NIST Technical Note 2158

NIST Software Package for Obtaining Charpy Transition Curves

Enrico Lucon
Jolene Splett
Amanda Koepke
David Newton

This publication is available free of charge from:
https://doi.org/10.6028/NIST.TN.2158

National Institute of
Standards and Technology
U.S. Department of Commerce

NIST Technical Note 2158

NIST Software Package for Obtaining Charpy Transition Curves

Enrico Lucon
Applied Chemicals and Materials Division
Material Measurement Laboratory

Jolene Splett
Amanda Koepke
David Newton
Statistical Engineering Division
Information Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.TN.2158

May 2021

U.S. Department of Commerce
Gina Raimondo, Secretary

National Institute of Standards and Technology
James Olthoff, NIST Director and Undersecretary of Commerce for Standards and Technology

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

National Institute of Standards and Technology Technical Note 2158
Natl. Inst. Stand. Technol. Tech. Note 2158, 41 pages (May 2021)
CODEN: NTNOEF

This publication is available free of charge from:
https://doi.org/10.6028/NIST.TN.2158

Abstract

NIST has developed a software package that allows users to fit models to test results obtained from Charpy or toughness tests as a function of test temperature, thereby obtaining so-called transition curves. Non-linear fitting is available for five regression models, two symmetric and three asymmetric, selected from the literature. The NIST software package includes: an Excel spreadsheet to establish initial values for model parameters, allowing users to obtain meaningful results for the asymmetric models, a Shiny App (web-based interface for R code) that provides an advanced analysis, including confidence intervals and uncertainties for the estimated parameters, and a macro-enabled Excel spreadsheet for a basic-level analysis (regressions curves, characteristic temperatures, and model selection statistics). Example data sets are also included in the NIST software package.

The NIST software is tested on three Charpy absorbed energy datasets. Two datasets correspond to actual impact test results performed on various pipeline steels, while the third is a simulated dataset generated from predetermined model parameters. These datasets allow a thorough demonstration of the NIST software. The software is freely available to the scientific community for research purposes.

The full software package is available for free download from the NIST Charpy Machine Verification Program web page (https://www.nist.gov/programs-projects/charpy-machine-verification-program).

Key words

Charpy testing; confidence bounds; non-linear fitting; parameter uncertainties; toughness tests; transition curves.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

Table of Contents

1. Introduction 1
2. Background 2
3. Establishment of Charpy Transition Curves 4
3.1. Mathematical Models Selected to Fit Charpy Test Results 5
3.1.1. Hyperbolic Tangent (HT – Symmetric) 5
3.1.2. Asymmetric Hyperbolic Tangent (AHT – Asymmetric) 6
3.1.3. Burr Model (BUR – Asymmetric) 7
3.1.4. Arctangent Model (ACT – Symmetric) 8
3.1.5. Asymmetric Kohout Model (KHT – Asymmetrical) 8
3.1.6. Other Regression Models Not Selected 9
4. NIST Software Package 10
4.1. Excel Spreadsheet for Initial Values 10
4.1.1. Instruction for Use 11
4.2 The Shiny App 14
4.3 Excel Spreadsheet for Basic Analysis 23
4.3.1. Sheet "*Input Data*" (Fig. 19) 23
4.3.2 Sheets "*HT*" (Fig. 20), "*AHT*", "*BUR*", "*ACT*", and "*KHT*" 24
4.3.3 Sheet "*Summary of Results*" 26
4.3.4 Sheet "*Characteristics Temperatures*" 26
4.3.5 Sheet "*Comparison Fitting Curves*" 26
5 Shiny App Examples 26
5.1 Dataset #1: Simulated 26
5.2 Dataset #2: "Variable" Shelves 28
5.3 Dataset #3: Outliers 29
6 Comparison of Output from Excel Spreadsheet and Shiny App 32
7 Conclusions 34
References 35
Appendix A: Computational and Statistical Methods 37
A.1 Convergence 37
A.2 Model Selection Statistics 37
A.3 Temperature Variability 38
A.4 Finding Standard Errors: Parametric Bootstrap Analysis 38
A.5 Uncertainty of Fixed Shelves 39
A.6 Confidence Intervals and Confidence Bounds 39
A.7 Diagnostic Plots 39

A.8 DBTT Calculation 40
A.9 Calculation of Characteristic Temperatures 40

List of Tables

Table 1. Values of LS, US, and initial values of model parameters used in the analyses performed in this document. 14
Table 2 - Selected numerical results for the HT model fit to the dataset #2 KV response for fixed shelf and variable shelf analyses. Standard errors of parameter estimates are shown in parentheses. 28
Table 3 - Comparison between Excel spreadsheet and Shiny App. 33
Table 4 - Features available in the Excel spreadsheet and in the Shiny App. 34

List of Figures

Fig. 1. Absorbed energy transition curve for a low-strength steel. 2
Fig. 2. Schematic transition behavior for various types of metals [5]. 3
Fig. 3. Example of AHT transition curve. 6
Fig. 4. Example of BURR transition curve. 7
Fig. 5. Example of ACT transition curve. 8
Fig. 6. Example of KHT transition curve. 9
Fig. 7. Left: Absorbed energy for simulated dataset #1. Right: Absorbed energy for experimental dataset #2. 10
Fig. 8. Worksheet for the AHT model (upper part). 11
Fig. 9. Worksheet for the AHT model (lower part). 12
Fig. 10. Effect of k parameter on the BUR regression curve (m = 0.1). 13
Fig. 11. Effect of m parameter on the BUR regression curve (k = 1). 13
Fig. 12. Effect of p parameter on the KHT regression curve. 13
Fig. 13. Screenshot of the top right portion of the Shiny App. 15
Fig. 14. Shiny App output: “Regression Results” tab (top section). 18
Fig. 15. Shiny App output: “Regression Results” tab (middle section). 19
Fig. 16. Shiny App output: “Regression Results” tab (bottom section). 20
Fig. 17. Shiny App output: “Diagnostic Plots” tab for the HT model. 22
Fig. 18. Transition curve for HT model with 95 % confidence bounds for Dataset #2 (absorbed energy). 23
Fig. 19. Sheet “*Input Data*”. 25
Fig. 20. Sheet “*HT*”. 25
Fig. 21. Simulated data and estimated transition curves for dataset #1 (KV). Model selection statistics are also shown. 27
Fig. 22. Transition curve for HT model with confidence bounds for dataset #2 (absorbed energy). 28
Fig. 23. Experimental data and estimated transition curves for dataset #3 (KV). 29
Fig. 24. Model selection statistics and parameter estimates for models fit to dataset #3. 30
Fig. 25. DBTT table and additional characteristic temperatures for dataset #3. 31
Fig. 26. Diagnostic plots for ACT model fit to dataset #3. 32

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

1. Introduction

This document describes software developed by the National Institute of Standards and Technology (NIST) to determine temperature transition curves for metals and steels. The temperature transition curves provide crucial information about the ductile-to-brittle transition of metals and steels used to build bridges and buildings and to monitor the degradation of metals and steels used in nuclear reactors.

The objective of our effort is not just to provide software for fitting models to Charpy or toughness test results (there are many available packages that can do that, and calculations are relatively easy to implement in a spreadsheet-based software such as Microsoft Excel[1]), but to provide a tool that can quantify uncertainties associated with Charpy parameters such as DBTT or characteristic temperatures and establish confidence bounds for the calculated regression curves. This is of particular interest whenever Charpy results are used in the framework of safety-related applications or structural assessments.

The NIST software package provides tools to fit five different models to temperature transition data, examine diagnostic plots, compare several fitted curves using model selection statistics, estimate uncertainties and confidence intervals for all parameters of interest, and plot confidence bounds for fitted curves. The NIST software package is comprised of three parts. The first part is a Microsoft Excel [1] spreadsheet utility that provides reasonable initial values of parameters. Utilizing appropriate initial values is of paramount importance in obtaining meaningful results for specific combinations of datasets and regression models. The second part, the Shiny App [2], is a user-friendly, web-based application (based on R software[3]) that performs all the statistical and graphical procedures needed to analyze temperature transition data. The third part is a macro-enabled Excel spreadsheet that offers a basic tool for determining regression curves for selected models, as well as for ranking the models in terms of fitting capability. The basic analysis tool does not produce uncertainties or confidence intervals. The most common implementation of the NIST software package is to use either the Shiny App for a complete analysis or the Excel spreadsheet for basic analysis, with initial values from the first Excel spreadsheet if needed. Sample data sets are also provided as part of the software package.

Section 2 provides background information regarding ductile-to-brittle temperature transition curves and the parameters of interest. Details regarding the response variables and the mathematical models available in the NIST software package are described in section 3. The NIST software package and the statistical techniques employed are described in section 4. Additional data analyses are described in section 5. A comparison of output from the Shiny App and the Excel spreadsheet for basic analysis is included in section 6. Section 7 provides a summary of the work completed. Computational considerations and statistical methods are provided in Appendix A.

[1] Certain commercial equipment, instruments, software packages, or materials are identified in this document in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

2. Background

The ductile-to-brittle transition is a phenomenon commonly observed in metals and steels, which corresponds to the change in the fracture behavior of metals from ductile (stable) fracture at high temperatures to brittle (unstable) fracture at low temperatures. The change in fracture mode generally occurs over a range of temperatures, centered on a specific temperature that is denominated ductile-to-brittle transition temperature, or DBTT. An example transition curve for the energy absorbed from Charpy tests on a low-strength ferritic steel (KV, in Joules) is illustrated in **Fig. 1**. Note that the temperature range where the fracture is fully ductile is generally known as *upper shelf* (US) and the temperature range where the fracture is fully brittle is typically indicated as *lower shelf* (LS). In both shelves, the transition curve is conventionally horizontal (*i.e.*, the toughness of the material is considered approximately constant).

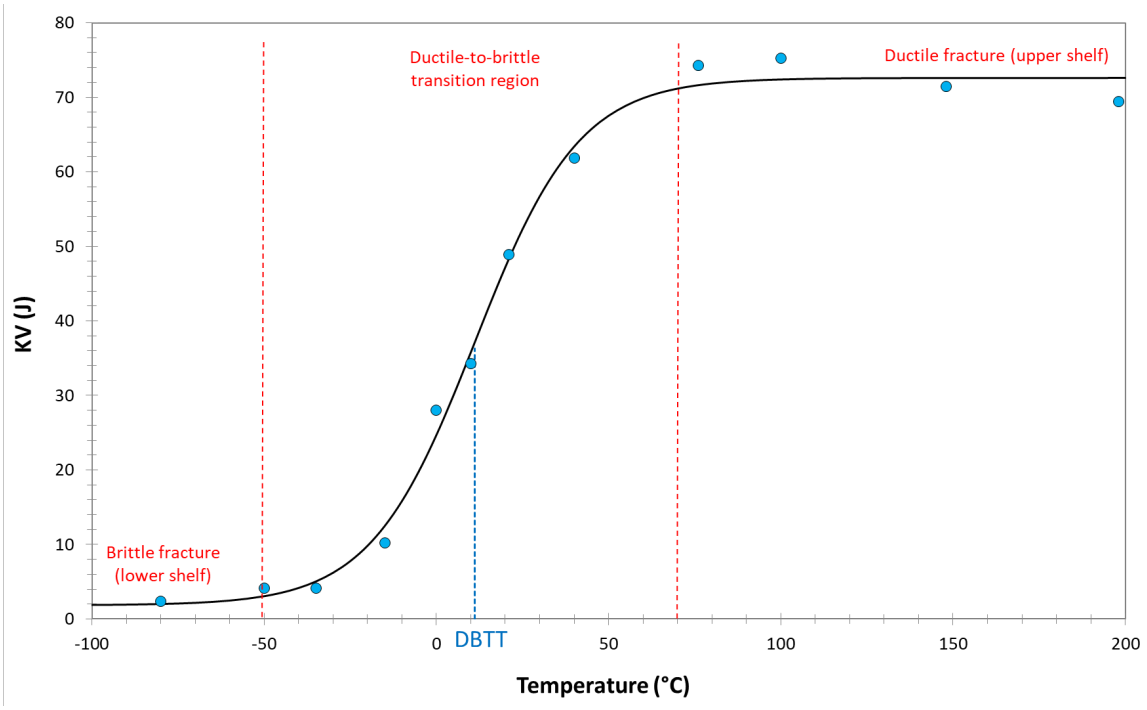


Fig. 1. Absorbed energy transition curve for a low-strength steel.

Sigmoidal regression curves such as the one depicted in **Fig. 1** are used to model the temperature dependence of Charpy test results, such as:

- absorbed energy, KV (J);
- lateral expansion, LE (mm);
- shear fracture appearance, SFA (%).

The same type of transition curve can also be used for illustrating the temperature dependence of other toughness-related parameters, such as stress-intensity factor K (MPa$\sqrt{m}$) or J-integral (kN/m^2). Although temperature transition curves are used to describe several possible Charpy test results, only KV will be used in this document to illustrate the use of the NIST software package.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

The phenomenon of ductile-to-brittle transition occurs because the development of the plastic zone in some classes of metals is a temperature-dependent process [4]. At upper-shelf temperatures, there is sufficient thermal energy in the crystal structure to facilitate the movement of dislocations under an externally applied stress. This allows the plastic zone to develop at the crack tip, thus facilitating crack formation and propagation by ductile (stable) fracture. However, as the temperature decreases, the thermal energy available to assist dislocation movement drops, making it more difficult to develop a significant plastic zone at the tip of the notch or pre-existing crack. Below the DBTT, dislocation mobility and plastic zone size rapidly decrease, resulting in a large loss of fracture toughness. Once dislocation slip is virtually impossible, the metal fractures by brittle (unstable) crack propagation.

As shown in **Fig. 2** [5], the transition behavior of metals falls into three categories determined by their crystal structure:

- Metals with FCC (face-centered cubic) crystal structure do not undergo a ductile-to-brittle transition and retain considerable ductility at low temperatures; the large number of slip systems in their crystal structure allows dislocation movement to occur even at low temperatures.
- Most hexagonal close packed (HCP) metals, including magnesium and α-titanium alloys, also do not exhibit a transition behavior.
- On the other hand, metals with a body-centered-cubic (BCC) crystal structure typically display ductile-to-brittle transition properties, as they do not exhibit close-packed planes that allow for easy dislocation migration, Therefore, dislocation movement in these materials require a thermal activation in order to slip, and more drastic mechanisms, such as breaking of bonds, are triggered at low temperatures to accommodate the applied stress [5].

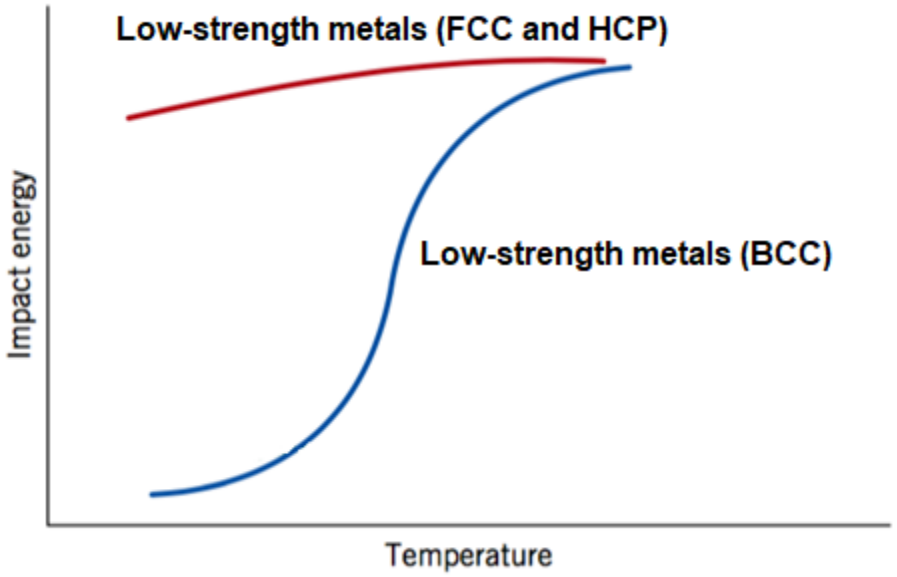


Fig. 2. Schematic transition behavior for various types of metals [5].

3. Establishment of Charpy Transition Curves

The Charpy impact test, developed over 120 years ago [6,7], is the simplest and most conventional test used to characterize the transitional behavior of BCC metals and steels. In its simplest form, this test measures the amount of energy expended to break a standardized bar-shaped specimen, notched on one side, upon impact by a swinging hammer. The absorbed energy, which must be corrected for the losses due to friction and windage, is calculated from the difference in the potential energies of the hammer (pendulum) before and after fracturing the specimen placed in the path of its swing. The most commonly used Charpy specimen is a bar with nominal length of 55 mm and a square cross section of 10 mm × 10 mm, with a V-shaped notch. The symbol of absorbed energy is KV, and its unit is Joules (J).

Besides absorbed energy, two additional post-test measurements are typically performed on Charpy specimens and reported.

- Lateral Expansion (LE, in mm): the increase in specimen thickness caused by plastic deformation induced by fracture. LE is measured on the broken specimen halves (separately or together) as the cumulative height of the shear lips caused by plastic deformation.

- Shear Fracture Appearance (SFA, in %): the percentage of ductile (shear) fracture surface measured on a tested specimen using optical or digital methods. SFA ranges from 0 % (fully brittle) in the lower shelf to 100 % (fully ductile) in the upper shelf.

All three measurements (KV, LE, and SFA) exhibit a transitional behavior for BCC metals similar to the one shown in **Fig. 1**. Regression (transition) curves are obtained by fitting mathematical models to results obtained as a function of test temperature, generally with a least-squares approach. This consists of minimizing the sum of squared residuals, with the residual defined as the difference between observed and fitted value. The method of least squares is a standard approach in regression analysis to approximate the solution of overdetermined systems (sets of equations in which there are more equations than unknowns).

Once the Charpy transition curve is mathematically established, the parameters that are commonly extracted are the following.

- DBTT (°C): the ductile-to-brittle transition temperature corresponding to the abscissa of the point that is equidistant from the lower and upper shelf, *i.e.*, DBTT = T[(LS+US)/2]. Conventionally, one can say that ductile behavior prevails at temperatures above the DBTT, while brittle behavior prevails below the DBTT.
- US (J or mm): upper-shelf value that is related to the toughness achieved by the material under fully ductile conditions. US can be expressed in terms of absorbed energy (in J) or lateral expansion (in mm). As for SFA, the upper-shelf value is, by definition, equal to 100 %.
- Additional characteristic temperatures corresponding to predefined values of the dependent variable. For example, within the nuclear energy community, characteristic temperatures corresponding to specific values of absorbed energy (27 or 28 J, 41 J, 68 J) or lateral expansion (0.89 mm) [8] are used. Moreover, material specifications might dictate maximum acceptable values for temperatures corresponding to predetermined values of KV, LE, or SFA.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

In the NIST software package, LS and US values can be either "fixed" or "variable" in the analysis. A "fixed" shelf analysis is performed when the user specifies the values of US and/or LS based on experimental data; shelf values are not obtained via nonlinear regression. A "variable" shelf analysis is performed when the values of US and/or LS are obtained by nonlinear regression.

3.1. Mathematical Models Selected to Fit Charpy Test Results

Undoubtedly, the most popular regression model used for fitting Charpy or toughness data is the Hyperbolic Tangent (HT) model, originally proposed by Oldfield in the 1970s [9,10]. The HT equation is an example of a symmetric model, as the regression curve is perfectly symmetric with respect to the flexural point that corresponds to X = DBTT and Y = (LS+US)/2.

However, some Charpy datasets exhibit a different shape in the two "knee" regions that connect the DBTT region with the lower and upper shelf (see **Fig. 1**). These two regions are commonly referred to as "lower transition" (T < DBTT) and "upper transition" (T > DBTT). In order to accommodate different curve shapes in these two regions, an additional coefficient is introduced in some regression models, and these models are therefore labeled "asymmetric".

Starting in 2020, NIST staff from the Advanced Chemicals and Materials and Statistical Engineering Divisions reviewed available/published regression models for Charpy and/or toughness data, including symmetric and asymmetric models. Based on this literature review, five selected models (two symmetric and three asymmetric) are implemented in the NIST software package that is freely available to the scientific community. The five selected models are described in sections 3.1.1 – 3.1.5.

3.1.1. Hyperbolic Tangent (HT – Symmetric)

The original model proposed by Oldfield [9,10] has the form:

$$Y = A + B \cdot tanh\left(\frac{T-DBTT}{C}\right) \quad , \tag{1}$$

where Y is the dependent variable and T is temperature (independent variable).

In its most general form, the model has four parameters (A, B, C, $DBTT$) that are determined by least-squares fitting. The A and B parameters, however, can be rewritten in terms of lower shelf (LS) and upper shelf (US) values as:

$$Y = \frac{\mathrm{LS+US}}{2} + \frac{\mathrm{US-LS}}{2} \cdot tanh\left(\frac{T-\mathrm{DBTT}}{\mathrm{C}}\right) \quad . \tag{2}$$

Therefore, if both LS and US are "fixed" (as we recommend in this report), only two parameters, C and $DBTT$, are left to estimate, which can be helpful in case of sparse datasets (8 data points or less). An example of HT transition curve is shown in **Fig. 1**. The two remaining parameters have the following physical interpretation:

- C (°C) corresponds to the half-width of the transition region, and
- $DBTT$ (°C) is the Ductile-to-Brittle Transition Temperature defined above. It corresponds to the abscissa of the point Y = (LS+US)/2.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

3.1.2. Asymmetric Hyperbolic Tangent (AHT – Asymmetric)

The model [11] is derived from the HT model and includes the parameter, D, whose value determines how much the shape of the regression curve differs between lower and upper transition regions. The AHT model has the form:

$$Y = A + B \cdot tanh\left(\frac{T-DBTT}{C+D \cdot T}\right) \quad , \tag{3}$$

or alternatively, similar to Eq.(2) above:

$$Y = \frac{\text{LS+US}}{2} + \frac{\text{US−LS}}{2} \cdot tanh\left(\frac{T-\text{DBTT}}{\text{C+D}\cdot T}\right) \quad . \tag{4}$$

The additional parameter D quantifies the asymmetry of the regression curve. If $D = 0$, the curve becomes symmetrical and coincides with HT; if $D < 0$, the curvature in the lower transition region is larger than in the upper transition region; the opposite occurs when $D > 0$.

The curve is not defined when $T = -C/D$, which may occur in the upper shelf when $D < 0$, or in the lower shelf if $D > 0$. If $D < 0$, Y = LS for $T > -C/D$; if $D > 0$, Y = US for $T < -C/D$.

An example of AHT transition curve (with $D < 0$) is given in **Fig. 3**.

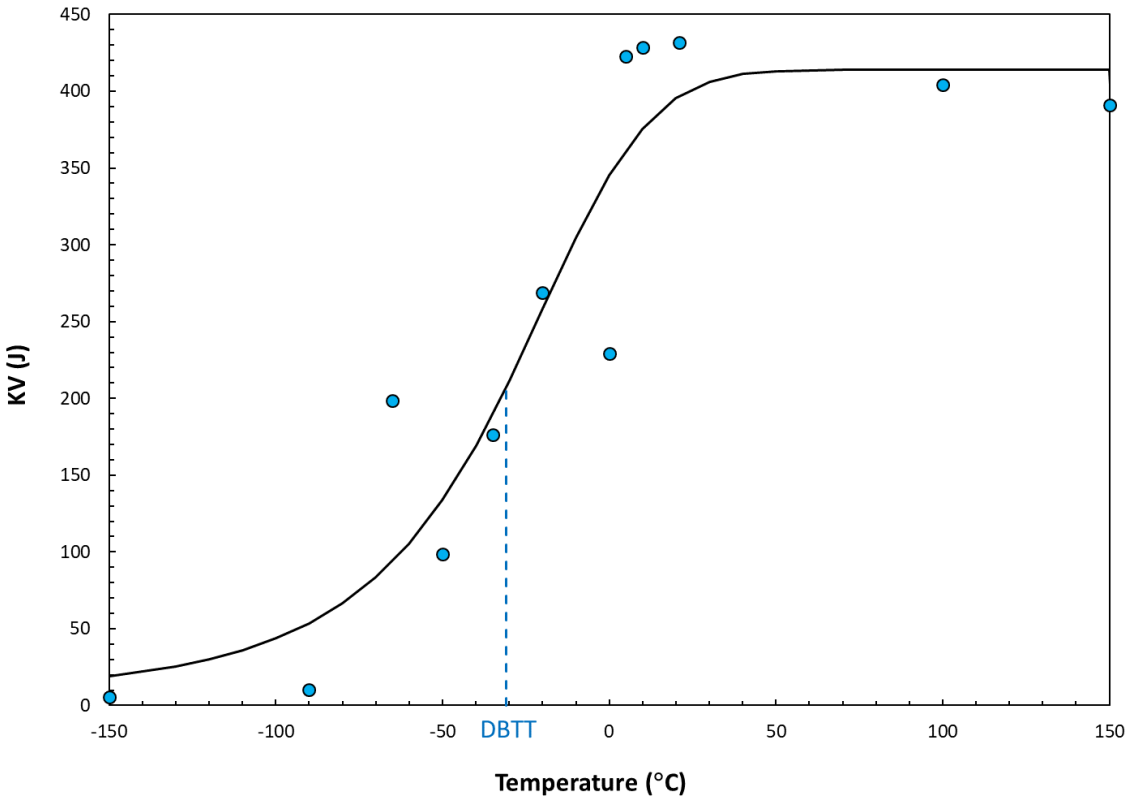


Fig. 3. Example of AHT transition curve.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

3.1.3. Burr Model (BUR – Asymmetric)

In probability theory, statistics, and econometrics, the *Burr Type XII distribution* or simply the *Burr distribution* [12] is a continuous probability distribution for a non-negative random variable. It is most commonly used to model household income. In 1996, Windle [13] proposed modeling Charpy data with the Burr distribution. Cao [14] found this approach to provide a better fit than HT for both sparse and large datasets, as long as the data are well distributed across the transition region.

The BUR model is:

$$Y = \mathrm{LS} + (\mathrm{US} - \mathrm{LS})\left[1 + e^{-k(T-T_0)}\right]^{-m}, \tag{5}$$

where the last term, $\left[1 + e^{-k(T-T_0)}\right]^{-m}$, is the actual Burr distribution. Besides LS and US, the model parameters that need to be determined through the regression process are:

- k, the scale parameter of the distribution ($k \geq 0$);
- m, the shape parameter of the distribution ($m \geq 0$); and
- T_0 (°C), a location parameter that determines the position of the curve along the temperature axis and does not generally correspond to *DBTT*.

The parameters k and m do not have a straightforward physical meaning; however, the following can be stated:

- as k and/or m increase, the curvature of the lower transition region and the slope of the transition region increase;
- as k and/or m decrease, the curvature of the lower transition region and the slope of the transition region decrease.

Further details about these parameters are provided in section 4.1.1. An example of BUR transition curve is given in **Fig. 4**.

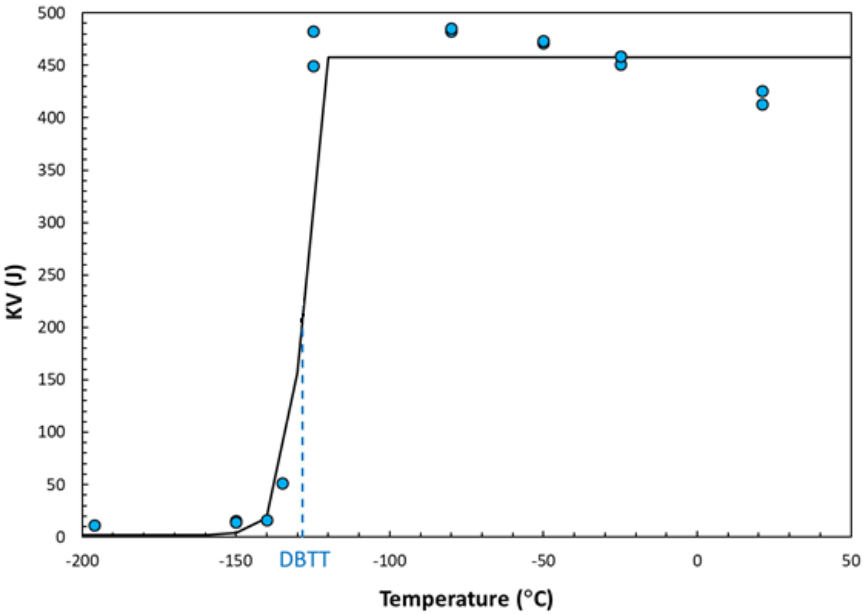


Fig. 4. Example of BURR transition curve.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

3.1.4. Arctangent Model (ACT – Symmetric)

Kohout [15] proposed a relatively simple regression model based on the arctangent trigonometric function as an alternative to the widely popular HT model, with the form:

$$Y = \frac{\text{LS+US}}{2} + \frac{\text{US−LS}}{2} \cdot arctan\left[\frac{\pi}{2C}(T - \text{DBTT})\right], \tag{6}$$

where the parameters *C* and *DBTT* have the same meaning as for the HT model.

The most significant difference between the HT and ACT models is that the curvature of the latter transition curve in the lower and upper transition regions is much larger than for the HT model. Moreover, lower and upper shelves can often exhibit gently decreasing/increasing trends, as can be seen in the example of ACT transition curve given in **Fig. 5**.

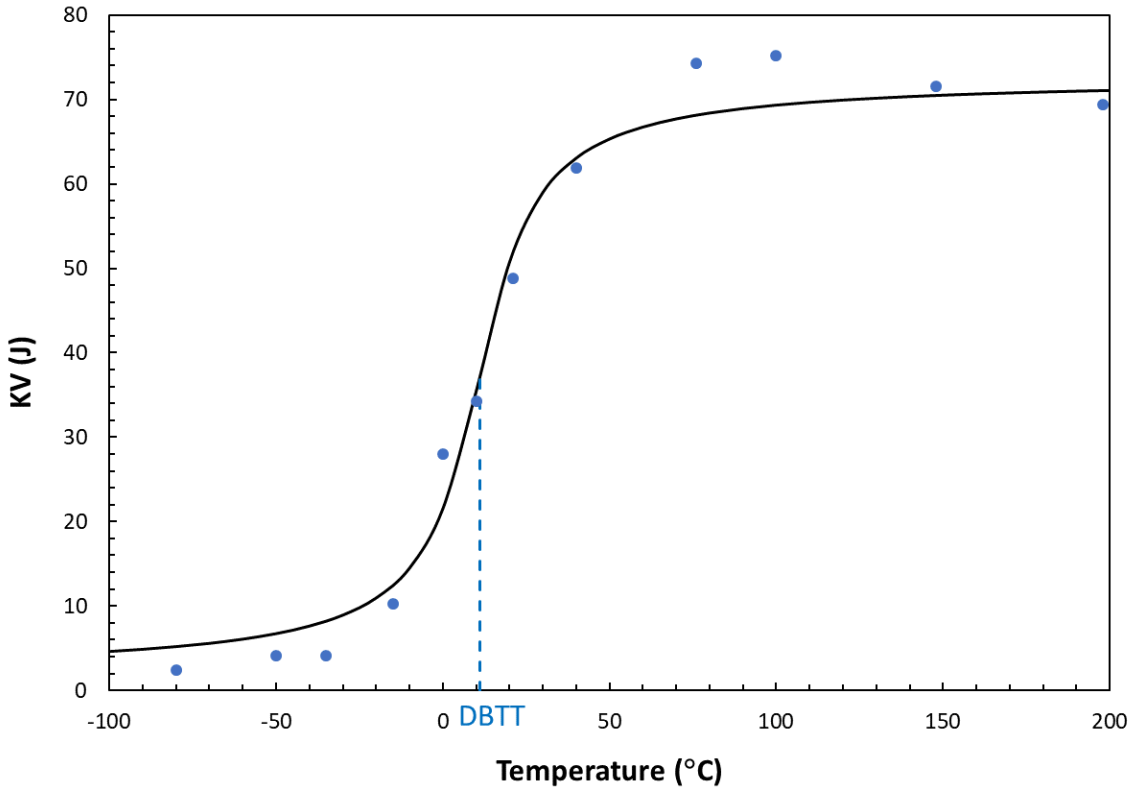


Fig. 5. Example of ACT transition curve.

3.1.5. Asymmetric Kohout Model (KHT – Asymmetrical)

Another regression model evaluated by Kohout in his 2012 Internal Report [15] was an asymmetrical model having the following "two-part" form:

$$Y = \text{LS} + \frac{\text{US−LS}}{1+p} \cdot e^{\frac{1+p}{2C}(T-T_0)} \qquad \text{for } T \le T_0\,, \quad \text{and} \tag{7}$$

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

$$Y = \mathrm{LS} - \; p \cdot \frac{\mathrm{US}-\mathrm{LS}}{1+p} \cdot e^{-\frac{1+p}{2C}(T-T_0)} \qquad \text{for } T > T_0\,. \tag{8}$$

Here, p is the parameter of asymmetry and corresponds to the ratio between the curvatures of the regression curve in the upper and lower transition region (namely, if $p > 1$, the curvature in the lower transition region is larger than in the upper transition region).

As in the case of the BUR model, T_0 is a location parameter for the curve on the x axis, and it generally does not correspond to the value of DBTT. C, on the other hand, has the same meaning as with the HT, AHT, and ACT models.

In spite of the "two-part" form of the KHT model, there is no discontinuity for $T = T_0$, as Eqs. (7) and (8) coincide in that point. An example of a KHT transition curve is given in **Fig. 6**.

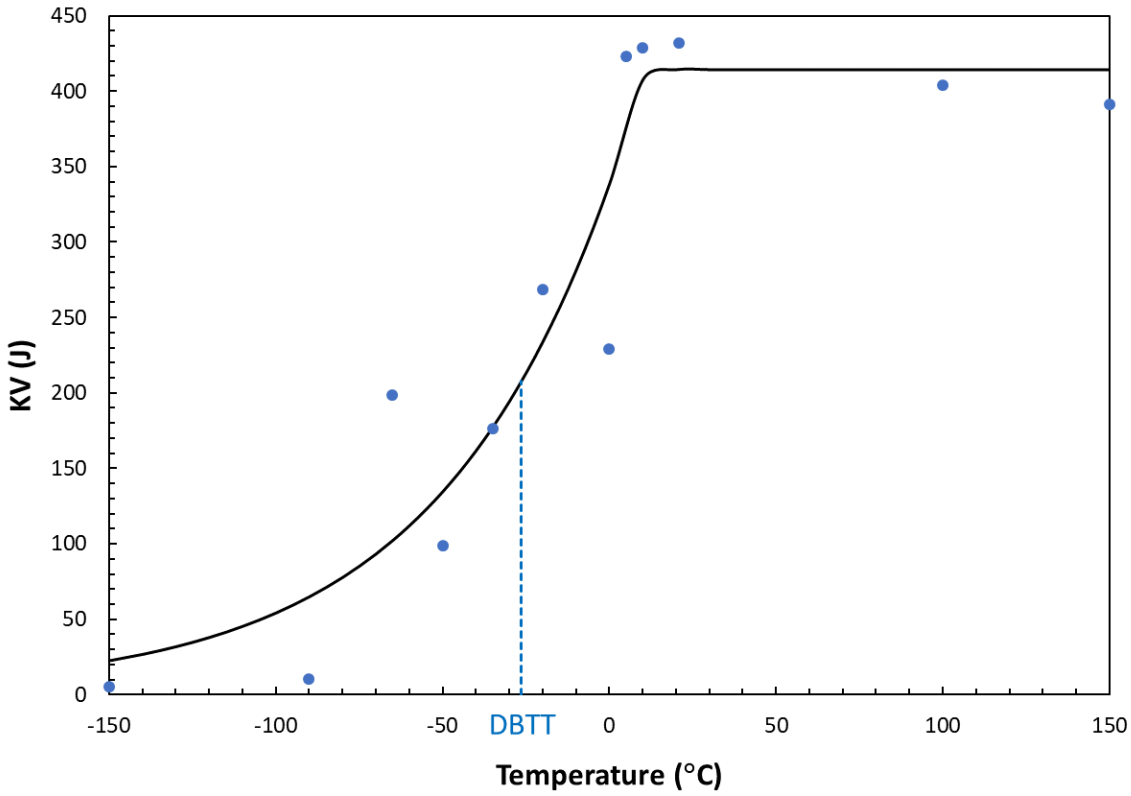


Fig. 6. Example of KHT transition curve.

3.1.6. Other Regression Models Not Selected

The literature review revealed several other less common regression models. In some cases, exponential-type curves are used to fit Charpy impact data in the transition region [16,17], but those are obviously unable to represent data in the lower and upper shelves. Other authors use error functions to describe Charpy data versus temperature [18], but data processing is overwhelmingly complex.

A popular commercial software uses a Boltzmann function, which is a simplified version of the Burr function. Other researchers [19,20] use Weibull distributions for transition

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

curve fitting. Another published approach [19] consists of dividing the regression interval into three different sections where the LS and US are provided by horizontal lines obtained from weight calculations, while regression in the transition region is based on an Avrami equation.

None of these models are selected for further analysis, as each one of them is deemed less user-friendly and/or more computationally intensive than the five approaches that are investigated in this report.

4. NIST Software Package

This section is dedicated to demonstrating the functionality of the NIST software package by analyzing example data. Instructions for using the NIST software package are also provided. The statistical methods included in the software package are briefly described and results are presented and interpreted.

In section 4.1, dataset #1 is used to demonstrate the use of the Excel spreadsheet to obtain initial values for model parameters. Dataset #1 is a simulated dataset of absorbed energy values, generated using the HT model with the following model parameters: LS = 0 J, US = 230 J, C = 50 °C, and DBTT = 0 °C. Test temperatures range between -200 °C and 200 °C.

Dataset #2, analyzed in sections 4.2 and 4.3, is a well-behaved set of Charpy test results from API X52, a vintage pipeline steel manufactured in the 1960s. The dataset comprises values of absorbed energy, lateral expansion, and shear fracture appearance; however, only absorbed energy will be examined.

All datasets are provided in the NIST software package. Datasets #1 and #2, displayed in **Fig. 7**, both have relatively gentle slopes between the lower and upper shelves.

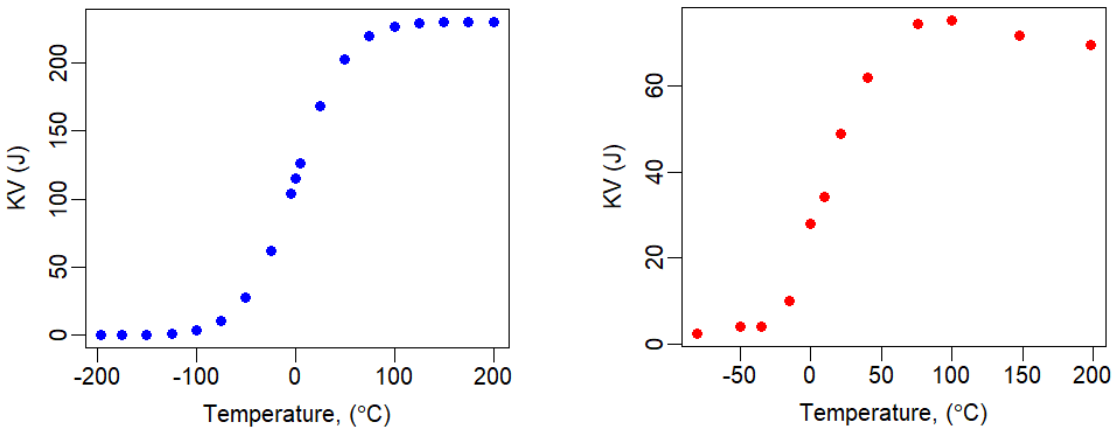


Fig. 7. Left: Absorbed energy for simulated dataset #1. Right: Absorbed energy for experimental dataset #2.

4.1. Excel Spreadsheet for Initial Values

Model parameters determined by nonlinear regression of Charpy test results are extremely sensitive to their initial values. In some cases, completely meaningless results can be obtained when the initial values of the model parameters are inadequate, as various local minima of the

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

root-mean-square error (RMSE) are often encountered during the regression process. Providing good initial values is therefore essential to ensure that optimization algorithms converge to a global (rather than local) minimum. However, even if adequate initial values are specified, there is no guarantee that any model will fit the data and provide acceptable results: in this case, the specific model "does not converge". (See Appendix A.1 for information regarding convergence.)

To facilitate the determination of reasonable initial values for the model parameters, the NIST software package includes a macro-enabled Excel spreadsheet called "*Establishment of initial values for Charpy regressions.xlsm*". The spreadsheet only considers the three asymmetric fitting models, *i.e.*, AHT, BUR, and KHT. The two symmetric models (HT and ACT) are simpler, and it should not be hard, even for a relatively inexperienced user, to choose good initial values for DBTT (transition temperature) and C (transition region half-width)[2]. The use of the Excel spreadsheet for initial values is illustrated using dataset #1.

4.1.1. Instruction for Use

The spreadsheet has three worksheets, one for each asymmetric model (AHT, BUR, KHT). The worksheet for the AHT model is where the user enters the experimental values of test temperature and Charpy result (KV, LE, or SFA) in the two columns with yellow background[3] (**Fig. 8**). The spreadsheet can accommodate up to 32 experimental data points. These input data are automatically copied into the remaining two worksheets (BUR and KHT).

In the central portion of the screen, the equation of the specific model is provided, with the $Y(T)$ vs. T plot below. The latter shows the experimental results, the model curve drawn using the current parameter values, and the position of the ductile-to-brittle transition temperature (indicated by a dashed line). The right side of the screen contains the values of temperature and $Y(T)$ for the fitted curve.

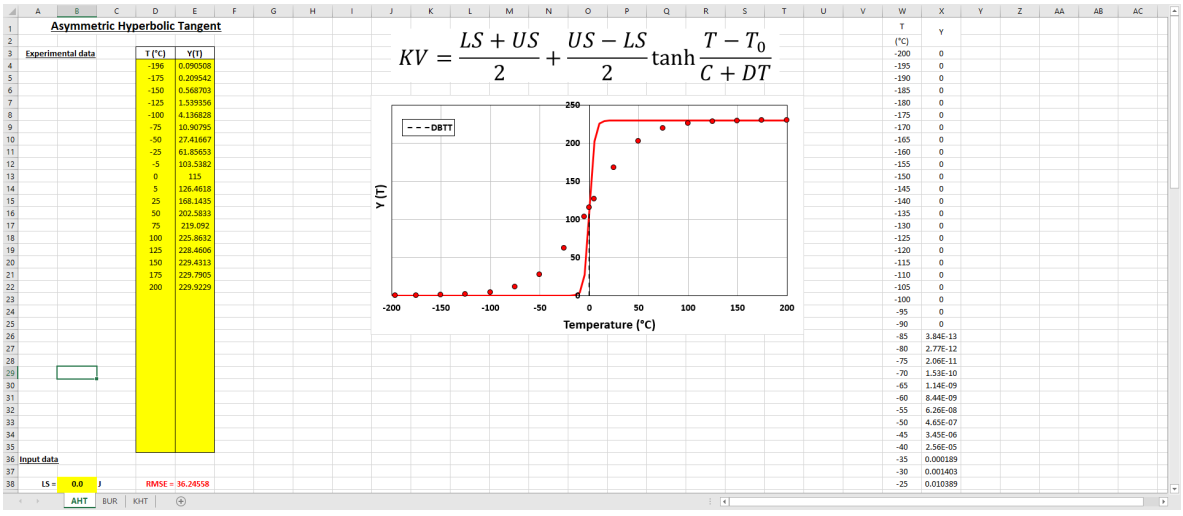


Fig. 8. Worksheet for the AHT model (upper part).

[2] Moreover, the initial values of C and DBTT obtained for the AHT model can also be used as initial values for both HT and ACT.
[3] In this spreadsheet, all cells that require input from the user are highlighted in yellow.

In the lower part of the worksheet (**Fig. 9**), the user must input lower shelf (LS) and upper shelf (US) values. These can be specific numbers, or the averages of specific test results. On the right side, the current value of RMSE is displayed in red. Note that, for the BUR and KHT models, the value of DBTT (which does not correspond to a specific model parameter) is also displayed above RMSE.

Below the input data, the initial values of the model parameters are displayed. The user can input values directly in the yellow cells or click on the buttons on the right side in order to quickly increase/decrease/multiply/divide etc. the parameters.

Some parameters are subject to constraints, namely: C, k, m, and p cannot be negative or zero. If the user tries to input a value ≤ 0, an error message appears, and the entry is rejected. Clicking the button "RESET" on the right causes parameters to assume generic default values (AHT: C = 25 °C, D = 0.0001, $DBTT$ = 0 °C; BUR: k = 1, T_0 = 0 °C, m = 0.1; KHT: C = 10 °C, T_0 = 0 °C, p = 1).

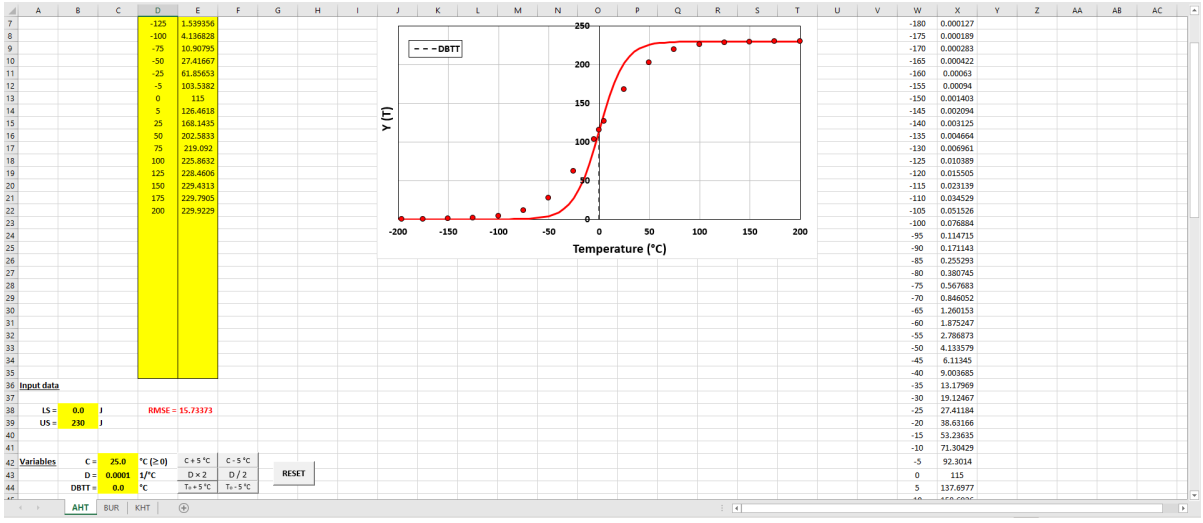


Fig. 9. Worksheet for the AHT model (lower part).

When trying different initial values of the model parameters, the user can check the acceptability of the corresponding regression curve in the plot and in terms of RMSE value (lower is better). Once initial values appear reasonable for the dataset under consideration, iterations can stop, and the initial values can be used for the actual fitting process.

This spreadsheet can also help clarify the practical meaning of some model parameters for two of the three asymmetric models (BUR and KHT), as shown in the examples below.

BUR: parameters k and m (**Fig. 10** and **Fig. 11**)

Both k and m affect the slope of the curve in the transition region (the higher k and/or m, the steeper the transition); lower values of k and/or m also enhance the asymmetry of the curve.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

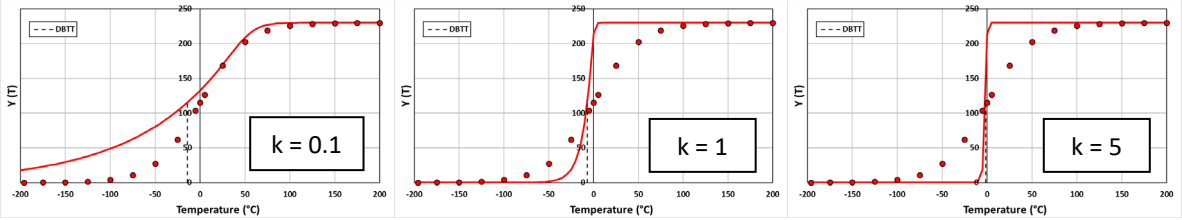


Fig. 10. Effect of *k* parameter on the BUR regression curve (*m* = 0.1).

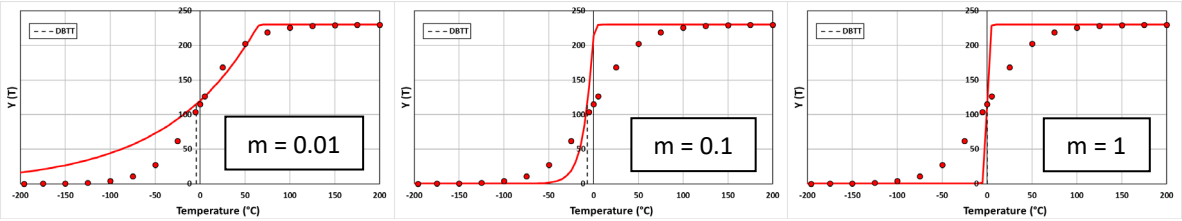


Fig. 11. Effect of *m* parameter on the BUR regression curve (*k* = 1).

KHT: parameter *p* (**Fig. 12**)

p = 1 corresponds to a symmetric transition curve. When *p* < 1, the curvature in the lower transition region is larger than in the upper transition region, which exhibits a "sharp knee". When *p* > 1, the situation is reversed (sharp knee in the lower transition region, smooth curve in the upper transition region).

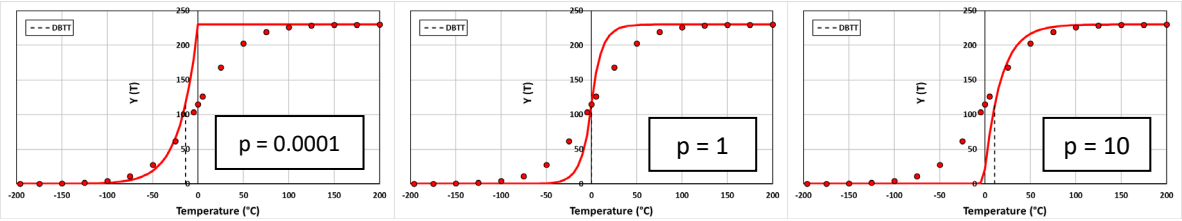


Fig. 12. Effect of *p* parameter on the KHT regression curve.

The spreadsheet is used to establish initial values for datasets #1 and #2 as well as for dataset #3 analyzed in section 5.3. The initial values are summarized in **Table 1**, along with LS and US values used in the analyses.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

Table 1. Values of LS, US, and initial values of model parameters used in the analyses performed in this document.

Dataset	LS	US	Model	Parameters					
				C	D	T_0	k	m	p
#1	0 J	230 J	HT	15		25			
			AHT	15	0.0001	25			
			BUR			35	0.5	0.05	
			ACT	15		25			
			KHT	35		0			2
#2	2 J	72.60 J	HT	50		10			
			AHT	30	0.0001	10			
			BUR			45	0.5	0.05	
			ACT	50		10			
			KHT	20		5			1
#3	2 J	441.87 J	HT	30		-80			
			AHT	30	0.0001	-80			
			BUR			-70	2	0.05	
			ACT	30		-80			
			KHT	5		-80			2

4.2 The Shiny App

The Shiny App is currently hosted at https://dnewt.shinyapps.io/charpyapp/.

The purpose of the Shiny App is to fit nonlinear models to temperature transition data for Charpy impact specimens, with the aim of generating parameter estimates and their standard errors, plots of data with fitted curves, diagnostic plots, and some statistics for model selection. In addition, the Shiny App can provide characteristic temperature estimates and their associated uncertainties for selected values of KV, LE, or SFA. Details of the computational and statistical aspects of the Shiny App are discussed in Appendix A. In this section, we describe the use of the Shiny App and present the dataset #2 analysis.

In the Shiny App, each model can have either fixed or variable (non-fixed) lower and/or upper shelves, unlike the Excel spreadsheet for a basic analysis that only accommodates the fixed-shelves configuration (recommended option). Mixed options, such as fixed lower shelf and variable upper shelf, are also possible.

For Charpy test results, we recommend using the following approach to estimate the values of LS and/or US used in an analysis with one or more "fixed" shelves:

- Lower shelf: should be set to a predetermined value, based on typical behavior for the material tested (for example: for KV, LS = 2 J, and for LE, LS = 0 mm are good choices for steels). Alternatively, and if enough data points are available in the lower shelf, LS can be set equal to the average of KV or LE for specimens with SFA ≤ 5 % [21,22] or with SFA = 0 %.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

- Upper shelf: should be set to the average value of KV or LE for the specimens that exhibited SFA ≥ 95 % [21,22]. Other threshold SFA values that have been used are 90 % and 100 %.

If SFA values are not available to define the shelf regions, the test data to be averaged for determining LS and/or US levels are generally selected from a visual analysis of the dataset, based on engineering judgement. Fixing one or more shelves reduces the number of model parameters to be determined by the regression process, which can be helpful in case of sparse datasets (8 data points or less).

While shelves can be "fixed" in the regression analysis for KV and LE, the user-specified values of LS and US are estimated with some uncertainty; we do not have perfect knowledge of the values of the fixed shelves. Thus, the user must input uncertainties associated with each fixed shelf. The uncertainties of fixed shelves are discussed in Appendix A.5.

Once data are uploaded to the Shiny App, models are selected, and shelves are specified as either "Fixed" or "Variable", the user must input initial values for each model parameter. The default initial values provided in the Shiny App for model parameters are based on engineering judgment, so these values may need to be adjusted to accommodate the user's data set. The initial values for dataset #2 are listed in **Table 1** for all models.

The user-input section of the Shiny App is on the left side of the screen, while the output (organized in tabs labelled "*Regression Results" and "Diagnostic Plots"*) is displayed on the right (**Fig. 13**).

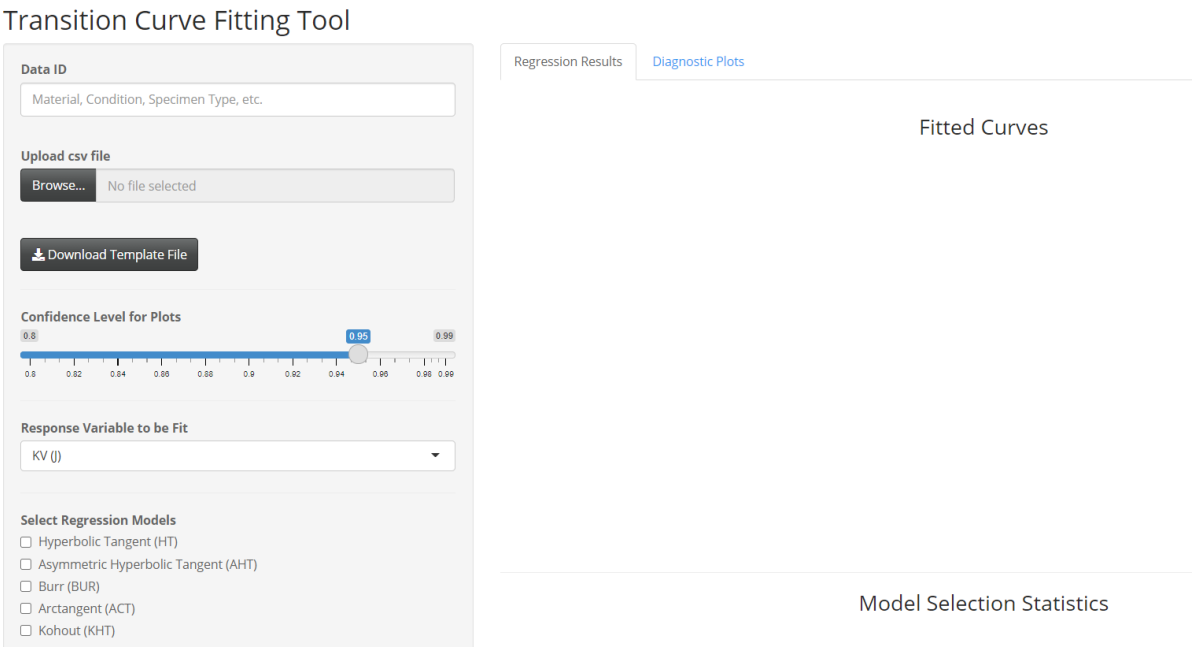


Fig. 13. Screenshot of the top right portion of the Shiny App.

The following information, from top to bottom, must be entered by the user:

(1) Data ID: information about the dataset to be analyzed.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

(2) Upload csv file: experimental data points, in the form of a csv (comma separated values) file containing temperature and *Y* values organized in separate columns, with simple headings ("temperature" and "*Y*"). A template file (*template_file.csv*) can be downloaded by clicking the button just below the file selection field. (The template file contains LE responses for dataset #2.) Some of the input boxes described below do not appear until the data has been uploaded.

(3) Confidence Level for Plots: the user can drag the slider to select the desired confidence level for the bounds of the transition curves. It is recommended to keep the default confidence level (95 %).

(4) Response Variable to Be Fit: possible choices are KV (J), LE (mm), SFA (%), and "Other" (for example: fracture toughness).

(5) Select Regression Model: all five models may be selected, or just a subset of them.

(6) Lower Shelf: available options are "Fixed" (default) and "Variable".
   a. If "Fixed" is selected, the user is asked to enter the fixed value of LS. For KV and LE, the default LS value is the minimum value of the response variable.
   b. For the SFA response, only the "Fixed" option is available and LS = 0; LS cannot be changed by the user.
   c. If the "Variable" option is selected, the initial value of the LS is input here.

(7) Upper Shelf: available options are "Fixed" (default) and "Variable".
   a. If "Fixed" is selected, the user is asked to enter the fixed value for the US. For KV and LE, the default US value is the maximum value of the response variable.
   b. For the SFA response, only the "Fixed" option is available and US = 100; US cannot be changed by the user.
   c. If the "Variable" option is selected, the initial value of the US is input here.

(8) Lower Shelf Uncertainty: input box only appears if the "Fixed" shelf option is selected. (See Appendix A.5 for details.)
   a. For KV and LE responses, the user must specify the uncertainty of the LS value. The default values of uncertainty for KV and LE are 0.3 J and 0.03 mm, respectively, based on engineering judgement. These values may not be appropriate for all data.
   b. For "Other" responses, the user must specify the uncertainty of the LS; there is no default value provided for "Other" responses.
   c. No LS uncertainty is required for the SFA response since the LS is assumed to be known without error.

(9) Upper Shelf Uncertainty: input box only appears if the "Fixed" shelf option is selected. (See Appendix A.5 for details.)
   a. For KV and LE responses, the user is asked to enter the uncertainty of the US. The default value, 0.05·US (5 % of the US value) for both KV and LE responses, may not be appropriate for all data.
   b. For "Other" Responses, the user must specify the uncertainty of the US; there is no default value for "Other" responses.
   c. No US uncertainty is required for the SFA response since the US is assumed to be known without error.

(10) Number of Additional Characteristic Temperatures to be Estimated: the user can choose up to 3 temperatures. If 1, 2, or 3 are selected from the drop-down menu, the user must enter the corresponding response value(s), *e.g.*, 28 J, 41 J, etc.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

(11) Initial Values for Hyperbolic Tangent and Arc Tangent Models (only if HT, AHT, and/or ACT models are selected): the user must enter initial values of the model parameters *C*, *D*, and *DBTT*. Default values are: 25 °C, 0.0001 1/°C, and the median value of test temperature in the dataset.
(12) Initial Values for Burr Model (only if the BUR model is selected): the user must enter initial values of the model parameters *k*, *m*, and T_0. Default values are: 1, 0.1, and the median value of test temperature in the dataset.
(13) Initial Values for Kohout Model (only if the KHT model is selected): the user must enter initial values of the model parameters *C*, *p*, and T_0. Default values are: 25 °C, 1, and the median value of test temperature in the dataset.

For the dataset #2 KV data analyzed in this section, all five models are fit with the "Fixed" shelf option for both shelves. The value of LS is set to 2 J, while the US is estimated using the recommended approach, yielding 72.60 J. Uncertainties of LS and US are set to the default values of 0.3 J and 3.63 J (5 % of US), respectively. Initial values for all parameters are listed in **Table 1**.

Clicking "*Go*" launches calculations. After a few seconds, the program output appears on the right side of the screen. The results obtained for dataset #2 (absorbed energy) are presented in **Fig. 14**, **Fig. 15**, and **Fig. 16**. Numerical and graphical results are presented under three separate tabs:

(a) Regression Results Tab (**Fig. 14**, **Fig. 15**, and **Fig. 16):** from top to bottom,

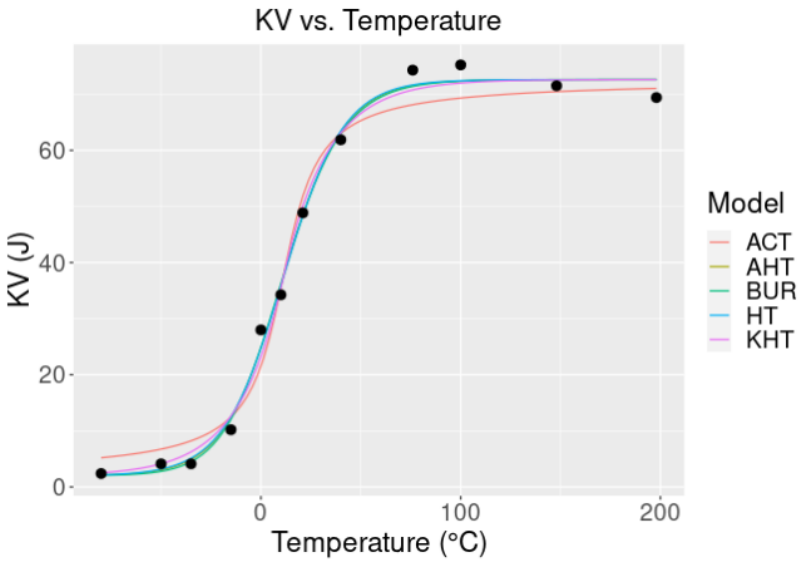
- Experimental data points and transition curves (the user can choose which curves are actually displayed by checking or unchecking the boxes under "Fits to show"); curves can be shown with or without confidence bounds by selecting "*Yes*" or "*No*" from a drop-down menu.
- Model selection statistics for each selected model (RMSE, AIC, BIC), and whether convergence is achieved. The models are listed in order of decreasing RMSE by default. Results can be organized in terms of the increasing or decreasing order of other model selection statistics by clicking the arrows next to that model selection statistic.
- Table of regression parameter estimates for each selected model, with standard errors (*S.E.*), lower bounds of the confidence intervals (*Lower CI*), and upper bounds of the confidence intervals (*Upper CI*).
- Table of Ductile-to-Brittle Transition Temperature (DBTT) values, with standard errors (*S.E.*), lower bounds of the confidence intervals (*Lower CI*), and upper bounds of the confidence intervals (*Upper CI*).
- Table of additional characteristic temperatures, with standard errors (*S.E.*), lower bounds of the confidence intervals (*Lower CI*), and upper bounds of the confidence intervals (*Upper CI*).

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

Regression Results | Diagnostic Plots | Download

Fitted Curves



Fits to Show

- [x] HT
- [x] BUR
- [x] AHT
- [x] KHT
- [x] ACT

Show Confidence Bounds?

No

Model Selection Statistics

	Model	RMSE	AIC	BIC	Converged?
1	HT	2.231	23.067	24.521	Yes
3	BUR	2.322	24.765	26.705	Yes
2	AHT	2.348	25.03	26.97	Yes
5	KHT	2.835	29.56	31.5	Yes
4	ACT	4.085	37.587	39.041	Yes

Showing 1 to 5 of 5 entries

Fig. 14. Shiny App output: “Regression Results” tab (top section).

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

Parameter Estimates

	Model	Parameter	Estimate	S.E.	Lower CI	Upper CI
1	HT	C	30.1411	3.7906	24.2328	38.6718
2	HT	DBTT	11.2705	2.3949	6.8726	16.4404
3	BUR	k	0.0609	0.0514	0.0409	0.1349
4	BUR	T0	4.5479	20.0738	-50.1607	25.992
5	BUR	m	1.3394	4.2627	0.3634	10.8191
6	AHT	C	29.9782	3.1197	24.5359	36.82
7	AHT	DBTT	11.1687	1.7391	8.0136	14.6074
8	AHT	D	0.0141	0.1646	-0.1566	0.4082
9	KHT	C	21.8974	4.1089	15.3058	31.4752
10	KHT	T0	11.1343	6.8388	-2.9751	25.0502
11	KHT	p	1.0059	1.2878	0.2644	2.7297
12	ACT	C	20.6024	5.2075	12.4584	32.8808
13	ACT	DBTT	11.1153	2.4831	6.5866	16.398

Showing 1 to 13 of 13 entries

DBTT Table (°C)

	Model	Estimate	S.E.	Lower CI	Upper CI
1	HT	11.2705	2.3949	6.8726	16.4404
2	BUR	10.9369	1.785	6.8989	14.975
3	AHT	11.1687	1.7391	8.0136	14.6074
4	KHT	11.1989	2.993	4.4284	17.9695
5	ACT	11.1153	2.4831	6.5866	16.398

Showing 1 to 5 of 5 entries

Fig. 15. Shiny App output: “Regression Results” tab (middle section).

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

Additional Characteristic Temperatures (°C)

	Model	Ref Value	Temperature Est	SE	Lower CI	Upper CI
1	HT	28	3.14	2.47	-2.36	8.63
2	HT	41	14.44	2.42	9.05	19.83
3	BUR	28	2.86	2.5	-2.8	8.52
4	BUR	41	14.15	2.43	8.66	19.64
5	AHT	28	3.07	2.51	-2.6	8.74
6	AHT	41	14.34	2.55	8.58	20.11
7	KHT	28	4.52	3.15	-2.61	11.66
8	KHT	41	13.38	3.47	5.53	21.22
9	ACT	28	5.35	3.82	-3.17	13.88
10	ACT	41	13.29	3.35	5.83	20.76

Showing 1 to 10 of 10 entries

Fig. 16. Shiny App output: "Regression Results" tab (bottom section).

Unlike linear least-squares regression, nonlinear least-squares regression is implemented using a numerical optimization algorithm. Thus, model parameters do not always converge to a globally optimal solution, and it's possible that the algorithm won't be able to converge at all. (See Appendix A.1 for a discussion of model convergence.) Often, convergence is not achieved because of "bad" initial values. The spreadsheet described in section 4.1 provides a tool for obtaining initial values for AHT, BUR, and KHT models. Results will not be printed and curves will not be plotted for models that do not converge to the optimal solution.

Three model selection statistics, RMSE, AIC (Akaike Information Criterion) [23], and BIC (Bayesian Information Criterion) [24] are provided for each model. The RMSE, AIC, and BIC statistics are relative measures, so there is no absolute threshold indicating whether the model provides a good fit to the data. These statistics simply allow models to be compared: the model with the smallest values of RMSE, AIC, and/or BIC is preferred. See Appendix A.2 for a discussion regarding the model selection statistics used in the Shiny App.

The model selection statistics (RMSE, AIC, and BIC) for dataset #2 are listed in **Fig. 14** under the heading "Model Selection Statistics". The models are listed in increasing order of RMSE, which means the best-fitting model according to RMSE is the HT model. The HT model also has the smallest AIC and BIC values.

The three model selection statistics often agree on which model is the best, but occasionally there can be some discrepancies since they are designed to capture different aspects of the models. For example, BIC more strongly favors parsimony and penalizes models based on the number of parameters. When discrepancies occur, it is up to the user to select the

"best" model based on these statistics and other sources of information, such as the diagnostic plots.

For dataset #2, estimated model parameters for each model selected are listed in **Fig. 14** under the heading "Parameter Estimates". **Fig. 15** ("DBTT Table") also displays DBTT estimates for all models selected. (Appendix A.8 provides information regarding the computation of DBTT for the BUR and KHT models. For all other models, DBTT corresponds to one of the model parameters and does not need to be separately estimated.)

Characteristic temperature estimates corresponding to 28 J and 41 J are listed in **Fig. 16** ("Additional Characteristic Temperatures").[4] The calculation of characteristic temperatures and their uncertainties is discussed in Appendix A.8.

(b) Diagnostic Plots Tab (**Fig. 17**): this tab displays diagnostic plots for each selected model. Diagnostic plots include residual plots, a lag plot of residuals, and a normal Q-Q plot of residuals. The results of the Shapiro-Wilk test for the normality of residuals is also provided.

Diagnostic plots should be carefully examined for all models considered to verify that regression assumptions are not violated and to ensure model adequacy. Plots should indicate that the residuals are independent and identically distributed as a normal distribution with a mean of zero and variance σ^2. For the "Residual Plot" and "Standardized Residuals" plots, the residuals should vary randomly about zero and should not have a discernible pattern. In the "Lag Plot", make sure there is no evidence of correlation. If the residuals are normally distributed, as desired, the data in the Q-Q plot should roughly fall on the reference line that indicates normality. Additionally, the results of the Shapiro-Wilk test for the normality of residuals [25] (test statistic W and probability p) are provided below the Normal Q-Q plot. If p is smaller than the level of significance $\alpha = 0.05$, the residuals cannot be considered normally distributed. A discussion of the diagnostic plots used in the Shiny App is included in Appendix A.7.

Fig. 17 displays diagnostic plots for the top-ranked HT model. The diagnostic plots for dataset #2 do not indicate gross assumption violations or model inadequacy.

[4] All these characteristic temperatures are generally of interest to the nuclear community [5].

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

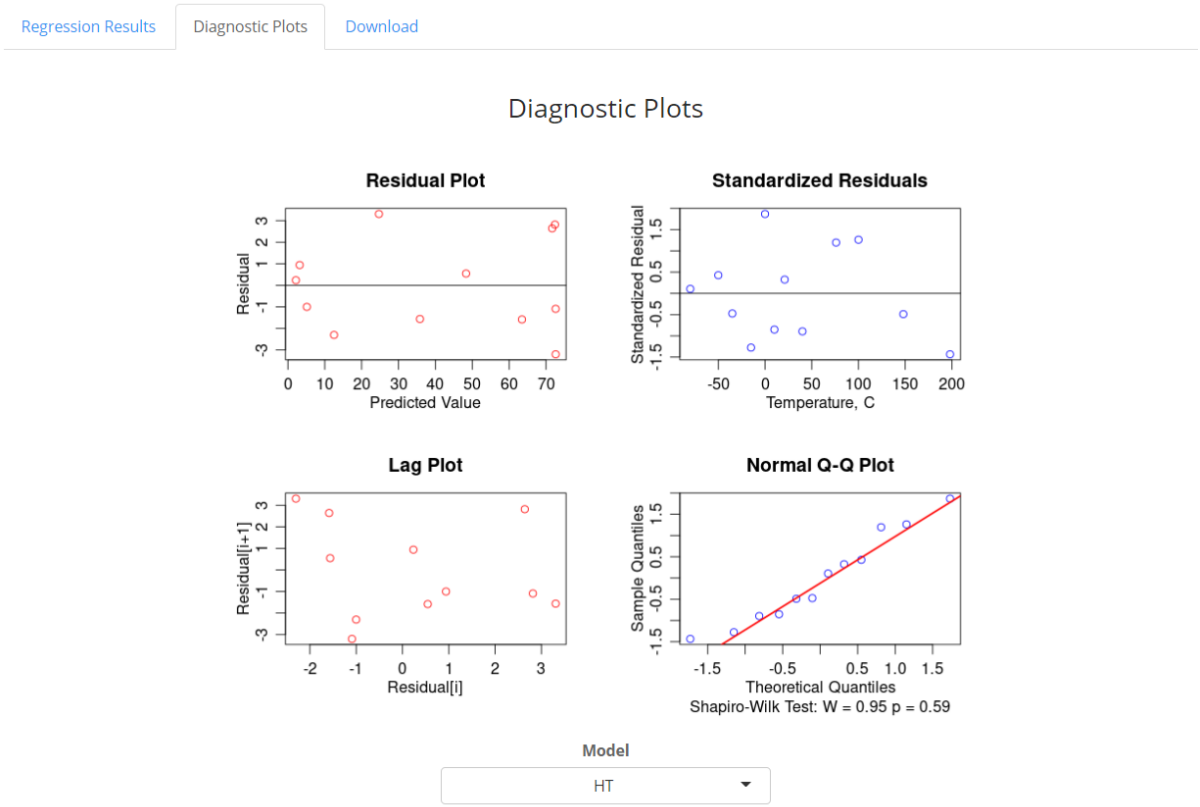


Fig. 17. Shiny App output: "Diagnostic Plots" tab for the HT model.

After examining model selection statistics and diagnostic plots, the well-established HT model is selected to represent the data. **Fig. 18** shows the fitted curve and 95 % confidence bounds for the HT model fit to dataset #2. Confidence bounds may not be symmetric about the fitted line. Confidence bounds are discussed in Appendix A.6.

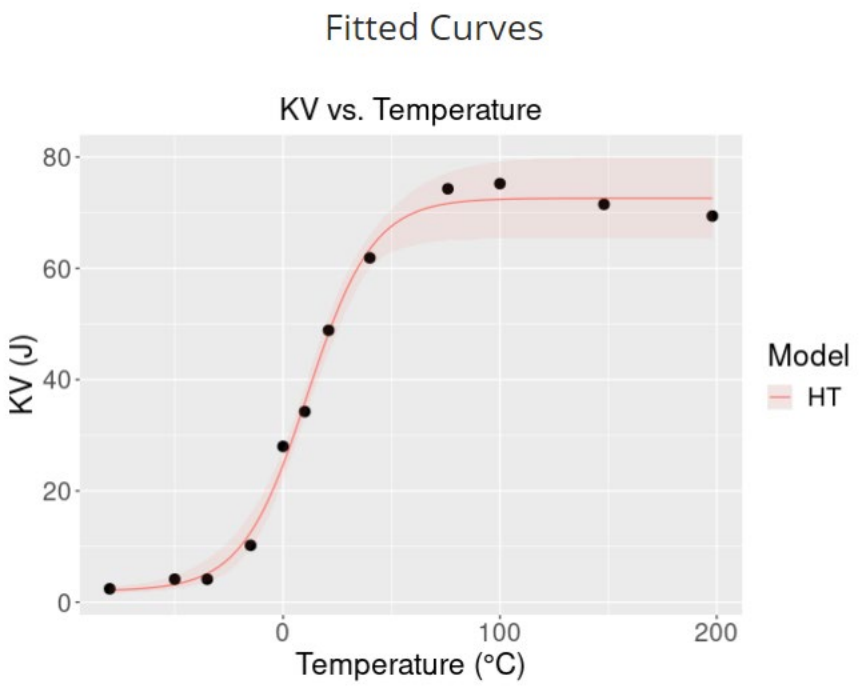


Fig. 18. Transition curve for HT model with 95 % confidence bounds for Dataset #2 (absorbed energy).

(c) Download Tab: after clicking the button "*Download All Results*", a PDF file is generated containing all tables and plots from the tabs "Regression Results" and "Diagnostic Plots". The file can be downloaded on the user's computer and printed. This tab does not appear on the Shiny App until the data is uploaded.

4.3 Excel Spreadsheet for Basic Analysis

For a user who is only interested in fitting Charpy data with different models and picking the best-fitting model, another macro-enabled Excel spreadsheet ("*Charpy data regression.xslm*") is provided for a basic-level analysis. The spreadsheet can also provide characteristic temperatures for any value of the independent variable (KV, LE, or SFA), and offers a graphical comparison of the different regression curves. Standard errors of parameter estimates and confidence bounds are not available in this spreadsheet. The results of the Shiny App and the outcome of the Excel spreadsheet are very close and often identical, as explained in section 6.

Note that all worksheets are protected, so that users may only edit cells with yellow background, which correspond to input cells, and charts. Sheets can however be easily unprotected without having to enter a password (*Review* → *Unprotect Sheet*).

4.3.1. Sheet "*Input Data*" (Fig. 19)

Charpy test results (T, KV, LE, SFA) are entered in cells A3:D32. The spreadsheet is currently able to accommodate data for a maximum of 30 tests. The user must select the

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

variable (KV, LE, or SFA) to fit by means of a drop-down menu located on the right. **Fig. 19** displays the data input for dataset #2.

If SFA data are available, column D indicates which tests correspond to upper shelf behavior (SFA ≥ 95 %), and the corresponding average value is displayed in cell B35. The lower shelf default value in cell B36 is 2 J for KV, 0 mm for LE, or 0 % for SFA. However, the user can also enter other values of US and LS in cells F35 and F36. If these cells are left empty, the program will use the values in B35:B36 for the regression.

This sheet also displays a button "*SAVE FILE AS…*", which allows the user to save the spreadsheet under a name of their choice in their preferred location.

4.3.2 Sheets "*HT*" (Fig. 20), "*AHT*", "*BUR*", "*ACT*", and "*KHT*"

Values of temperature and selected Y variable (KV, LE, or SFA) are automatically copied in columns A and B (rows 3 to 32). The Y variable from dataset #2 selected for analysis in **Fig. 20** is KV. Column C shows the corresponding values of the regression (transition) curve ($\hat{Y}$), and column D the calculated residuals $Y - \hat{Y}$.

In the case of the BUR model, values of the function

$$\xi_i = \left[1 + e^{-k(T_i - T_0)}\right]^{-m} \tag{9}$$

are displayed in column C. For the KHT model, columns C and D contain values of the functions shown in Eqs. (10) and (11) for $T_i \leq T_0$ and $T_i > T_0$, respectively.

The chart on the right displays experimental data points, transition curve, and DBTT. Charts can always be modified/reformatted, but two buttons are provided to the user for quickly formatting the Y axis ("*Auto Y scale*": bounds and units are set automatically; "*SFA Y scale*": maximum = 100, major unit = 25, minor unit = 5).

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

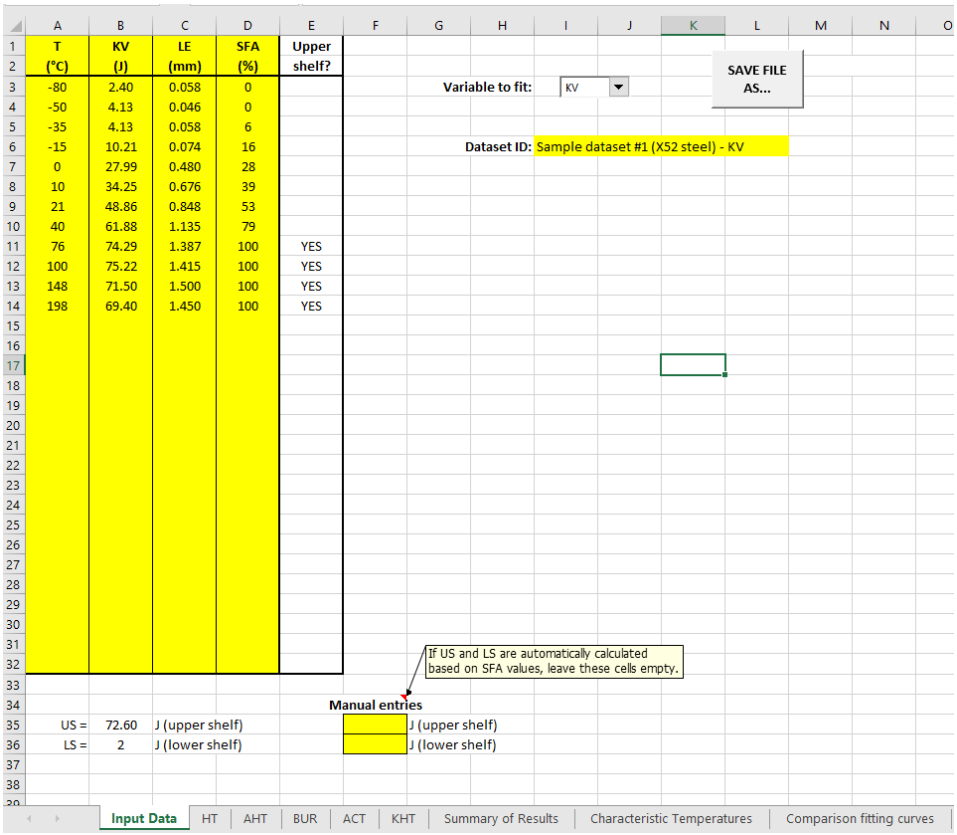


Fig. 19. Sheet "*Input Data*".

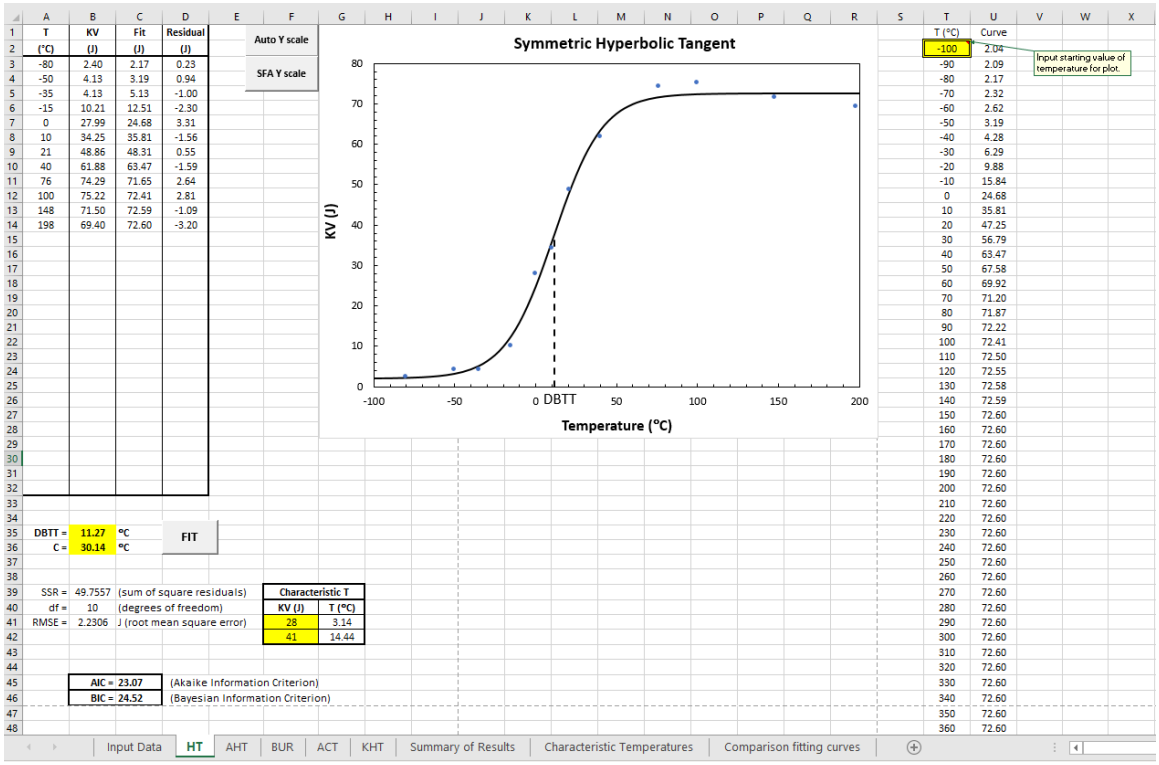


Fig. 20. Sheet "*HT*".

Temperatures and *Y* values (response variables) for the regression curve are given on the right side (columns T and U). The user can input the initial temperature value in cell T2. Values of the dependent variable *Y* are calculated every 10 °C.

In the bottom part of the sheet, the user enters initial values of the model parameters in cells B35-B36 (symmetric models) or B35-B37 (asymmetric models). Clicking the button "*FIT*" launches the *Solver* tool. NOTE: particularly for asymmetric models (AHT, BUR, and KHT), it is recommended to click "*FIT*" at least three times until the value of RMSE (B41) becomes stable and does not change appreciably.

If characteristic temperatures are required (max 2), the user must enter the corresponding *Y* values (KV, LE, or SFA) in cells F41-F42. Calculated temperatures are displayed in cells G41-G42. If no characteristic temperatures are required, the input cells F41-F42 can just be left empty.

Finally, the very bottom of the sheet displays values of the AIC and BIC model selection statistics.

For the BUR and KHT models, the value of DBTT, which is not one of the model parameters, is provided in cell E35.

4.3.3 Sheet "*Summary of Results*"

The results of the regression analyses (with the exception of characteristic temperatures) are collected in this sheet and can be printed on the default system printer by clicking the button "*PRINT RESULTS*". The following outputs are displayed/printed:

- DBTT and C values for all the models.
- Model selection statistics (RMSE, AIC, and BIC).
- Estimated model parameters for all the models.
- Comparison chart[5] for the different transition curves, including experimental data points.

4.3.4 Sheet "*Characteristic Temperatures*"

If any characteristic temperatures have been calculated, they are displayed in this sheet and can be printed on the default system printer by clicking the button "*PRINT*".

4.3.5 Sheet "*Comparison Fitting Curves*"

As previously mentioned, the chart that compares transition curves for the different models is also reproduced as an individual sheet at the end of the spreadsheet.

5 Shiny App Examples

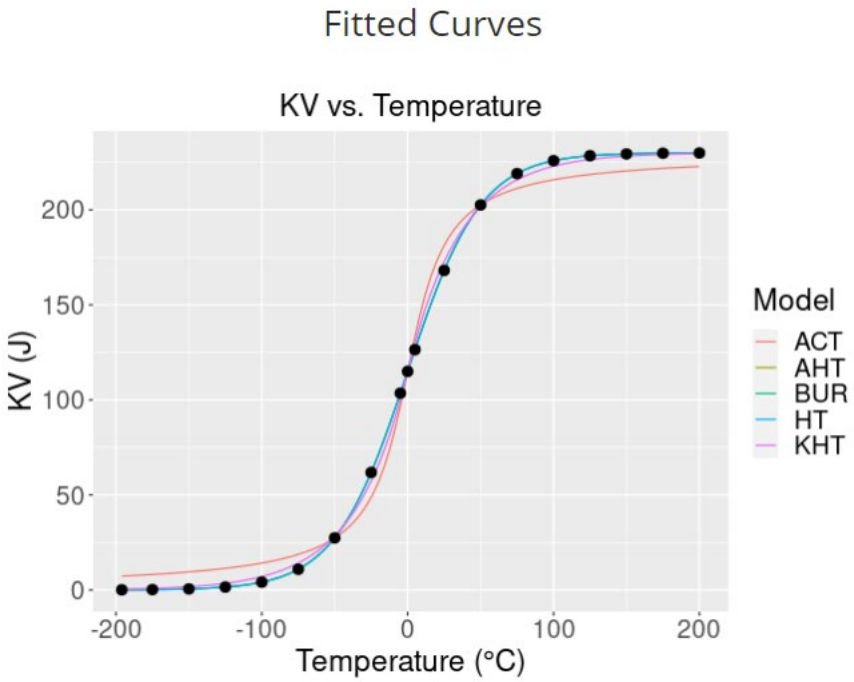
5.1 Dataset #1: Simulated

We analyze dataset #1 to illustrate the validity of the Shiny App. Because the data is simulated from an HT model with no noise, HT and AHT models are expected to reproduce the original

[5] This chart is also reproduced as a separate sheet (*Comparison Fitting Curves*) at the end of the worksheet.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

parameters with very small standard errors. Experimental data and estimated transition curves for the different models with fixed shelves are compared in **Fig. 21**.



Model Selection Statistics

	Model	RMSE	AIC	BIC	Converged?
1	HT	0.001	-255.262	-252.429	Yes
2	AHT	0.001	-253.253	-249.475	Yes
3	BUR	0.001	-253.273	-249.495	Yes
5	KHT	2.737	42.998	46.776	Yes
4	ACT	8.898	86.95	89.783	Yes

Showing 1 to 5 of 5 entries

Fig. 21. Simulated data and estimated transition curves for dataset #1 (KV). Model selection statistics are also shown.

The model selection statistics (RMSE, AIC, and BIC) for the different models are listed at the bottom of **Fig. 21**. The HT, AHT, and BUR model yielded RMSE ≈ 0 J, indicating a nearly perfect fit to the experimental data points. This is not surprising since the underlying model for the simulated data is HT. The remaining two models, KHT and ACT, yielded non-zero RMSE values, with ACT providing the worst fit. All models correctly identified DBTT = 0 °C to within 0.0005 °C , while more scatter (between 2 °C and 7 °C) is observed for T_{28J} and T_{41J}, with the ACT model yielding the most discordant values.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

5.2 Dataset #2: "Variable" Shelves

To illustrate the use of the "variable shelves" option, we fit the HT model to dataset #2 absorbed energy values. **Fig. 22** shows the fitted curve and 95 % confidence bounds.

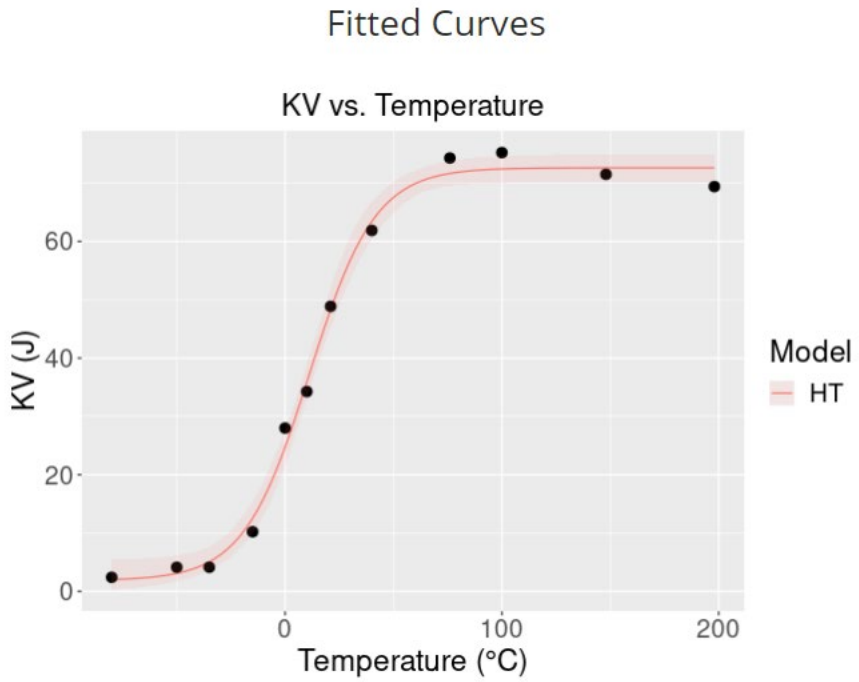


Fig. 22. Transition curve for HT model with confidence band for dataset #2 (absorbed energy).

The most noticeable difference between the "Fixed" and "Variable" shelf options is seen in the confidence bounds in the shelf areas. The confidence bounds associated with fixed shelves (**Fig. 13**) reflect the specified shelf uncertainties of 0.3 J and 3.63 J, while the bounds for the variable shelves (**Fig. 22**) are narrower and reflect the standard errors of the LS and US parameter estimates. Numerical results for the fixed and variable shelves options are shown in **Table 2**.

Table 2 - Selected numerical results for the HT model fit to the dataset #2 KV response for fixed shelf and variable shelf analyses. Standard errors of parameter estimates are shown in parentheses.

Statistic	Fixed Shelves	Variable Shelves
RMSE (J)	2.768	2.493
LS (J)	2 (0.3)	1.8561 (1.3949)
US (J)	72.60 (3.63)	72.6228 (1.2736)
C (˚C)	31.7780 (3.0261)	30.2865 (3.3029)
DBBT (˚C)	13.0196 (1.3775)	11.2039 (1.5337)
T_{28J} (˚C)	3.14 (2.75)	3.11 (2.74)
T_{41J} (˚C)	14.44 (2.87)	14.44 (2.51)

5.3 Dataset #3: Outliers

Dataset #3 consists of absorbed energy data from a X70 pipeline steel. Test results are relatively well behaved. However, data points between -100 °C and -80 °C are significantly scattered and the first one (or the remaining two) could be regarded as an outlier. This situation generally poses a significant challenge for establishing transition curves. The results from the Shiny App are compared in **Fig. 23**. All models have fixed shelves.

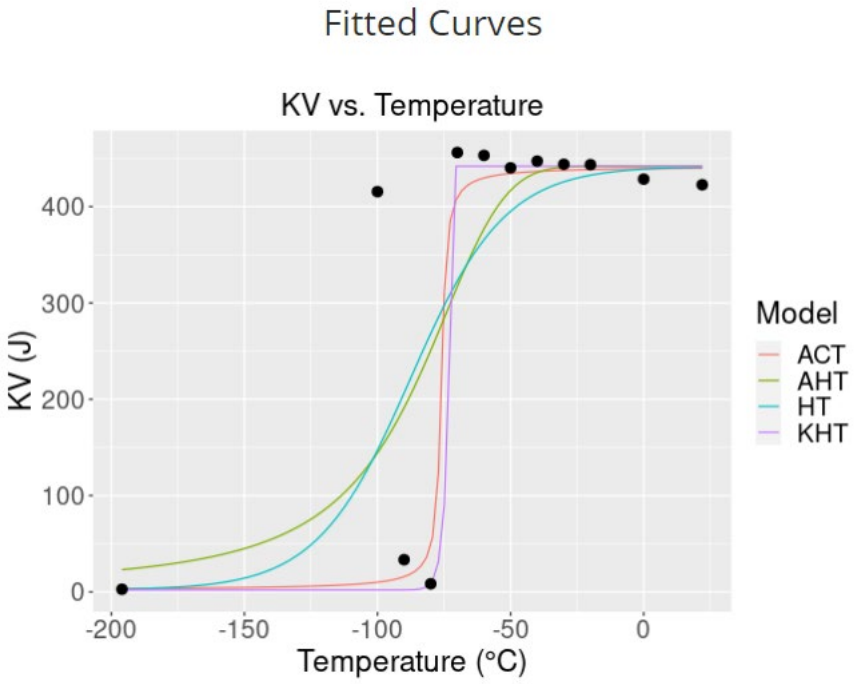


Fig. 23. Experimental data and estimated transition curves for dataset #3 (KV).

As **Fig. 23** shows, two models (HT and AHT) exhibit a gentle slope in transition region, as opposed to the ACT and KHT models, which yield a very steep slope. The BUR model does not converge. The model selection statistics (RMSE, AIC, and BIC) for the different models are shown in **Fig. 24** along with estimated model parameters. The best-fitting model, according to all fit metrics, is ACT.

Estimated $DBTT_{KV}$ values and characteristic temperatures are listed in **Fig. 25** As could be expected from the dichotomy of transition curves (shallow versus steep) seen in **Fig. 23**, the range of characteristic temperatures is huge (more than 100 °C for T_{28J}), although in the case of $DBTT_{KV}$ the range is less than 15 °C.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

Model Selection Statistics

	Model	RMSE	AIC	BIC	Converged?
4	ACT	130.229	120.675	122.13	Yes
5	KHT	138.602	122.906	124.846	Yes
1	HT	140.988	122.58	124.035	Yes
2	AHT	142.939	123.646	125.586	Yes
3	BUR				No

Showing 1 to 5 of 5 entries

Parameter Estimates

	Model	Parameter	Estimate	S.E.	Lower CI	Upper CI
1	ACT	C	2.1818	44.3243	0	64.2374
2	ACT	DBTT	-75.9327	20.6635	-93.6428	-67.545
3	KHT	C	1.0945	4.8804	1.0945	12.5718
4	KHT	T0	-71.8308	8.3832	-78.8909	-60.6112
5	KHT	p	0.0873	3.5193	0.0131	2.3677
6	HT	C	35.2369	31782509.2611	0.6967	294.109
7	HT	DBTT	-87.5186	21441813.5012	-204.9495	-68.8294
8	AHT	C	7.5281	141.8645	0	147.2999
9	AHT	DBTT	-84.5931	44.4059	-153.0338	-68.7381
10	AHT	D	-0.3422	1.2721	-1.6022	0.534

Fig. 24. Model selection statistics and parameter estimates for models fit to dataset #3.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

DBTT Table (°C)

	Model	Estimate	S.E.	Lower CI	Upper CI
1	ACT	-75.9327	20.6635	-93.6428	-67.545
2	KHT	-73.0576	2.6912	-79.1454	-66.9697
3	HT	-87.5186	21441813.5012	-204.9495	-68.8294
4	AHT	-84.5931	44.4059	-153.0338	-68.7381

Showing 1 to 4 of 4 entries

Additional Characteristic Temperatures (°C)

	Model	Ref Value	Temperature Est	SE	Lower CI	Upper CI
1	ACT	28	-83.33	22.19	-132.76	-33.89
2	ACT	41	-80.79	20.57	-126.63	-34.95
3	KHT	28	-77.36	10.46	-101.03	-53.68
4	KHT	41	-76.54	4.71	-87.2	-65.88
5	HT	28	-136.28	30.72	-204.73	-67.83
6	HT	41	-128.57	31.37	-198.46	-58.68
7	AHT	28	-180.45	33.02	-255.14	-105.76
8	AHT	41	-155.26	32.37	-228.48	-82.04

Fig. 25. DBTT table and additional characteristic temperatures for dataset #3.

For dataset #3, standard errors calculated by the Shiny App for the C and DBTT parameters of the HT model (**Fig. 24**) are 31,782,509 °C and 21,441,813 °C, respectively! The extremely large standard errors indicate that the HT model does not provide an adequate fit to the data probably due to the outlying data point at -100 ˚C that aligns with the upper shelf rather than the lower shelf. The nonlinear regression procedure cannot be expected to produce good results when the data are not well behaved. Thus, both the model selection statistics and the standard errors must be taken into account when selecting the most suitable regression model. In addition, model adequacy and regression assumptions need to be checked based on the diagnostic plots. The diagnostic plots in **Fig. 26** for the ACT model clearly show the outlying point and indicate that assumptions are violated. Perhaps more data points are needed to clarify the structure of the data, or the outlier should be removed before fitting.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

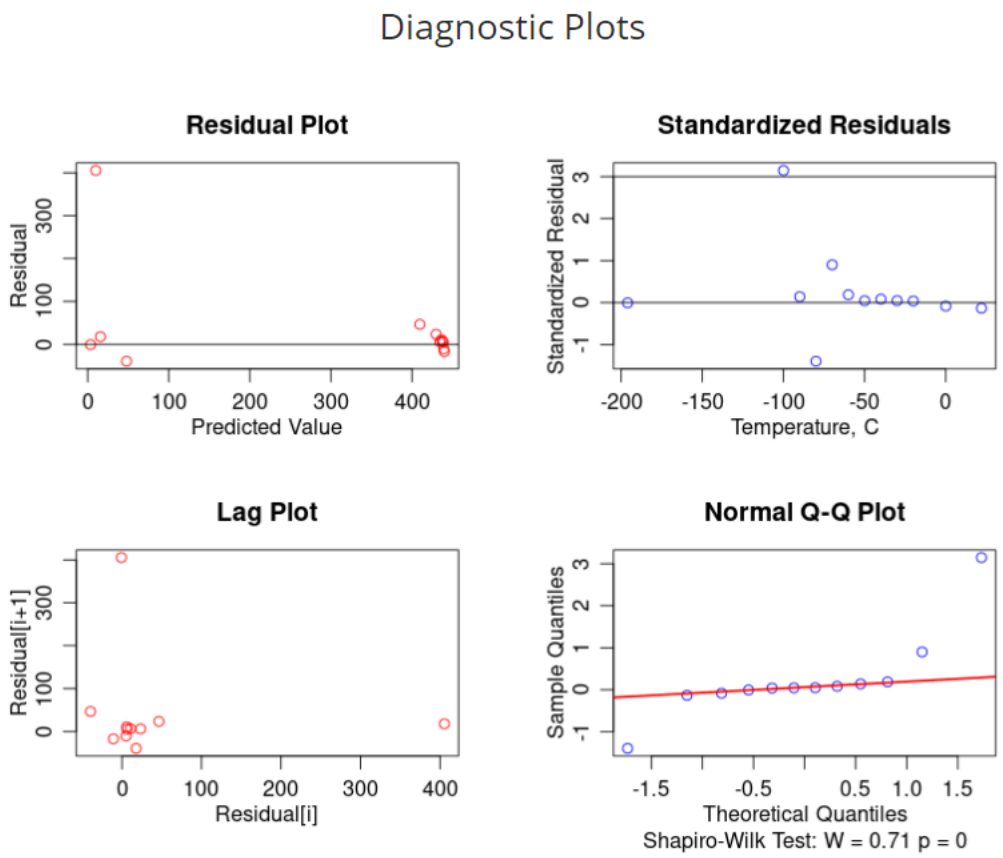


Fig. 26. Diagnostic plots for ACT model fit to dataset #3.

6 Comparison of Output from Excel Spreadsheet and Shiny App

Table 3 compares estimated model parameters, RMSE values, and characteristic temperatures provided by the Excel spreadsheet and the Shiny App for all three KV datasets (real and simulated) and all investigated models with fixed shelves. The Excel spreadsheet and the Shiny App use different fitting algorithms for estimating model parameters, so the two methods might not produce the same results depending on the data and initial parameter values.

In **Table 3**, values of DBTT and characteristic temperatures are color-coded for quick assessment of the consistency between the two programs:

(a) Values coinciding within ±1 °C are highlighted in green, and if they are equal to within ±0.01 °C, cells have a double green border.
(b) When differences are between 1 °C and 5 °C, cells are highlighted in yellow.
(c) For values differing by more than 5 °C, the cell background is pink, and the font is red.

For datasets #1 and #2, the Excel spreadsheet and Shiny App produce nearly identical estimates. For dataset #3, in which an outlier is present in the data, large temperature differences are observed for the AHT and KHT models. The AHT model appears to be the most sensitive to discrepancies between Excel and the Shiny App.

This comparison demonstrates the equivalence of using either software, if the user only requires a basic level of analysis.

Table 3 - Comparison between Excel spreadsheet and Shiny App.

Data Set 1 (virtual)

Variable	Symmetric TANH			Asymmetric TANH			Burr			ARCTAN			K-ASYM		
	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app
KV (J)	C	50.00	50.00	C	50.00	50.00	k	0.040000	0.0400	C	30.80	30.80	C	35.83	35.83
	DBTT	**0.00**	**0.00**	**DBTT**	**0.00**	**0.00**	T_0	0.00	0.00	**DBTT**	**0.00**	**0.00**	T_0	0.00	0.00
				D	1.29E-07	0	m	0.999995	1				p	1.000055	1.0001
	RMSE	9.58E-06	0.001	RMSE	0.003	0.001	RMSE	2.16E-05	0.001	RMSE	8.899	8.898	RMSE	2.737	2.737
	T_{28J}	**-49.40**	**-49.41**	T_{28J}	**-49.40**	**-49.41**	T_{28J}	**-49.40**	**-49.41**	T_{28J}	**-48.74**	**-48.74**	T_{28J}	**-50.62**	**-50.62**
	T_{41J}	**-38.20**	**-38.21**	T_{41J}	**-38.20**	**-38.21**	T_{41J}	**-38.20**	**-38.21**	T_{41J}	**-31.27**	**-31.27**	T_{41J}	**-36.95**	**-36.96**
							DBTT	**0.00**	**0.00**				**DBTT**	**0.00**	**0.00**

Data Set 2 (gentle slope, low scatter)

Variable	Symmetric TANH			Asymmetric TANH			Burr			ARCTAN			K-ASYM		
	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app
KV (J)	C	30.14	30.14	C	29.98	29.98	k	0.0608	0.0608	C	20.60	20.60	C	21.90	21.90
	DBTT	**11.27**	**11.27**	**DBTT**	**11.17**	**11.17**	T_0	4.54	4.54	**DBTT**	**11.12**	**11.12**	T_0	11.13	11.13
				D	0.01417	0.01417	m	1.3398	1.3398				p	1.006	1.006
	RMSE	2.231	2.231	RMSE	2.3477	2.3477	RMSE	2.322	2.322	RMSE	4.0843	4.0840	RMSE	2.835	2.835
	T_{28J}	**3.14**	**3.14**	T_{28J}	**3.07**	**3.07**	T_{28J}	**2.86**	**2.86**	T_{28J}	**5.35**	**5.35**	T_{28J}	**4.52**	**4.52**
	T_{41J}	**14.44**	**14.44**	T_{41J}	**14.34**	**14.34**	T_{41J}	**14.15**	**14.15**	T_{41J}	**13.30**	**13.30**	T_{41J}	**13.63**	**13.37**
							DBTT	**10.94**	**10.94**				**DBTT**	**11.20**	**11.20**
LE (mm)	C	33.04	33.04	C	30.07	30.07	k	0.042948		C	23.04	23.04	C	21.37	21.37
	DBTT	**14.56**	**14.56**	**DBTT**	**13.22**	**13.22**	T_0	-186.048		**DBTT**	**13.58**	**13.58**	T_0	-4.31	-4.31
				D	0.1682	0.1682	m	3579.919					p	3.733	3.733
	RMSE	0.059	0.059	RMSE	0.055	0.055	RMSE	0.0497		RMSE	0.088	0.088	RMSE	0.050	0.050
	$T_{0.89mm}$	**22.58**	**22.58**	$T_{0.89mm}$	**21.38**	**21.38**	$T_{0.89mm}$	**21.59**		$T_{0.89mm}$	**19.33**	**19.33**	$T_{0.89mm}$	**20.21**	**20.21**
							DBTT	**13.02**					**DBTT**	**11.06**	**11.06**
SFA (%)	C	35.42	35.47	C	36.20	36.30	k	0.06827	0.06930	C	23.81	23.84	C	24.87	24.89
	DBTT	**17.40**	**17.39**	**DBTT**	**18.01**	**18.04**	T_0	27.75	28.41	**DBTT**	**17.23**	**17.21**	T_0	22.60	22.76
				D	-0.07658	-0.082	m	0.644631	0.6244				p	0.7433	0.736
	RMSE	1.734	1.746	RMSE	1.524	1.491	RMSE	1.542	1.504	RMSE	5.184	5.138	RMSE	2.447	2.389
							DBTT	**18.12**	**18.16**				**DBTT**	**18.68**	**18.70**

Data Set 3 (outliers)

Variable	Symmetric TANH			Asymmetric TANH			Burr			ARCTAN			K-ASYM		
	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app	Parameter	Excel	Shiny app
KV (J)	C	35.24	35.24	C	6.31	7.53	k	23.90508		C	2.18	2.18	C	0.41	1.09
	DBTT	**-87.52**	**-87.52**	**DBTT**	**-82.34**	**-84.59**	T_0	-70.31		**DBTT**	**-75.93**	**-75.93**	T_0	-78.95	-71.83
				D	-0.28684	-0.3422	m	0.01738					p	1.5662	0.0873
	RMSE	140.988	140.988	RMSE	143.214	142.939	RMSE	138.6004		RMSE	130.229	130.229	RMSE	138.603	138.602
	T_{28J}	**-136.28**	**-136.28**	T_{28J}	**-151.01**	**-180.44**	T_{28J}	**-77.12**		T_{28J}	**-83.33**	**-83.33**	T_{28J}	**-79.55**	**-77.36**
	T_{41J}	**-128.57**	**-128.57**	T_{41J}	**-134.71**	**-155.25**	T_{41J}	**-76.14**		T_{41J}	**-80.79**	**-80.79**	T_{41J}	**-79.42**	**-76.54**
							DBTT	**-71.98**					**DBTT**	**-78.84**	**-73.06**

As discussed in Appendix A.1, non-convergence when fitting a regression model to a specific dataset may stem from convergence to a local rather than global minimum, or from the lack of a unique solution. For the three data sets considered here, only the BUR model failed to converge using the Shiny App. However, the same datasets fit using the Excel spreadsheet produce reasonable results for all models. Caution is advised when using results from the Excel spreadsheet if convergence is not also achieved for the Shiny App; standard errors of the parameter estimates and diagnostic plots are not available in the Excel spreadsheet, so the quality of the parameter estimates cannot be evaluated.

For the BUR model with fixed shelves, convergence may depend on the steepness of the slope in the transition region as well as the amount of scatter in the transition region data. The BUR model converged using the Shiny App for all data sets when both shelves are variable.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

7 Conclusions

We developed at NIST a software package to obtain transition curves for Charpy and toughness results as a function of test temperature. The package includes an Excel macro-enabled spreadsheet for a basic-level analysis and a web-based Shiny App for a more advanced-level analysis. Both programs perform non-linear data regression according to five fitting models identified from a literature review. summarizes the features available in the Excel spreadsheet and the Shiny App.

Table 4 - Features available in the Excel spreadsheet and in the Shiny App.

Program	Parameter Estimates	RMSE	AIC	BIC	Characteristic Temperatures	Standard Errors	Diagnostic Plots	Confidence Bounds
Excel spreadsheet	√	√	√	√	√	X	X	X
Shiny App	√	√	√	√	√	√	√	√

The package also includes another macro-enabled Excel spreadsheet that allows determining reasonable initial values for the three asymmetric models (AHT, BUR, and KHT), which are needed in order to increase the likelihood of obtaining reliable fits.

The NIST software package is demonstrated for two experimental Charpy datasets consisting of absorbed energy values. Additionally, a simulated dataset, generated based on specified parameters for the HT model, is analyzed to verify that the software returns parameter estimates that are very close to the specified model parameters for the HT model fit. Although only KV data are analyzed here, other responses such as LE and SFA can also be analyzed using the NIST software package.

The output of the Excel spreadsheet and of the Shiny App are compared, and large differences in calculated temperatures (greater than 5 °C) are only observed for the dataset containing outliers. In general, either software appears adequate for a basic-level data analysis.

This report recommends fixing both upper and lower shelves. Specifically, lower shelf should be set to a physically meaningful value considering the material being tested (*e.g.*, LS = 2 J for KV values, and LS = 0 mm for LE values). Upper shelf values should be set equal to the average of specimens corresponding to SFA ≥ 95 % or SFA = 100 %. In the case of Shear Fracture Appearance, shelves must be set to 0 % and 100 %, respectively. The Shiny App allows the user to choose between fixed or variable shelves, while the Excel spreadsheet only performs analyses where both shelves are fixed.

Finally, fitting the Burr (BUR) model using the Shiny App can be difficult for the fixed-shelves case if the slope of the transition region is very steep and/or if the scatter in the data is relatively large. In these cases, the Shiny App might be able to produce a good fit if variable shelves are used instead of fixed shelves. The Excel spreadsheet does not seem to have the same difficulty fitting the BUR model.

We believe this freely available software package developed at NIST will be useful both for a basic-level analysis, which only requires transition curves and some characteristic temperatures (such as DBTT), and for a more advanced data treatment, aiming at establishing

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

confidence bounds for the regression curves and standard errors/uncertainties on the calculated parameters.

References

[1] Microsoft Corporation. *Microsoft Excel* [Internet]. 2002. Available from: https://office.microsoft.com/excel.

[2] Chang W, Cheng J, Allaire J, Xie Y, McPherson J (2020) shiny: Web Application Framework for R. R package version 1.5.0. https://CRAN.R-project.org/package=shiny

[3] R Core Team (2018) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (https://www.R-project.org/).

[4] Mouritz AP (2012) *Introduction to aerospace materials* (Woodhead Publishing Limited).

[5] Jordan J (2016) *Ductile to brittle transitions in materials.* https://www.jeremyjordan.me/ductile-to-brittle-transitions-in-materials/ (retrieved on February 1, 2021).

[6] Russell SB (1898) Experiments with a New Machine for Testing Materials by Impact. *Transactions ASCE*, 39:237–250.

[7] Charpy G (1901) Note sur l'essai des metaux à la flexion par choc de barreaux. *Soc. Ing. Civ. de Français*, pp 848–877.

[8] International Atomic Energy Agency, IAEA (2009) Integrity of Reactor Pressure Vessels in Nuclear Power Plants: Assessment of Irradiation Embrittlement Effects in Reactor Pressure Vessel Steels. IAEA Nuclear Energy Series No. NP-T-3.11, Vienna.

[9] Oldfield W (1975) Curve fitting impact test data: a statistical procedure. *Standardization News*, Vol. 3, No. 11, pp. 24-29.

[10] Oldfield W (1979) Fitting Curves to Toughness Data. *Journal of Testing and Evaluation*, JTEVA, Vol. 7, No. 6, pp. 326-333.

[11] Wallin KRW (2011) *Fracture Toughness of Engineering Materials*. EMAS Publishing, UK.

[12] Burr IW (1942) Cumulative frequency functions. *Annals of Mathematical Statistics*, Vol. 13, No. 2, pp. 215–232.

[13] Windle PL, Crowder M, Moskovic R (1996) A statistical model for the analysis and prediction of the effect of neutron irradiation on Charpy impact energy curves. *Nuclear Engineering* Design, Vol. 165, No. 1-2, pp. 43-56.

[14] Cao LW, Wu SJ, Flewitt PEJ (2012) Comparison of ductile-to-brittle transition curve fitting approaches. *International Journal of Pressure Vessels and Piping*, Vol. 93-94, pp. 12-16.

[15] Kohout J (2012) Various Regression Functions Fitting Transition Curves. Internal Report for the Ministry of Defense of the Czech Republic.

[16] Akselsen OM, Grong Ø (1992) Prediction of weld metal Charpy V notch toughness. *Mater Sci Eng* A, Vol. 159, No. 2, pp.187-192.

[17] EricksonKirk MA, EricksonKirk MT, Rosinski S, Spanner J (2009) A Comparison of the tanh and Exponential Fitting Methods for Charpy V-Notch Energy Data. *Journal of Pressure Vessel Technology*, Vol. 131, pp. 031404-1–031404-13.

[18] Pontikis V, Gorse D (2009) Absorbed impact energy and mode of fracture: a statistical description of the micro-structural dispersion. *Journal of Nuclear Materials*, Vol. 394,

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

No. 1, pp. 109-113.
[19] Todinov MT (1999) Fitting impact fracture toughness data in the transition region. *Materials Science Engineering A*, Vol. 265, No.1-2, pp. 1-6.
[20] Todinov MT (2001) Estimating the parameters of the impact energy variation in the ductile/brittle transition region from complete and sparse data. *Computational Materials Science*, Vol. 21, No. 1, pp. 111-123.
[21] ASTM E185-16 (2016) Standard Practice for Design of Surveillance Programs for Light-Water Moderated Nuclear Power Reactor Vessels. ASTM International, West Conshohocken, PA.
[22] ASTM E2215-19 (2019) Standard Practice for Evaluation of Surveillance Capsules from Light-Water Moderated Nuclear Power Reactor Vessels. ASTM International, West Conshohocken, PA.
[23] Burnham KP, Anderson DR (2004) Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods and Research*, Vol. 33, No. 2, pp. 261-304.
[24] Hastie T, Tibshirani R, Friedman J (2001) Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York.
[25] Shapiro SS, Wilk MB (1965) An Analysis of Variance Test for Normality (Complete Samples), *Biometrika*, Vol. 52, No. 3/4, pp. 591-611.
[26] Elzhov TV, Mullen KM, Spiess A-N Bolker B. (2016) minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds, R package version 1.2-1 (https://CRAN.R-project.org/package=minpack.lm).
[27] Marquardt D (1963) An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *SIAM Journal on Applied Mathematics,* Vol. 11, No. 2, pp. 431-441.
[28] Banks HT, Joyner ML (2017) AIC under the framework of least squares estimation, *Applied Mathematics Letters*, Vol. 74, pp. 33-45.
[29] Fox J (2016) Applied Regression Analysis and Generalized Linear Models, 3rd Edition, Sage Publications, Inc., Los Angeles, CA.
[30] Johnson N, Kotz S, Balakrishnan N (1994) Continuous Univariate Distributions, Second Edition, John Wiley & Sons, Inc., Hoboken, NJ.
[31] Graybill FA, Iyer H K (1994) Regression Analysis: Concepts and Applications, Duxbury Press, Boston, MA.
[32] Riazoshams AH, Habshah Jr. BM, Adam CMB (2009) On the Outlier Detection in Nonlinear Regression, *International Journal of Mathematical and Computational Sciences*, Vol. 3, No. 12, pp. 243-250.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

Appendix A: Computational and Statistical Methods

The software is offered in the form of its web-based interface, which is generated using the R Shiny package (version 1.5.0) [2]. The software uses the nonlinear least-squares function, *nls.lm*, from the R library *minipack.lm* [26] to fit various models. This function uses the Levenberg-Marquardt optimization method [27] to estimate model parameters and allows for the specification of lower and upper bounds for such parameters.

A.1 Convergence

Not all the models considered (HT, AHT, BUR, ACT, and KHT) will provide a good fit to the input data. As with any numerical optimization routine, convergence to the optimal set of parameters for the given dataset is not guaranteed and is highly dependent upon the landscape of the underlying objective function. The objective function may be non-convex, ill-conditioned, or underdetermined. These issues tend to be exacerbated when the data size is small, the underlying model doesn't fit the data well, or the underlying model is overly complex. Additionally, successful performance of the optimization algorithm is often contingent upon choosing initial parameter values that aren't too "far away" from the optimal solution. The Shiny App will not return a solution if the optimization algorithm fails to converge. However, even if the application does not report any convergence issues, the user should carefully inspect the diagnostics presented within the application to ensure the algorithm has not settled onto an unacceptable stationary point. In most cases, this will be immediately obvious from the plot showing the fitted models.

A.2 Model Selection Statistics

To compare various model fits, we provide the RMSE from the fit as well as two statistics called the Akaike Information Criterion (AIC) [23] and the Bayesian Information Criterion (BIC) [24]. The RMSE, AIC, and BIC statistics are relative measures, so there is no absolute threshold that indicates whether the model provides a good fit to the data. The RMSE, AIC, and BIC statistics simply allow models to be compared to each other; the model with the smallest values of RMSE, AIC, and BIC is preferred. RMSE, AIC, and BIC are defined as:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n-p}} = \sqrt{\frac{\mathrm{RSS}}{n-p}}, \tag{A.1}$$

$$\mathrm{AIC} = 2(p+1) - 2\ln\left(\hat{L}\right), \tag{A.2}$$

$$\mathrm{BIC} = (p+1)\ln(n) - 2\ln\left(\hat{L}\right), \tag{A.3}$$

where $\hat{Y}_i$ is the value of the response at X_i predicted by the regression equation, p is the number of parameters estimated by the regression analysis (excluding σ, the residual standard deviation), n is the number of data points, and $\hat{L}$ is the maximum value of the log-likelihood function. For least-squares estimation [28], the AIC can be written in terms of the residual sum of squares (RSS) as

$$AIC = 2(p+1) + n \ln\left(\frac{RSS}{n}\right). \quad (A.4)$$

Similarly, BIC can be written as

$$BIC = (p+1)\ln(n) + n\, ln\left(\frac{RSS}{n}\right). \quad (A.5)$$

A.3 Temperature Variability

As in least-squares regression, nonlinear least-square regression assumes that each independent variable (test temperature, in this case) is known without error. In reality, temperatures are measured with error, but we assume that such errors are small relative to the errors in the response variable (KV, LE, or SFA). We also assume that the variability in temperature will not substantially influence the results.

A.4 Finding Standard Errors: Parametric Bootstrap Analysis

The R code estimates the standard errors of the parameter estimates using a parametric bootstrap method. The bootstrap procedure involves generating samples of residuals for each value of observed temperature, based on the parameter estimates and RMSE from the original fit [29]. For example, consider a straight line fit:

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i \ . \quad (A.6)$$

The b^{th} bootstrap sample ($b = 1, \ldots, n_b$) for X_i is given by:

$$Y_{bi} = \hat{\alpha} + \hat{\beta} X_i + e_{bi} = \hat{Y}_i + e_{bi} \ , \quad (A.7)$$

where e_{bi} is a random draw from a normal distribution with a mean of zero and a standard deviation equal to the RMSE from the original fit. A straight line is fit to each bootstrap sample and the n_b sets of parameter/coefficient estimates are saved (along with predicted values at each observed temperature) for further analysis.

For our data, all bootstrapped sample values must be greater than zero, so a truncated normal distribution is used to generate data. For SFA, the normal distribution is truncated at 0 % on the left and at 100 % on the right.

The number of bootstrap iterations in the Shiny App is set to be 1000 for each model. This choice is based on the results of numerical simulations investigating the number of iterations needed for the bootstrap interval estimates to stabilize.

For the case of one or more fixed shelves, the bootstrap samples are generated based on the original fitted model parameters along with the original fixed shelves. However, the fixed-shelf values specified for fitting each bootstrapped sample are generated randomly. The bootstrap procedure incorporates variability of all parameters included in the model, as well as variability in the fixed shelves.

The standard errors provided by the Shiny App represent sources of uncertainty due to random effects. The Shiny App does not address sources of uncertainty due to systematic effects.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

A.5 Uncertainty of Fixed Shelves

In case of variable shelves, the R code only requires initial values of LS and/or US. For models with one or both fixed shelves, values for the shelves must be specified, along with the uncertainty of the specified shelf. The fixed-shelf values and their uncertainties are used to define the distributions for generating random shelf values for the parametric bootstrap method.

A fixed-shelf uncertainty may be estimated based on the standard deviation of the observed shelf data (type A uncertainty evaluation) or based on engineering judgement (type B uncertainty evaluation) if data are insufficient to estimate the standard deviation.

For KV, LE, and "Other" response variables, random US values are generated using a normal distribution with mean equal to the user-specified fixed upper-shelf value and standard deviation equal to the user-specified uncertainty of the fixed upper shelf, N(US, u(US)). The default value of u(US) is 5 % of the US value for both KV and LE responses, based on engineering judgement. The default values may not be realistic for all data sets.

For KV, LE, and "Other" variables, the randomly generated LS values cannot be less than zero (that would be non-physical). Thus, the LS values are generated using a truncated normal distribution that is truncated at zero with mean equal to the user-specified fixed lower-shelf value, LS, and standard deviation equal to the user-specified uncertainty of the fixed lower shelf, u(LS). Information regarding truncated normal distributions can be found in [30]. The default value of u(LS) is 0.3 J for the KV response and is 0.03 mm for LE response, based on engineering judgement. The default values may not be realistic for all data sets.

The user may specify u(US) = 0 to obtain results for the fixed-shelves case where there is no error in the upper shelf. Similarly, the user may specify u(LS) = 0 for the case where there is no error in the fixed lower shelf. However, fixing shelves without error is not recommended for KV and LE responses.

A.6 Confidence Intervals and Confidence Bounds

All confidence intervals in the Shiny App are computed using the percentiles of the bootstrap distribution implied by the user's selected confidence level (such that each tail probability is the same). For example, if the user selects a 95 % confidence level, the application will compute the 2.5 % and 97.5 % percentiles of the simulated bootstrap distribution. The choice of using the percentile bootstrap ensures that confidence intervals are robust when the fitted parameters lie near their boundary (e.g. upper and lower shelves).

A.7 Diagnostic Plots

The following four diagnostic plots (see **Fig. 17**) are provided to display information regarding the quality of the fit and to expose assumption violations:

1. Residual versus predicted value.
2. Standardized residual versus temperature.
3. Residual lag plot (residual[i+1] versus residual[i]).
4. Normal quantile-quantile plot (Q-Q plot) of residuals.

Residual analysis is needed to check model adequacy and to verify that regression assumptions are not violated. (See [31] section 3.5 for a discussion of residual analysis.)

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

Standardized residuals are computed using the formulas provided by Riazoshams [32] for nonlinear models. However, residuals are called "studentized" instead of "standardized" in [32]. The term "studentized" generally implies that the normalizing variance is computed by leaving out the data point for which the residual is being calculated. Since the standardized residuals produced by the Shiny App are not computed using a "leave-one-out" type of procedure, the term "standardized" is more correct in this context.

The Shiny App uses a Shapiro-Wilk test [25] to test the assumption that the residuals from the fit are normally distributed.

A.8 DBTT Calculation

For the HT, AHT, and ACT models, $DBTT = T_0$. Therefore, the standard error of $DBTT$ for these models is provided by the fitting procedure.

For the BUR model, $DBTT$ is computed using:

$$DBTT = T_0 - \frac{1}{k}\ln\left(\left(\frac{1}{2}\right)^{-1/m} - 1\right). \tag{A.8}$$

For the KHT model, $DBTT$ is computed from:

$$DBTT_1 = T_0 + \frac{2C}{1+p}\ln\left(\frac{1+p}{2}\right) \tag{A.9}$$

and

$$DBTT_2 = T_0 - \frac{2Cp}{1+p}\ln\left(\frac{1+p}{2p}\right)\ , \tag{A.10}$$

where $DBTT = DBTT_1$ if $(DBTT_1 + DBTT_2)/2 \leq T_0$, and $DBTT = DBTT_2$ otherwise.

The standard error of $DBTT$ for the BUR and KHT models is computed from the parametric bootstrap.

A.9 Calculation of Characteristic Temperatures

The Shiny App also provides characteristic temperature estimates and their associated uncertainties for specific values of KV, LE, or SFA. Characteristic temperature estimates are obtained by inverting each model, so that temperature becomes a function of the response variable. The bootstrapped parameter estimates (and randomly generated shelves for models with one or more fixed shelves) are used to compute standard error estimates for each selected response value (Y). The equations used to compute the temperatures are provided in **Table A.1**.

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158

Table A.1 - Equations used to calculate characteristic temperatures.

Model	Equation
HT	$T_Y = T_0 + \frac{C}{2}\ln\left(\frac{Y - LS}{US - Y}\right)$
AHT	$T_Y = \frac{T_0 + \frac{C}{2}\ln\left(\frac{Y - LS}{US - Y}\right)}{1 - \frac{D}{2}\ln\left(\frac{\bar{Y} - LS}{US - Y}\right)}$
BUR	$T_Y = T_0 - \frac{1}{k}\ln\left[\left(\frac{Y - LS}{US - Y}\right)^{-\frac{1}{m}} - 1\right]$
ACT	$T_Y = DBTT + \frac{2C}{\pi}\tan\left(\frac{\frac{\pi}{2}}{US - LS}\right)[2Y - (LS + US)]$
KHT	$T_Y = T_0 + \frac{2C}{1+p}\ln\left[\frac{(1+p)(Y - LS)}{US - LS}\right]$ for $Y \leq \frac{US + p \cdot LS}{1+p}$ $T_Y = T_0 - \frac{2Cp}{1+p}\ln\left[\frac{(1+p)(US - Y)}{p(US - LS)}\right]$ for $Y > \frac{US + p \cdot LS}{1+p}$

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.2158