, and round-off errors on the master curve reference temperature. Journal of Testing and Evaluation, 54 (2). pp. 521-539. ISSN: 0090-3973

https://doi.org/10.1520/jte20250031

Journal of Testing and Evaluation

doi:10.1520/JTE20250031 / Vol. 54 / No. 2 / 2026 / available online at go.astm.org/pubs

Kamila K. Nowosad,[1] Eleanor Stillman,[2] and Daniel Cogswell[3]

Influence of Individual Specimens, Estimation Accuracy, and Round-off Errors on the Master Curve Reference Temperature

Reference

K. K. Nowosad, E. Stillman, and D. Cogswell, "Influence of Individual Specimens, Estimation Accuracy, and Round-off Errors on the Master Curve Reference Temperature," *Journal of Testing and Evaluation* 54, no. 2 (2026): 521–539. https://doi.org/10.1520/JTE20250031

Manuscript received January 23, 2025; accepted for publication July 11, 2025; published online September 12, 2025. Issue published March 1, 2026.

[1] School of Chemical, Materials and Biological Engineering, University of Sheffield, Sir Robert Hadfield Building, Mappin St., Sheffield, S1 3JD, UK (Corresponding author), e-mail: kknowosad1@sheffield.ac.uk, https://orcid.org/0000-0003-1245-1215

[2] School of Mathematical and Physical Sciences, University of Sheffield, Hicks Building, Hounsfield Rd., Sheffield, S3 7RH, UK https://orcid.org/0000-0002-5971-0985

[3] School of Chemical, Materials and Biological Engineering, University of Sheffield, Sir Robert Hadfield Building, Mappin St., Sheffield, S1 3JD, UK https://orcid.org/0000-0002-3401-2371

ABSTRACT

Fracture toughness of ferritic steels in the ductile-to-brittle transition region can be probabilistically described using the Master Curve approach standardized as ASTM E1921, *Standard Test Method for Determination of Reference Temperature, T0, for Ferritic Steels in the Transition Range*. The Master Curve reference temperature, T_0, is a point estimate based on direct fracture toughness testing results and can be used to generate lower bound fracture toughness values to be used in structural integrity assessments. Understanding the factors affecting the accuracy and variability of the T_0 estimation is crucial to comprehensively quantifying uncertainty in the T_0-based lower bound fracture toughness. In this work, simulated data were used to investigate the influence of individual specimens, which can exhibit vastly different fracture toughness levels, on the T_0 estimate. This was quantified with a simple novel leave-one-out statistic—*DFT0*. The effect of rounding errors in the ASTM E1921 equations, data set size, and test temperatures on the T_0 estimation accuracy were also assessed. It was found that high fracture toughness specimens shift the T_0 to lower values with considerably larger magnitude than the upward shifts caused by correspondingly low fracture toughness specimens. Specimens tested at lower temperatures tend to have stronger influence on T_0 than those tested at higher temperatures. The average *DFT0* values rarely exceeded 4°C in single-temperature datasets. The rounding errors in ASTM E1921 introduce a positive bias of around 0.2°C. The root mean square error of the reference temperature estimates can be reduced by increasing data set size, temperature, or both.

Introduction

Fracture toughness assessment is one of the key inputs to structural integrity assessments to ensure safe design, operation, and potential life extension of components. Using low-alloy steels in high integrity components requires that fracture behavior in the ductile-to-brittle transition (DBT) region is characterized to ensure the material would not become brittle when subjected to operational conditions. The main challenges to quantifying fracture toughness in the DBT region include the considerable degree of test data scatter because of the statistical nature of the cleavage initiation mechanism, the effect of thickness on fracture, and the expense of low temperature testing of sufficient number of valid specimens to describe the relationship between the fracture toughness and temperature adequately.

In 1997, the Master Curve method developed by Wallin[1–4] was standardized as ASTM E1921, *Standard Test Method for Determination of Reference Temperature, T0, for Ferritic Steels in the Transition Range*.[5] It introduced several benefits over the earlier indirect assessments via Charpy impact testing, such as use of precracked fracture toughness specimens tested in quasistatic conditions, the inbuilt framework for thickness scaling and the possibility of generation of fracture toughness values for a selected range of temperatures and defined failure probability levels using a single parameter, the reference temperature, which can be calculated using as few as six specimens[6–9]

Thickness scaling of the critical stress intensity factor is an important feature of the ASTM E1921 method, most notably in the context of testing specimens subjected to long-term neutron irradiation. Because of typically limited availability of irradiated material, significant research effort has been dedicated in recent years to the topic of applying the Master Curve approach to the fracture toughness results obtained from miniature specimens.[10–15]

Better understanding of uncertainty in the reference temperature would allow for more accurate definition of safe lifetime for safety-critical components, such as reactor pressure vessels in nuclear power plants or gas storage tanks and pipelines, if the T_0-based approach is applied to calculate lower bound fracture toughness. To achieve this, the effect of the test conditions and assumptions made requires quantification. Furthermore, given that the assessment can be made with the use of as few as six specimens tested at various temperatures, the influence of individual specimens on the calculated T_0 should also be understood.

In structural integrity assessments, a conservative reference temperature value is often used, typically calculated as the sum of T_0 and two standard deviations. ASTM E1921 provides an expression for estimating the standard deviation of T_0, which accounts for the T_0 estimation uncertainty, arising from applying the maximum likelihood (ML) estimation method, and experimental uncertainties. A value of $\sigma_{exp} = 4°C$ is recommended in ASTM E1921 to account for the latter, whereas the former is calculated as a function of the number of uncensored test results and the "sample size uncertainty factor," which is selected from a table in ASTM E1921 based on the median fracture toughness of the data set. The estimation uncertainty of T_0 has been a subject of previous research.[16–19] Notably, Sham and Eno[17] developed mathematical expressions to calculate the standard error of the T_0 estimate using a statistically rigorous approach based on the ML estimator's properties. It has been observed that the ASTM E1921 approach to quantification of the T_0 estimation uncertainty is conservative when standard specimen size (25.4 mm thick) is considered[17,19]; however, Quintin et al.[20] reported that the value is underestimated when smaller test specimens are used.

This paper aims to further understanding in this area by introducing a novel metric named *DFT0* ("difference in T_0") to quantify the influence of individual specimens on the T_0 estimate, quantifying the effect that rounding errors in the ASTM E1921 have on the T_0 estimation accuracy and assessing the effects of test temperature and sample size on the T_0 estimation accuracy. This is achieved by performing a comprehensive Monte Carlo simulation study aiming to model A508-4N—a modern pressure vessel steel.

The first two objectives provide a novel contribution to the understanding of the T_0 estimation uncertainties. Although the last objective has been previously addressed in the works cited above, this study applies a different methodology for data creation and analysis, thus supplementing the research performed in the aforementioned studies.

Master Curve and ASTM E1921

The Master Curve model, as per ASTM E1921, can be used to describe a wide range of ferritic steels (within the yield strength range 275 MPa and 825 MPa) in the DBT zone.[5] It addresses the main challenges of characterizing fracture toughness in the DBT region: the results scatter, the temperature dependence of fracture toughness, and the effect of specimen thickness. The reference temperature, T_0, is defined as the temperature at which the median fracture toughness $K_{Jc(P_f=0.5)}$ is equal to 100 MPa$\sqrt{\text{m}}$ for specimen thickness equal to the reference thickness, $B_0 = 25.4$ mm, often denoted as 1T (which slightly differs from the initial definition of T_0 which was for a 25 mm thick specimen[3,21]). The probability, P_f, that a B thick specimen will fail at stress intensity factor value of K_{Jc} is given by equation (1).

$$P_f = 1 - \exp\left[-\frac{B}{B_0}\left(\frac{K_{Jc} - K_{min}}{K_0 - K_{min}}\right)^m\right] \tag{1}$$

The model assumes that fracture toughness at a temperature T follows a three-parameter Weibull distribution. The three Weibull parameters are the shape parameter m, the scale parameter K_0, and the threshold parameter K_{min}. The shape parameter reflects the magnitude of scatter, the scale parameter is equal to the value of K_{Jc} for which the failure probability is 63.2 % (if $B = B_0$), and the threshold parameter represents the minimum stress intensity factor value below which cleavage fracture is not possible. The work leading to the derivation of equation (1) has been detailed in other publications.[1,4,22,23]

The Master Curve model is based on assumption that cleavage is a weakest-link event (i.e., for it to occur, at least one of critical fracture initiating particles must be sampled ahead of the macroscopic crack front leading to formation of a sharp microscopic crack). Additionally, for the initiated microscopic crack to propagate uncontrollably, adequate driving force is required, leading to incorporation of a threshold toughness in the model. The weakest-link assumption leads to a statistical thickness effect. A thicker specimen has a longer crack front meaning that the probability of sampling one of the randomly distributed fracture triggers increases, resulting in, on average, lower recorded fracture toughness. Equation (2), provided in ASTM E1921, allows for thickness correction of critical stress intensity factor to a 1T equivalent.

$$K_{Jc(B_0)} = K_{min} + (K_{Jc(B)} - K_{min})\left(\frac{B}{B_0}\right)^{\frac{1}{m}} \tag{2}$$

The temperature dependence of fracture toughness is defined by the Weibull scale parameter, K_0, through the following relationship

$$K_0 = K_{min} + a + b \cdot \exp[c \cdot (T - T_0)] \tag{3}$$

ASTM E1921 allows the use of compact tension, disk-shaped compact tension, or single-edge notched bend specimen geometries. The specimens can be tested either at the same temperature or at different temperatures. ASTM E1921 requires that all test temperatures must be within $T_0 \pm 50$°C for a test result to be included in the T_0 determination process.

The ML estimation method is used to obtain a T_0 estimate. Equation (4) must be iteratively solved to calculate the value of T_0. The steps to derive equation (4) from equation (1) are described in other works.[17,24] A right hand censoring procedure, detailed in the next section, is applied to fracture toughness values which exceed the specimen capacity limit. To account for this in the MLE calculations, each of N specimens is assigned a censoring indicator, δ, value of zero (censored) or one (uncensored). The number of uncensored specimens, r, required, as a minimum, by ASTM E1921 for a valid determination of reference temperature is between six and eight, depending on the test temperature.

$$\sum_{i=1}^{N} \delta_i \frac{exp[c \cdot (T_i - T_0)]}{a + b \cdot exp[c \cdot (T_i - T_0)]} - \sum_{i=1}^{N} \frac{(K_{Jc(1T,i)} - K_{min})^m exp[c \cdot (T_i - T_0)]}{\{a + b \cdot exp[c \cdot (T_i - T_0)]\}^{m+1}} = 0 \tag{4}$$

If the single-temperature (ST) method is used, the T_0 is typically calculated using a closed-form expression by solving equations (5) to (7).

$$K_0 = \left[\sum_{i=1}^{N} \frac{(K_{Jc(i)} - K_{min})^m}{r}\right]^{\frac{1}{m}} + K_{min} \tag{5}$$

$$K_{Jc(med)} = K_{min} + \ln(2)^{\frac{1}{m}}(K_0 - K_{min}) \tag{6}$$

$$T_{0(ST)} = T - \frac{1}{c}\ln\left(\frac{K_{Jc(med)} - K_{min} - a \cdot \ln(2)^{\frac{1}{m}}}{b \cdot \ln(2)^{\frac{1}{m}}}\right) \tag{7}$$

It is assumed that every material eligible for the ASTM E1921 assessment is described with identical and constant parameter values, which are presented in **Table 1**. The origin of these values can be found in articles by Wallin.[1–3]

The Master Curve model's applicability to different steels has been validated[4,19,25–27] leading to the consensus that fracture toughness scatter and temperature dependence can be described with the same model and parameters for most ferritic steels. T_0 is the only parameter that requires to be determined for a specific data set. Hence, the only difference between various materials is the location of the fracture toughness curve on the temperature axis.

It is noted that there are several round-off errors present in the equivalents of equations (4), (6), and (7) provided by ASTM E1921, which one should be conscious of when assessing the Master Curve uncertainties. To demonstrate the differences, the original equations were solved to computer system accuracy and are presented here to 5 decimal places. First rounding occurred with the establishment of the value of b. Through combining the T_0 definition with equations (1) and (3) substituted with the values of the Weibull distribution parameters from **Table 1**, it can be seen that for a 1T specimen at $T = T_0$:

$$a + b = \frac{K_{Jc(P_f=0.5)} - K_{min}}{\ln(2)^{\frac{1}{m}}} = \frac{100 - 20}{\ln(2)^{\frac{1}{4}}} \cong 87.67658 \tag{8}$$

When finding the values of a and b based on experimental data, the result of equation (4) has been rounded to the nearest integer and b was assumed to be equal to $(88 - a)$.[4,21] It is not known what the precise fitted value of a was, so the standard value of 11 $MPa\sqrt{m}$ is assumed in this work.

After this, the ASTM E1921 defines $K_{Jc(P_f=0.5,1T)}$ and $K_{Jc(P_f,1T)}$ as follows:

$$K_{Jc(P_f=0.5,1T)} = 30 + 70\exp[0.019((T - T_0)] \tag{9}$$

TABLE 1
Model parameters with fixed values in ASTM E1921

Parameter	m	K_{min}, $MPa\sqrt{m}$	a, $MPa\sqrt{m}$	b, $MPa\sqrt{m}$	c, °C^{-1}
ASTM E1921 value	4	20	11	77	0.019

$$K_{Jc(P_f,1T)} = 20 + \left[\ln\left(\frac{1}{1-P_f}\right)\right]^{\frac{1}{4}} \{11 + 77\exp[0.019(T - T_0)]\} \tag{10}$$

By solving equation (10) for $P_f = 0.5$, we can see that further rounding operations were performed to produce the parameters in equation (9) in the following manner:

$$20 + \ln(2)^{\frac{1}{4}} \cdot 11 \cong 30.03689 \cong 30 \tag{11}$$

$$\ln(2)^{\frac{1}{4}} \cdot 77 \cong 70.25821 \cong 70 \tag{12}$$

Methods

In this section, the process of the synthetic data generation and its analysis are described. A method of quantifying data points' influence is also introduced.

SYNTHETIC DATA GENERATION AND ANALYSIS

The first step of generating synthetic datasets for the ASTM E1921 analysis is to define the material parameters, specimen dimensions, and test conditions to be simulated. The objective was to simulate 1T specimens of the A508-4N material. The values of required parameters that were assumed in this work are presented in **Table 2**.

The algorithm for a data set generation is presented in **figure 1**. Every data set consisting of N specimens was generated by the calculation of K_{Jc} values using equation (13) for each of N randomly sampled values of failure probability, P_f, ranging between 10^{-10} and $(1 - 10^{-10})$. For each combination of data set size and test temperature, 30,000 datasets were simulated.

$$K_{Jc(Pf=0.xx,1T)} = 20 + \left[\ln\left(\frac{1}{1-0.xx}\right)\right]^{\frac{1}{4}} \left\{11 + \left(\frac{100-20}{\ln(2)^{\frac{1}{4}}} - 11\right)\exp[0.019(T - T_0)]\right\} [MPa\sqrt{m}] \tag{13}$$

To ensure the reproducibility and independence of each created data set, a unique seed value was generated for each combination of test temperature, data set size and iteration using the *hash()* function in Python. This approach was chosen for each simulation to represent an independent realization of data set generation process, allowing for a more generalized comparison between outcomes obtained for different conditions tested. Any data set can also therefore be independently fully reproduced if required, leading to higher computational efficiency because less information must be stored while the simulations are performed.

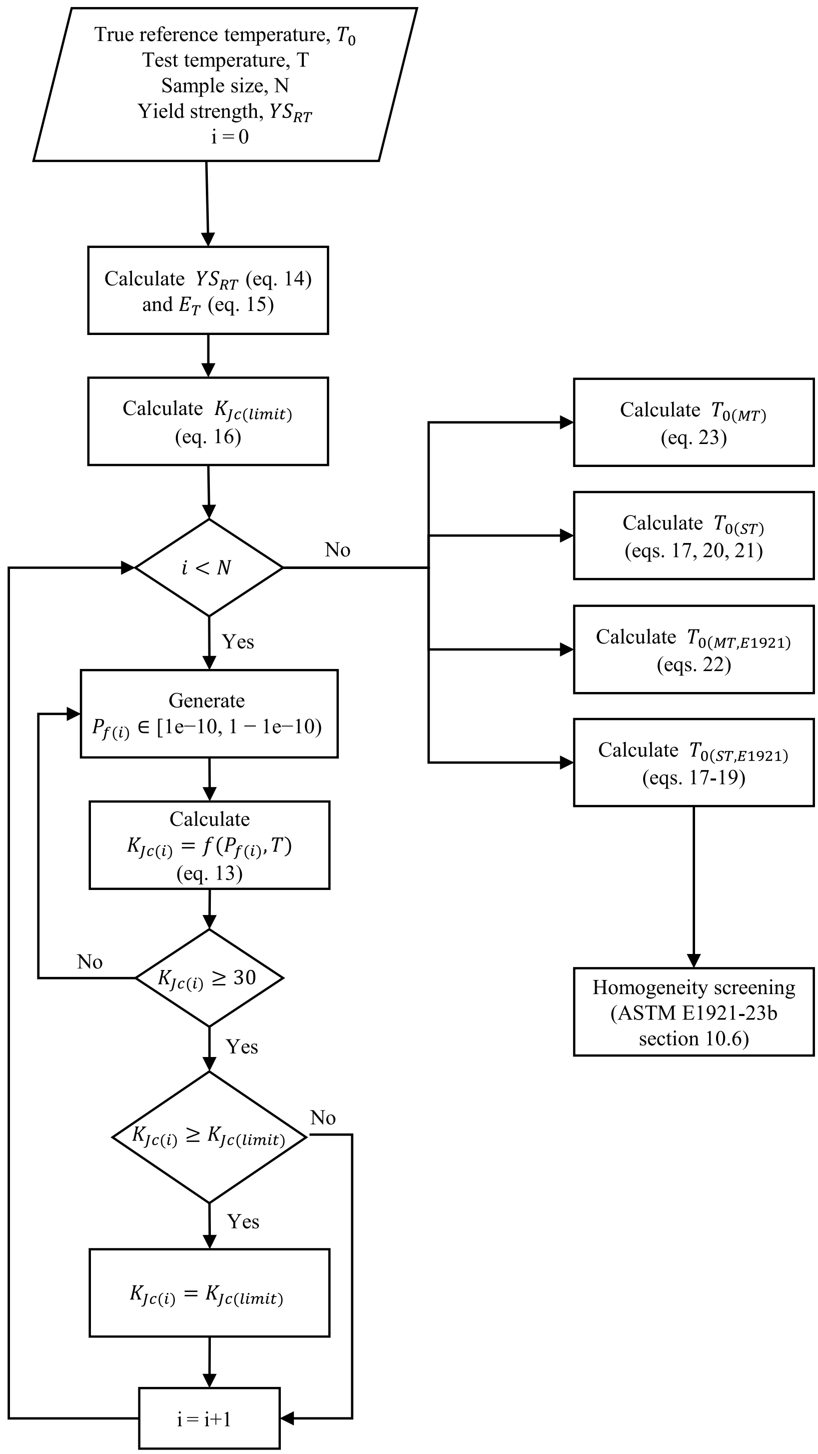
The minimum possible value of simulated K_{Jc} was set to 30 $\text{MPa}\sqrt{\text{m}}$ to avoid a computational issue when applying the procedure outlined in ASTM E1921 Appendix 5 to the data sets that are flagged as potentially inhomogeneous. This is a reasonable adjustment because lower fracture toughness values are extremely rare

TABLE 2
Assumptions used to generate the synthetic data

Parameter	Value (s)
Reference temperature, T_0, °C	−150
Room temperature yield strength, YS_{RT}, MPa	550
Test temperatures, $T - T_0$, °C	−50 to 50, in 10°C increments
Data set size, N	{6, 7, 8, 9, 10, 12, 14, 16, 20, 50}
Specimen thickness, B, mm	25.4
Specimen's remaining ligament, b_0, mm	25.4
Poisson's ratio, ν	0.3

FIG. 1

Algorithm for generation of synthetic data and calculation of reference temperature estimates.



for fully fractured specimens—there was only one instance reported in the analysis of 2,546 data points performed by Cogswell.[24] This critical stress intensity factor value of approximately 29 $MPa\sqrt{m}$ was recorded for a Cr-Ni-Mo-V reactor pressure vessel steel.[28]

Yield strength and elastic modulus values of the simulated material at any of the considered test temperatures were calculated using equations (14) and (15), which are provided in ASTM E1921.

$$YS_T = YS_{RT} + \frac{10^5}{491 - 1.8T} - 189[MPa] \tag{14}$$

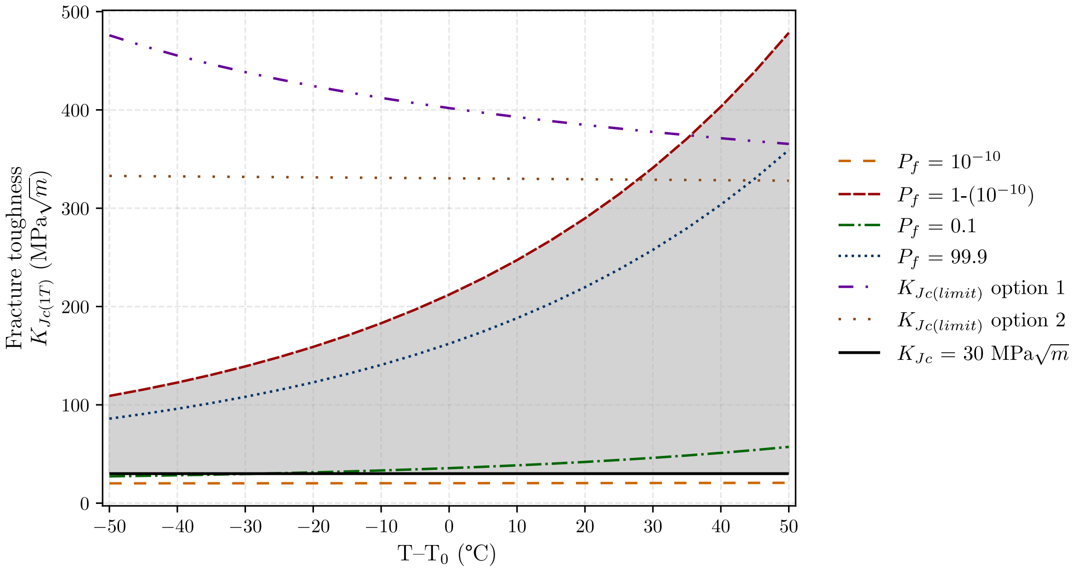
$$E_T = 204 - \frac{16}{T}[GPa] \tag{15}$$

All simulated values of K_{Jc} were tested against the ASTM E1921 specimen capacity limit defined by equation (16), which was used as the censoring value. After this, the correct values of the censoring indicator were assigned. As shown in **figure 2**, this limit would be reached in very rare instances for a material with high level of both strength and fracture toughness, such as the material simulated in this study. This figure illustrates the possible sampling space when simulating data for Master Curve analysis and highlights how the range of simulated data can be affected by some of the decisions one needs to take when creating a simulation. For instance, ASTM E1921 allows four methods of obtaining the yield strength value to be used in equation (16), the two most common of which are presented in **figure 2**.

$$K_{Jc(limit)} = \sqrt{\frac{E_T b_0 YS}{30(1-\nu^2)}} \tag{16}$$

$K_{Jc(limit)}$ – "option 1" refers to the use of temperature adjusted values of yield strength, YS_T, calculated for each test temperature using equation (14). This method was used in this study. In the "option 2" scenario, all of

FIG. 2 Master Curve-based sampling space (shaded) of fracture toughness values that can be used to create the synthetic datasets. $K_{Jc(limit)}$ – "option 1" refers to the use of temperature adjusted yield strength in equation (16), whereas option 2 refers to the use of room temperature yield strength value.



the $K_{Jc(limit)}$ values were calculated using the same yield strength value that was measured at a higher temperature than all test temperatures.

It can be also observed that, especially at higher test temperatures, there is a significant difference in the simulated high fracture toughness value depending on the definition of the upper bound (how close to $P_f = 1$ it is). For example, the maximum possible fracture toughness value in this study is equal to around 464 MPa$\sqrt{m}$, whereas at $T - T_0 = 50°C$, for the maximum of $P_f = 0.995$ used by Yount et al.,[18] the maximum possible simulated value would be around 339 MPa$\sqrt{m}$. This is a source of potential differences between the results obtained in different simulation studies, especially when the simulations involve test temperatures close to the upper limit prescribed by ASTM E1921.

Four estimates of T_0 were calculated for each data set. This was to quantify the effect that rounding of certain values in ASTM E1921 has on the reference temperature estimates and compare the results obtained using the multitemperature method and the single-temperature method. The first estimate of T_0 was calculated using equations (17)–(19), which constitute the ASTM E1921 single-temperature method.

$$K_0 = \left[\sum_{i=1}^{N} \frac{(K_{Jc(i)} - 20)^4}{r}\right]^{0.25} + 20 \tag{17}$$

$$K_{Jc(P_f=0.5,\,1T)} = 20 + 0.91(K_0 - 20) \tag{18}$$

$$T_{0(ST,E1921)} = T - \frac{1}{0.019}\ln\left(\frac{K_{Jc(P_f=0.5,\,1T)} - 30}{70}\right) \tag{19}$$

The second T_0 estimate was calculated using the single-temperature method again; however, the parameter values that are rounded in the aforementioned equations were replaced by the original expressions leading to equations (20) and (21). The logarithmic expression in equation (20) was computed to the computer precision and carried forward in this form forward to equation (21).

$$K_{Jc(P_f=0.5,\,1T)} = 20 + \ln(2)^{\frac{1}{4}}(K_0 - 20) \tag{20}$$

$$T_{0(ST)} = T - \frac{1}{0.019}\ln\left(\frac{K_{Jc(p_f=0.5,\,1T)} - 20 - 11\cdot\ln(2)^{\frac{1}{4}}}{\left(\frac{100-20}{\ln(2)^{\frac{1}{4}}} - 11\right)\cdot\ln(2)^{\frac{1}{4}}}\right) \tag{21}$$

The remaining two values of T_0 were calculated by solving the equation that can be applied to multitemperature datasets. The first version, described in equation (22), is as in ASTM E1921. A slightly different version, shown in equation (23), is again the maximum likelihood estimate iterative equation with the use of the Master Curve parameters but without the intermediate rounded values.

$$\sum_{i=1}^{N}\delta_i\frac{exp[0.019\cdot(T_i - T_{0(MT,E1921)})]}{11 + 77\cdot exp[0.019\cdot(T_i - T_{0(MT,E1921)})]} - \sum_{i=1}^{N}\frac{(K_{Jc(1T,i)} - 20)^4 exp[0.019\cdot(T_i - T_{0(MT,E1921)})]}{\{11 + 77\cdot exp[0.019\cdot(T_i - T_{0(MT,E1921)})]\}^5} = 0 \tag{22}$$

$$\sum_{i=1}^{N}\delta_i\frac{\exp[0.019\cdot(T_i-T_{0(MT)})}{11+\left(\frac{100-20}{\ln(2)^{\frac{1}{4}}}-11\right)\cdot\exp[0.019\cdot(T_i-T_{0(MT)})]}$$
$$-\sum_{i=1}^{N}\frac{(K_{Jc(1T,i)}-20)^4\exp[0.019\cdot(T_i-T_{0(\mathrm{MT})})]}{\{11+(\frac{100-20}{\ln(2)^{\frac{1}{4}}}-11)\cdot exp[0.019\cdot(T_i-T_{0(\mathrm{MT})})]\}^5}=0 \tag{23}$$

The datasets were subject to the homogeneity screening procedure outlined in ASTM E1921, using the $T_{0(ST,E1921)}$ value, and no potential inhomogeneity was detected, as would be expected for the datasets that are simulated from a single distribution.

The ASTM E1921 test temperature validity check was not performed because all data set sizes were simulated within the allowed test temperature range. It has been noted that single-temperature testing at the lower test temperature limit permitted by ASTM E1921 ($T-T_0=-50$°C) does not align with its requirement for the minimum median K_{Jc} of 58 $MPa\sqrt{\mathrm{m}}$ for a data set permitted to be used to calculate T_0. This requirement cannot be satisfied if $T-T_0=-50$°C because substituting this value into equation (9) leads to $K_{Jc(P_f=0.5,1T)}=57.07$ $\mathrm{MPa}\sqrt{\mathrm{m}}$.

The condition can only be met if T_0 is underestimated and, as a result, $T-T_{0(estimate)}$ is equal to at least −48.225°C. Around two thirds (68 %) of all datasets simulated at the lowest test temperature were characterized by the median fracture toughness of <58 $\mathrm{MPa}\sqrt{\mathrm{m}}$. For these datasets, homogeneity screening could not be performed because of the lack of a provided sample size uncertainty factor value. In rarer instances, this also occurred at other test temperatures (up to $T-T_0=-20$°C) when overestimation of T_0 led to to the $T-T_{0(estimate)}$ being lower than approximately −48.225°C. If censored values were included in a data set, its T_0 values were only recorded if they met the data set size validity criteria of the ASTM E1921; otherwise, a data set was resampled. It should be noted that the ASTM E1921 data set size criterion was impossible to satisfy for some of the low data set size and low test temperature combinations (e.g., $N=6$ at $T-T_0=-50$°C [because a minimum of eight specimens is required for valid T_0 determination at this test temperature]); however, the results are included in the paper because they still provide a valuable insight into the accuracy of the Master Curve T_0 estimation procedure.

DETECTING INFLUENTIAL OBSERVATIONS

There are a number of metrics to assess the influence of data points; however, they are typically presented in the context of linear regression, so adapting them for use in the Master Curve assessment is not necessarily trivial. Some of these statistics, such as *DFFIT* and *DFBETA* ("difference in fit" and "difference in beta," respectively), measure the effect that removing individual observations has on, respectively, model prediction and regression coefficient estimates.[29] This type of approach provides a very intuitive measure of the influence of an individual specimen on the calculated T_0 by simply calculating the difference in the estimate with and without a given specimen (i.e., difference in T_0). In this study, a metric called *DFT0* was defined by equation (24) and was calculated for each specimen in every data set.

$$DFT0_{(i)}=\widehat{T}_0-\widehat{T}_{0(-i)} \tag{24}$$

where,

$\widehat{T}_0$ is the T_0 estimate calculated using all observations, and

$\widehat{T}_{0(-i)}$ is the T_0 estimate calculated without the i-th observation.

At this stage, the authors have not studied the distribution theory of *DFT0*; however, the leave-one-out approach is widely applied as a measure of influence in statistics, and calculating the *DFT0* values for each specimen within a data set can clearly demonstrate whether any data point bears outstandingly large influence on the T_0 estimate. We illustrate our approach with a simple example. The *DFT0* values calculated for an assumed data set of seven specimens are presented in **Table 3**. This example illustrates that different specimens can have a

TABLE 3
DFT0 values calculated for an example data set

Specimen	K_{Jc}, MPa$\sqrt{m}$	T, °C	$\hat{T}_0$, °C	$\hat{T}_{0(-i)}$, °C	$DFT0$, °C
1	99.01	−150	−151.19	−152.33	1.14
2	101.39	−150	−151.19	−152.12	0.93
3	122.56	−150	−151.19	−148.88	−2.31
4	97.05	−150	−151.19	−152.49	1.30
5	118.01	−150	−151.19	−149.84	−1.35
6	71.96	−150	−151.19	−153.63	2.44
7	107.85	−150	−151.19	−151.42	0.23

positive or negative influence on the value of the calculated T_0, and the magnitude of this influence can vary. Note that, for this data set, tested at temperature near T_0, exclusion of even the most influential specimens shifted T_0 by <2.5°C. This is <7.9°C: the standard deviation of T_0 calculated for this data set using the ASTM E1921 procedure.

Results and Discussion

The results of a statistical analysis performed on the total of 2,970,000 simulated data sets are presented in this section. The specimens in each data set were simulated for the same test temperature. For all simulated datasets that were valid as per ASTM E1921, the homogeneity condition defined in the standard was satisfied. The shape of histograms showing the distribution of T_0 values obtained was similar for all simulated conditions with a normal-like curve, as shown in **figure 3**.

INFLUENCE OF INDIVIDUAL DATA POINTS

The influence of each simulated datum on the reference temperature estimate of the data set to which it belonged was assessed by calculating *DFT0*, as demonstrated in **Table 3**. To investigate the effect of test temperature, data set size, and failure probability on the sensitivity of T_0 estimates to individual specimens within different datasets,

FIG. 3 Distribution of T_0 estimates obtained for $T - T_0 = 0$°C and $N = 10$.

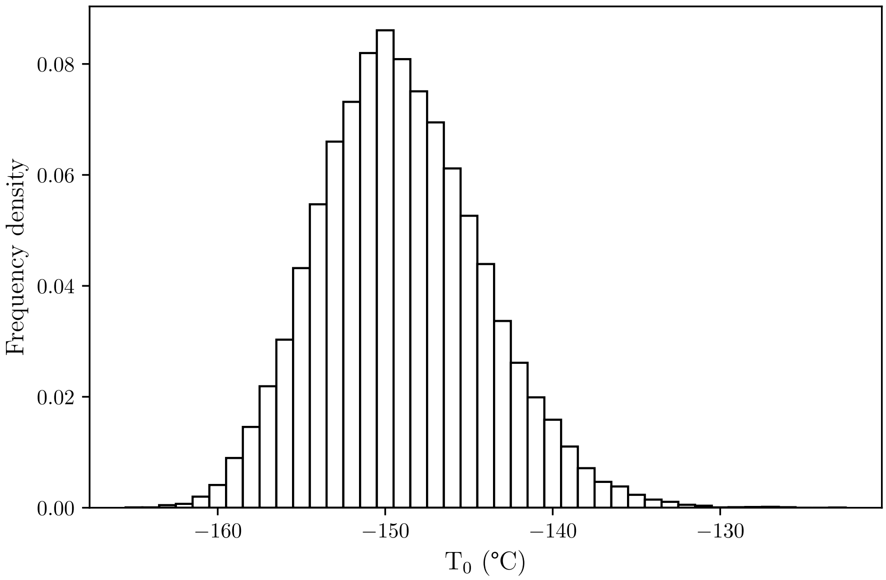
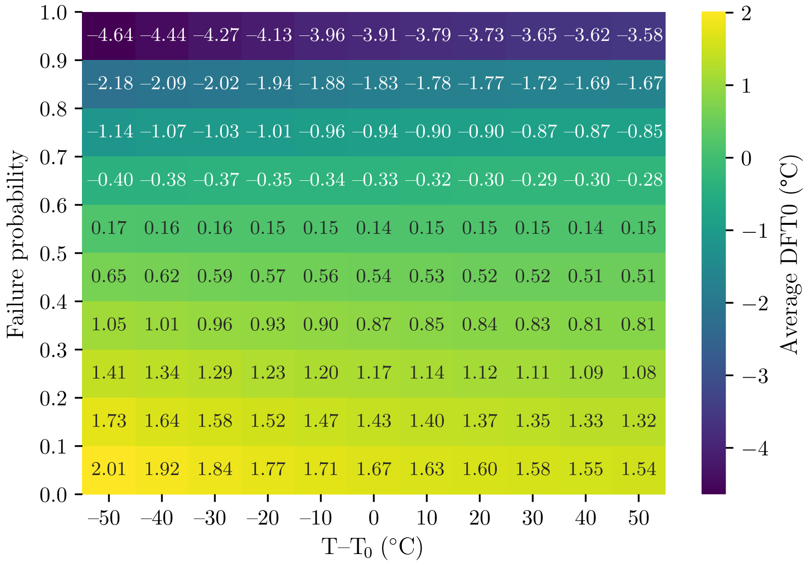


FIG. 4 Influence of individual specimens as a function of failure probability and test temperature for $N = 9$.



average *DFT0* values were calculated for ten discrete P_f bins (ranging from zero to one with increments of 0.1) for every data set size and test temperature combination. The full results for all datasets of nine specimens are presented in **figure 4**. To aid connection between the fracture toughness values and their influence, fracture toughness is presented as a function of probability and test temperature in **figure 5**, whereas **figure 6** presents results of the influence study for remaining simulated data set sizes. Common trends can be observed for all data set sizes analyzed. On average, specimens with the lowest influence on $\widehat{T}_0$ are characterized by fracture toughness corresponding the failure probability level between 0.5 and 0.6 and the *DFT0* values in that range do not vary much

FIG. 5 Relationship between failure probability, test temperature, and fracture toughness in the Master Curve model.

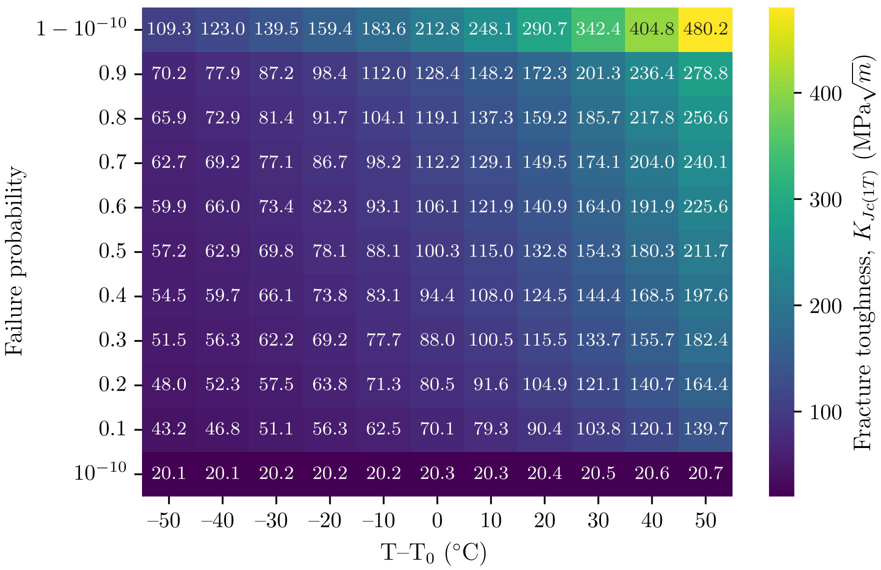
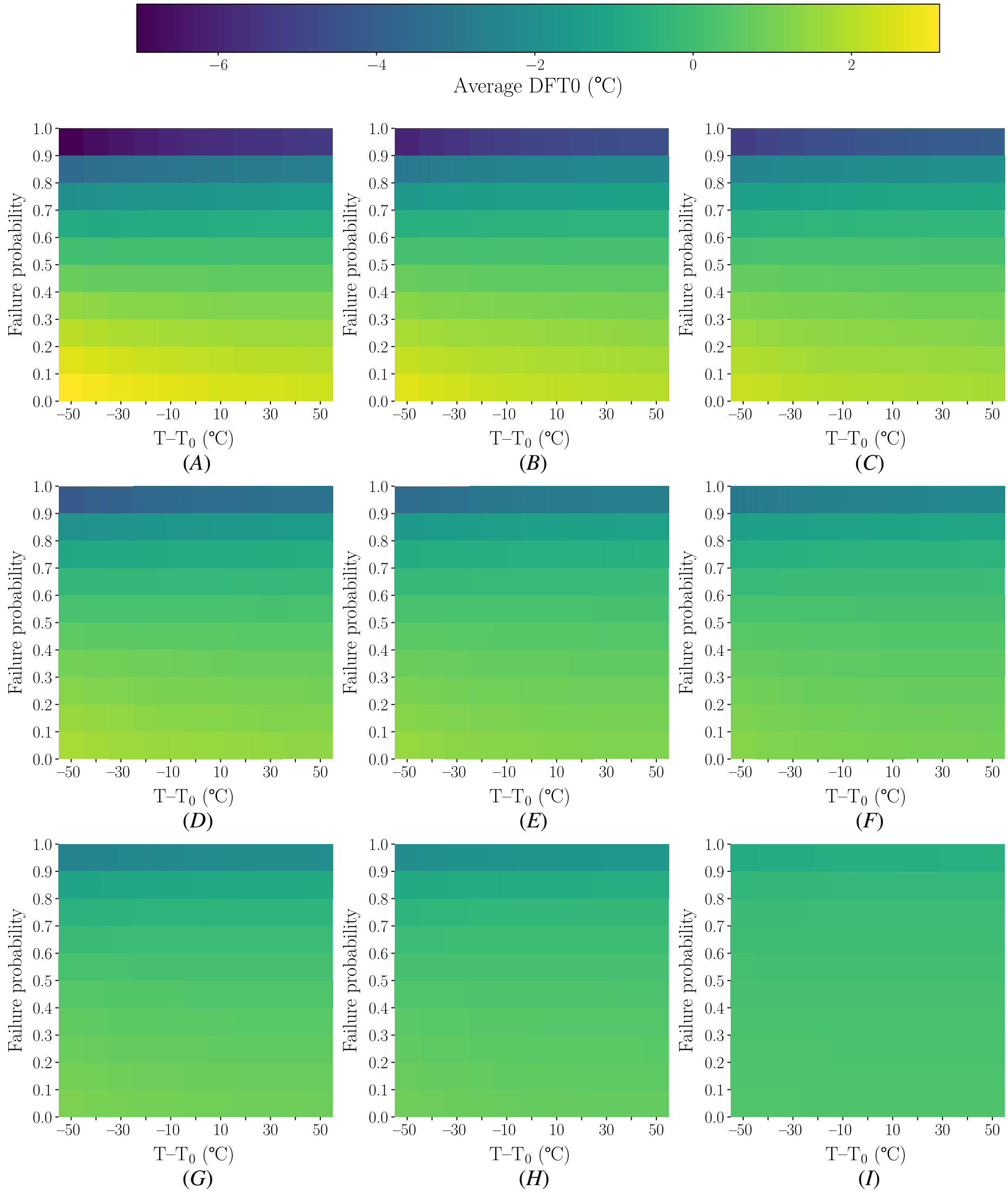


FIG. 6 Influence of individual specimens as a function of failure probability and test temperature for (*A*) $N=6$, (*B*) $N=7$, (*C*) $N=8$, (*D*) $N=10$, (*E*) $N=12$, (*F*) $N=14$, (*G*) $N=16$, (*H*) $N=20$, and (*I*) $N=50$.



with test temperature. From this position, the *DFT0* values exhibit asymmetric change in magnitude. In case of all data set sizes above $P_f \geq 0.6$, the average *DFT0* values were negative, with their magnitude increasing with P_f. Specimens with high fracture toughness value can skew T_0 estimates toward appreciably lower values; for $N=9$, data above the 90th percentile was characterized by the average *DFT0* of nearly −4°C. For the failure probability levels <0.5, the average *DFT0* values were increasing with decreasing P_f, and the average *DFT0* for data <10th percentile was equal to 1.7°C. Therefore, specimens with fracture toughness >90th percentile have significantly stronger influence on the T_0 estimate than those <10 percentile. This highlights that, although typically specimens that exhibited lower fracture toughness are perceived as a cause for concern, the Master Curve method is quite robust in dealing with them, pushing T_0 to a more conservative value.

On the other hand, the presence of fracture toughness specimens that performed exceptionally well in a data set can lead to a significantly less conservative T_0. The asymmetry is a result of the assumed distribution and temperature dependence as the range of fracture toughness values within equal probability bound bins is not the same, as illustrated in **figure 5**.

The test temperature also had a clear effect on the calculated average values of *DFT0*. It can be observed that as it increases the absolute *DFT0* values decrease across all failure probability ranges (however, at different rates for different P_f). The difference between the highest and lowest values of the *DFT0* average was also decreasing with the increase in $T - T_0$.

The range of the average *DFT0* values between the lowest and highest P_f bin at each test temperature decreases for larger datasets. This is expected because the more specimens there are in a data set, the lesser individual influence, on average, any one of them has on the result.

It should be noted that steel with high yield strength was simulated in this work. Yield strength affects the value of $K_{Jc(limit)}$, which provides an upper limit on constraint loss. It can be seen in **figure 2** that sampling a specimen that exceeds this limit is unlikely for the simulated steel and only possible at test temperatures near the maximum allowed by ASTM E1921. In datasets representing material with lower yield strength, the probability of having a censored datum in a data set is higher as the $K_{Jc(limit)}$ would be shifted down. As high fracture toughness specimens were found to be the most influential, having an upper limit on the fracture toughness value that can be used to calculate T_0 would lead to lower maximum *DFT0* values being reported if a steel characterized by lower yield strength was simulated.

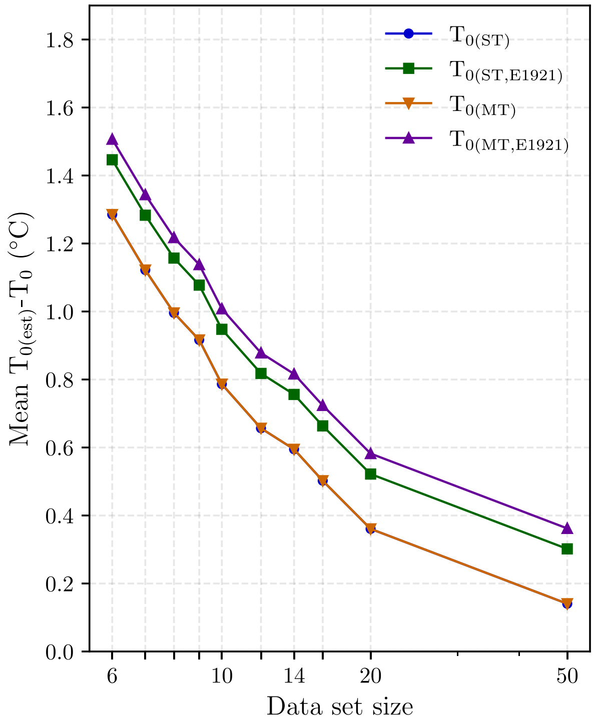
EFFECT OF THE ROUNDING IN ASTM E1921

The effects of the rounding operations applied to the results of equations (8), (11), and (12) as well as the estimation method chosen on the mean error of the reference temperature estimate are presented in **figure 7** and **figure 8**. For each of the four T_0 estimates, the mean error is defined here as the difference between the mean of the estimates and the true reference temperature of the simulated material and was calculated for every test temperature and data set size combination. The mean error observed for all cases was positive, and rounding of the parameter values in ASTM E1921 did introduce a certain level of bias. The same trends were observed for all test temperatures considered: there was no noticeable difference between the single-temperate and multitemperature methods when parameter values are not rounded; the bias introduced by rounding is always positive (i.e., leads to more conservative estimate), and it is slightly larger when the multitemperature method is applied. It is noted that the multitemperature equation is based on one rounded value, from equation (8), whereas the single-temperature method contains rounded parameters originating from equations (8), (11), and (12). The operation in equation (12) offsets some of the bias introduced by rounding of the equation (8) result, in the end leading to a more accurate T_0 estimate.

For a more detailed view, the difference between the T_0 estimates obtained by the ASTM E1921 equations and the estimates calculated with the parameter values that were not rounded to the nearest integer was calculated for each simulated data set. The maximum offset in each subset representing every data set size and test temperature combination, caused by rounding in the Standard for the single-temperature method, was plotted and shown in **figure 9** for each combination of data set size and test temperature. It can be observed that the offset increases with the test temperature, with the highest value reaching 0.167°C. For the multitemperature method, the offset caused by rounding was practically (to 8 decimal places) equal for all simulated datasets, at approximately 0.2215°C. In any case, rounding the parameters in the Standard leads to the T_0 estimates being more conservative, by roughly 0.2°C.

FIG. 7

Error of mean T_0 estimates obtained for datasets with simulated $T - T_0 =$ 0°C.



EFFECT OF THE TEST TEMPERATURE AND DATA SET SIZE ON T_0 ESTIMATION ERROR

The effect of the test temperature and data set size on the performance of the ASTM E1921 single-temperature reference temperature estimation procedure is presented in **figure 10**, where the root mean square error (*RMSE*) is plotted against test temperature for various data set sizes. The *RMSE* decreases as test temperature and sample size increase. The results in **figure 10** indicate that, for the data set sizes typically generated in the industry (up to ten specimens), using the single-temperature equations from ASTM E1921 produces reference temperature estimates that, on average, can differ from the actual value by 4.6°C to around 7°C, depending on the test temperature and exact number of specimens. To distinguish the contributions of accuracy and precision to the overall error, the relationship between the three is recalled in equation (25); bias and standard deviation (SD) of T_0 estimates were plotted in **figure 11**. Both metrics decrease with increasing test temperature and sample size, however, the test temperature dependence of the bias reduces with increasing sample size. It was negligible for $N = 50$, where for all test temperatures the average difference between the T_0 estimates and the known T_0 value was in the region of 0.3°C. It was also observed that all simulated estimates were, on average, positively biased (i.e., estimates exceeded the true value). The trends and values observed in the standard deviation of T_0 estimates are in accordance with what has been previously reported in the literature.[16,17]

$$RMSE = \sqrt{SD^2 + bias^2} \tag{25}$$

Conclusions

Almost three million fracture toughness datasets were simulated to perform the analysis described in this work. Design of the simulation program highlighted that the following information, in addition to the data set size,

FIG. 8 Error of mean T_0 estimates obtained for datasets with simulated $T - T_0$ values of (*A*) −50°C, (*B*) −40°C, (*C*) −30°C, (*D*) −20°C, (*E*) −10°C, (*F*) 10°C, (*G*) 20°C, (*H*) 30°C, (*I*) 40°C, and (*J*) 50°C.

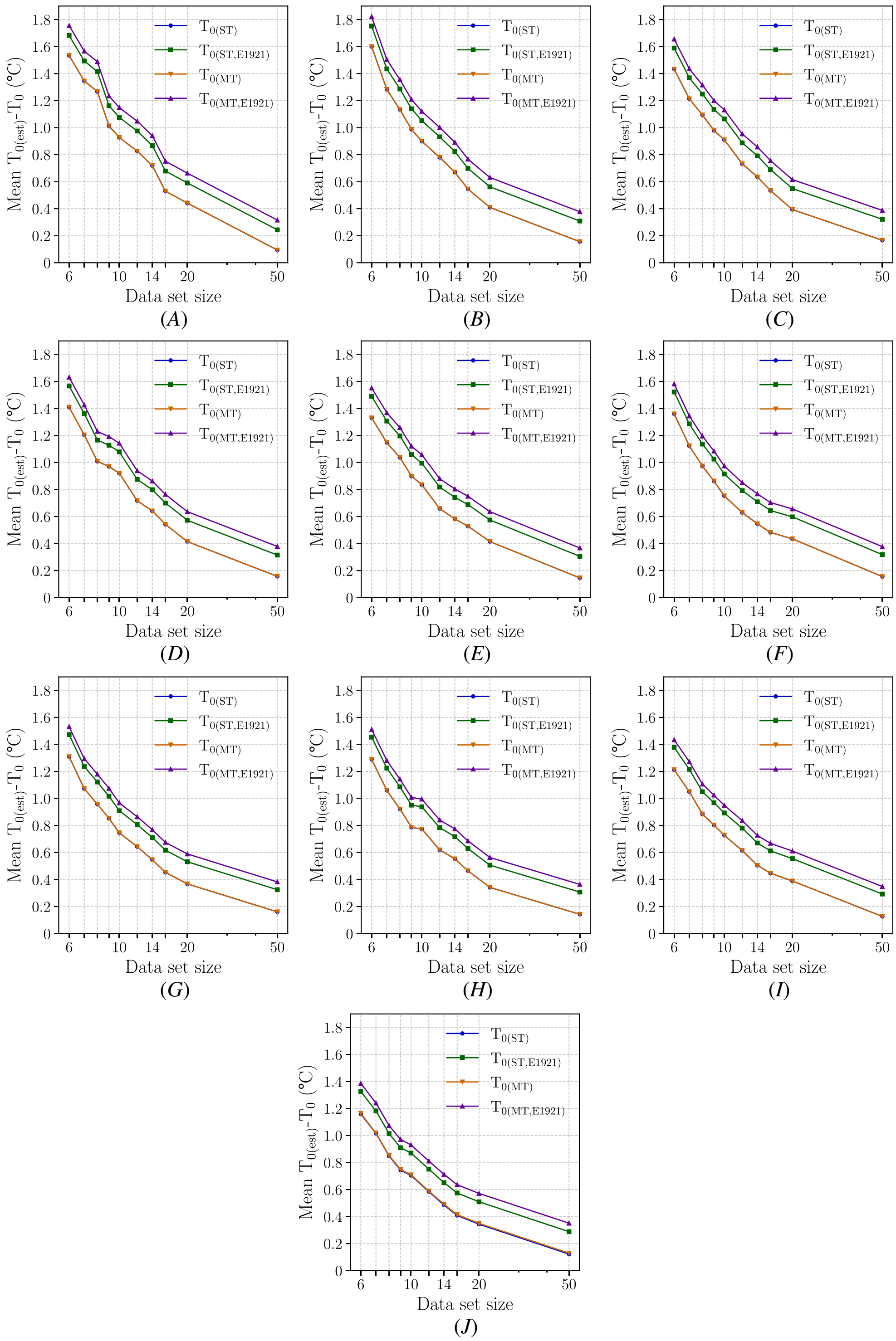


FIG. 9 Maximum difference between the two single-temperature T_0 estimates obtained for the simulated datasets.

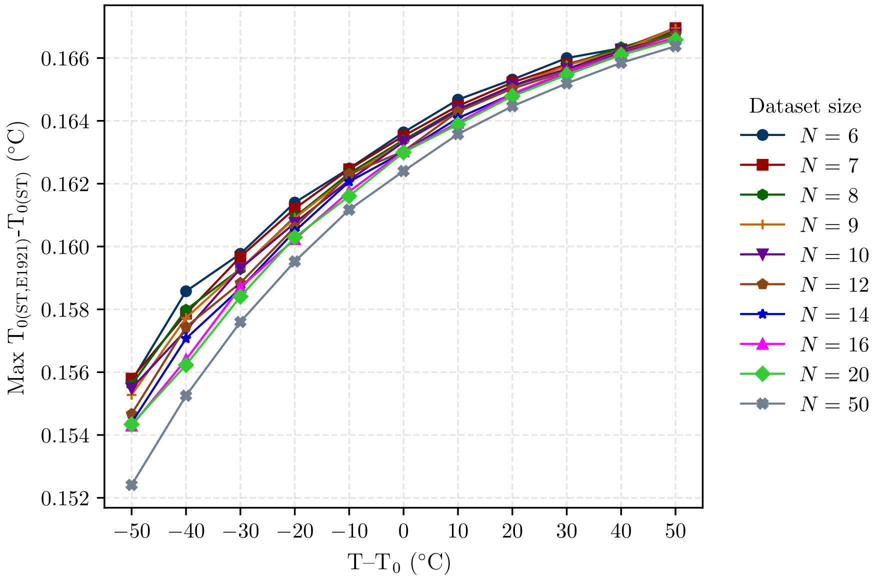
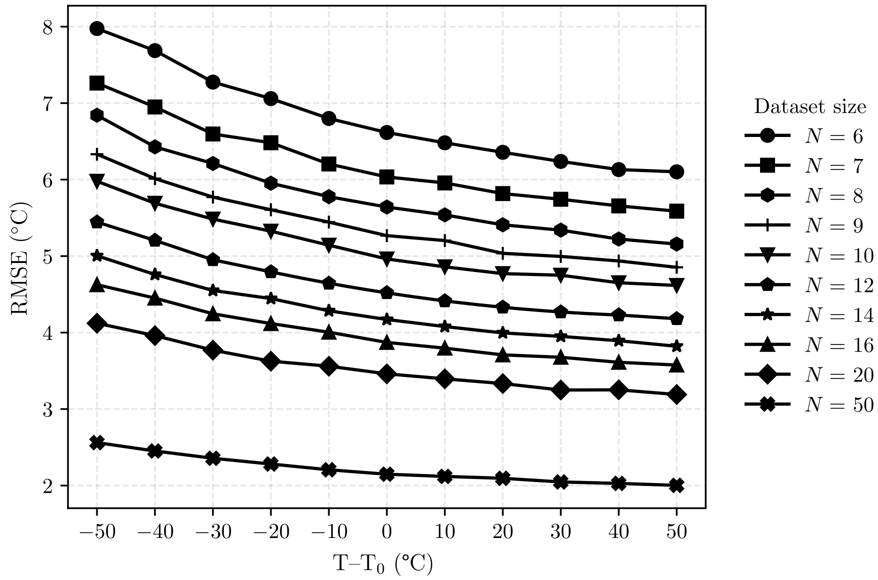
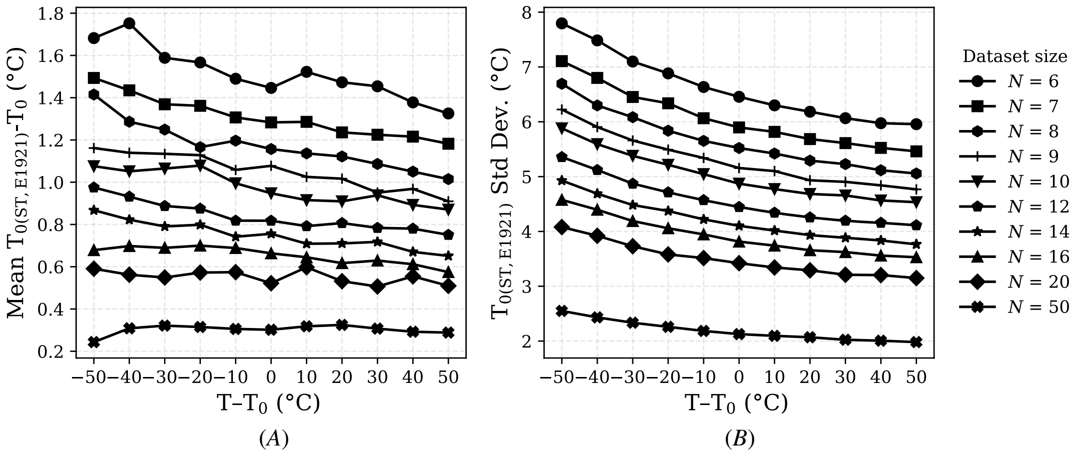


FIG. 10 Root mean square error on the T_0 estimates obtained by applying the ASTM E1921 single-temperature method.



specimen thickness, and test temperatures, must be provided in order to ensure the work describing Master Curve–obeying simulated data is reproducible: true reference temperature of the simulated material, how the yield strength, and elastic modulus values used to calculate $K_{Jc(limit)}$ of the simulated material were obtained, the remaining ligament length, b_0 characterizing the simulated specimens, and precise range of failure probability sample space. The treatment of the simulated datasets that fail some or all the ASTM E1921 validity criteria should also be stated.

FIG. 11 (*A*) Bias and standard deviation (*B*) of the T_0 estimates obtained by applying the ASTM E1921 single-temperature T_0 estimation method.



A measure of specimen influence, *DFT0*, based on the leave-one out principle was introduced, and its application to the simulated data provided new insights regarding the magnitude of effect that individual specimens can have on a T_0 estimate. *DFT0* is an easily understandable metric that can be used to quantify the influence of specimens in any fracture toughness datasets. It was found that high fracture toughness specimens have a larger impact on T_0 values than low fracture toughness specimens and lead to less conservative values. If exceptionally high fracture toughness specimens are present in a data set, this effect can be notable, especially in small datasets, and exceed 4°C (ASTM E1921 recommended value of σ_{exp}). In relatively large datasets of 20 fracture toughness specimens, the average *DFT0* values above 90 % probability bound were between −1.61 and −2.07°C (depending on the test temperature). Specimens tested at lower temperatures are generally more influential than those tested at higher temperatures, and this effect is more pronounced as the failure probability values move from the median toward the extremes. In summary, for the considered case of homogeneous datasets with specimens tested at the same temperature, the influence of any single specimen on the estimate of T_0 would not exceed the expected standard deviation of T_0. The *DFT0* values observed were notably lower than the differences in T_0 (exceeding 10°C), which can be observed because of issues such as specimen geometry type or material inhomogeneity.[5,19,25,30]

Regardless of whether multi or single-temperature equations are solved to estimate reference temperature, the T_0 estimates are, on average, overestimated. Having rounded parameters in the ASTM E1921 equations introduces an additional positive bias of around 0.2°C to the T_0 estimates. Increasing test temperature and data set size leads to lower root mean square error of T_0 estimates. This effect is also generally true for bias and variance examined separately, with the exception of bias on the estimates made for the largest data set tested ($N = 50$) not exhibiting temperature dependence, having plateaued at around 0.3°C. These effects are, however, negligible, particularly when compared to other sources of bias and variance in T_0 estimates.

Future research will focus on applying *DFT0* to multitemperature datasets and investigation of whether there is a relationship between presence of highly influential data points in a data set and T_0 accuracy and the result of the homogeneity screening procedure. *DFT0* will also be applied to inhomogeneous datasets to evaluate its potential as a complementary tool to the ASTM E1921 homogeneity screening procedure.

ACKNOWLEDGMENTS

This research was funded by The UK Engineering and Physical Sciences Research Council, grant number EP/S022635/1.

References

1. K. Wallin, "The Scatter in K_{IC}-Results," *Engineering Fracture Mechanics* 19, no. 6 (1984): 1085–1093, https://doi.org/10.1016/0013-7944(84)90153-X
2. K. Wallin, "The Size Effect in K_{IC} Results," *Engineering Fracture Mechanics* 22, no. 1 (1985): 149–163, https://doi.org/10.1016/0013-7944(85)90167-5
3. K. Wallin, "Fracture Toughness Transition Curve Shape for Ferritic Structural Steels," in *Fracture of Engineering Materials and Structures*, edited by S. H. Teoh and K. H. Lee (Dordrecht, the Netherlands: Springer, 1991), 83–88, https://doi.org/10.1007/978-94-011-3650-1_10
4. J. G. Merkle, K. Wallin, and D. E. McCabe, *Technical Basis for an ASTM Standard on Determining the Reference Temperature, T_0, for Ferritic Steels in the Transition Range, NUREG/CR-5504* (Oak Ridge, TN: Oak Ridge National Laboratory, 1998).
5. *Standard Test Method for Determination of Reference Temperature, T0, for Ferritic Steels in the Transition Range* (Superseded), ASTM E1921-24 (West Conshohocken, PA: ASTM International, approved May 1, 2024), https://doi.org/10.1520/E1921-24
6. M. Erickson, M. Kirk, and G. Stevens, *Technical Basis for ASME Code Case N-830, Revision 1 (MRP-418): Direct Use of Master Curve Fracture Toughness Curve for Pressure-Retaining Materials of Class 1 Vessels, Section XI* (New York, NY: ASME, 2014).
7. *Guide to Methods for Assessing the Acceptability of Flaws in Metallic Structures*, BS7910:2019 (London: British Standards Institution, 2019), https://doi.org/10.3403/30369478
8. International Atomic Energy Agency, *Guidelines for Application of the Master Curve Approach to Reactor Pressure Vessel Integrity in Nuclear Power Plants, Technical Reports Series No. 429* (Vienna, Austria: International Atomic Energy Agency, 2005).
9. *Fitness-for-Service*, API 579-1/ASME FFS-1 (Washington, DC: American Petroleum Institute, 2021).
10. M. Yamamoto, K. Onizawa, K. Yoshimoto, T. Ogawa, Y. Mabuchi, M. Valo, M. Lambrecht, H.-W. Viehrig, N. Miura, and N. Soneda, "International Round Robin Test on Master Curve Reference Temperature Evaluation Utilizing Miniature C (T) Specimen," in *Small Specimen Test Techniques: Sixth Volume*, ed. M. Sokolov and E. Lucon (West Conshohocken, PA: ASTM International, 2014), 53–69, https://doi.org/10.1520/STP157620140020
11. D. Andres, M. Serrano, R. Hernandez, Y. Wang and M. Richardson, "Determination of Eurofer97 Fracture Toughness by Testing Small C(T) Specimens," in *Proceedings of the ASME 2021 Pressure Vessels and Piping Virtual Conference, Volume 1: Codes and Standards* (New York: ASME, 2021), PVP2021-62040, V001T01A025, https://doi.org/10.1115/PVP2021-62040
12. M. Sánchez, S. Cicero, M. Kirk, E. Altstadt, W. Server, and M. Yamamoto, "Using Mini-CT Specimens for the Fracture Characterization of Ferritic Steels within the Ductile to Brittle Transition Range: A Review," *Metals* 13, no. 1 (January 2023): 176, https://doi.org/10.3390/met13010176
13. M. Sánchez, S. Cicero, B. Arroyo, and A. Cimentada, "On the Use of Mini-CT Specimens to Define the Master Curve of Unirradiated Reactor Pressure Vessel Steels with Relatively High Reference Temperatures," *Theoretical and Applied Fracture Mechanics* 124 (2023): 103736, https://doi.org/10.1016/j.tafmec.2022.103736
14. F. Naziris, R. Hernandez Pascual, T. Metzler, E. Gaganidze, I. Uytdenhouwen, and M. Kolluri, "Master Curve Evaluation Using Miniature C(T) Specimens as Part of a Round Robin Program within the FRACTESUS Project," in *Proceedings of the ASME 2023 Pressure Vessels and Piping Conference: Volume 1, Codes & Standards* (New York: ASME, 2023), PVP2023–107254, V001T01A045, https://doi.org/10.1115/PVP2023-107254
15. M. Sánchez, S. Cicero, and B. Arroyo, "Mini-C(T) Specimens for Master Curve Analysis of Structural Steels Operating within Their Ductile-to-Brittle Transition Region," *Engineering Fracture Mechanics* 298 (2024): 109917, https://doi.org/10.1016/j.engfracmech.2024.109917
16. C. A. J. Miranda and J. D. Landes, "Defining the Confidence Level in the Reference Temperature Determination for a Finite Set of Toughness," *Engineering Fracture Mechanics* 68, no. 8 (May 2001): 1061–1076, https://doi.org/10.1016/S0013-7944(01)00004-2
17. T.-L. S. Sham and D. R. Eno, "Estimation Uncertainty in the Determination of the Master Curve Reference Temperature," in *Proceedings of the ASME 2010 Pressure Vessels and Piping Division/K-PVP Conference: Volume 6, Parts A and B* (Bellevue, WA: ASME, 2010), 107–117, https://doi.org/10.1115/PVP2010-25332
18. T. Yount, W. Collins, S. Lindqvist, and K. Wallin, "Statistical Analysis of Master Curve Sampling, Screening, and Weighting Methodologies," *Journal of Testing and Evaluation* 52, no. 2 (March 2024): 1163–1188, https://doi.org/10.1520/JTE20230372
19. E. Lucon, M. Scibetta, and E. Van Walle, "Assessment of the Master Curve Approach on Three Reactor Pressure Vessel Steels," *International Journal of Fracture* 119, no. 2 (January 2003): 161–178, https://doi.org/10.1023/A:1023910610553
20. A. Quintin, T. Petit, R. Chocat, C. Mattrand, and J.-M. Bourinet, "Uncertainty Quantification of the Reference Temperature T_0 of 16MND5 Steel from Experimental and Numerical Fracture Toughness Tests," *Engineering Fracture Mechanics* 315 (2025): 110755, https://doi.org/10.1016/j.engfracmech.2024.110755

21. K. Wallin, "Recommendations for the Application of Fracture Toughness Data for Structural Integrity Assessments," in *Proceedings of the Joint IAEA/CSNI Specialists' Meeting on Fracture Mechanics Verification by Large-Scale Testing, NUREG/CP-0131* (Washington, DC: US Nuclear Regulatory Commission, 1993), 465–494.
22. K. Wallin, "Statistical Modelling of Fracture in the Ductile-to-Drittle Transition Region," in *Defect Assessment in Components–Fundamentals and Applications* (London: Elsevier, 1991), 415–445.
23. D. E. McCabe, J. G. Merkle, and K. Wallin, *An Introduction to the Development and Use of the Master Curve Method* (West Conshohocken, PA: ASTM International, 2005), https://doi.org/10.1520/MNL52-EB
24. D. J. Cogswell, "Statistical Modelling of the Transition Toughness Properties of Low Alloy Pressure Vessel Steels" (PhD diss., University of Birmingham, 2010).
25. K. Wallin, "Master Curve Analysis of the "Euro" Fracture Toughness Dataset," *Engineering Fracture Mechanics* 69, no. 4 (March 2002): 451–481, https://doi.org/10.1016/S0013-7944(01)00071-6
26. K. K. Yoon, J. B. Hall, W. A. Van Der Sluys, M. Higuchi, and K. Iida, "Japanese Fracture Toughness Data Analysis Using Master Curve Method," in *Proceedings of the ASME 2001 Pressure Vessels and Piping Conference* (New York: ASME, 2001), 167–174.
27. M. Kirk, *The Technical Basis for Application of the Master Curve to the Assessment of Nuclear Reactor Pressure Vessel Integrity, ADAMS ML093540004* (Washington, DC: US Nuclear Regulatory Commission, 2009).
28. M. Holzmann, J. Man, L. Válka, and B. Vlach, "*R*-Curves and Fracture Toughness Transition Behaviour at Static, Rapid and Impact Loading of Cr-Ni-Mo-V Reactor Pressure Vessel Steel," *International Journal of Pressure Vessels and Piping* 62, no. 1 (1995): 39–47, https://doi.org/10.1016/0308-0161(95)93827-R
29. D. A. Belsley, E. Kuh, and R. E. Welsch, "Detecting Influential Observations and Outliers," in *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity* (Hoboken, NJ: John Wiley & Sons, 1980), 6–84, https://doi.org/10.1002/0471725153
30. E. Lucon, "Assessment of Macroscopically Inhomogeneous Fracture Toughness Data Sets Using the Simplified and Multimodal Master Curve Methods," *Theoretical and Applied Fracture Mechanics* 125 (2023): 103861, https://doi.org/10.1016/j.tafmec.2023.103861