Solid State Phenomena
ISSN: 1662-9779, Vol. 390, pp 121-135
doi:10.4028/p-oDeJ6s

Submitted: 2026-01-19
Revised: 2026-02-24
Accepted: 2026-02-25
Online: 2026-04-17

Predictive Modeling of 42CrMo4 Hardness after Two-Step Heat Treatment Using Symbolic Regression

ALBERS Johann[1,a*], DITTRICH Fabian[1,b], and HÄRTEL Sebastian[1,c]

Fachgebiet Hybride Fertigung, BTU, Cottbus-Senftenberg 03046, Germany

[a]johann.albers@b-tu.de, [b]fabian.dittrich@b-tu.de, [c]sebastian.haertel@b-tu.de

Keywords: 42CrMo4, Heat Treatment, Material Hardness, Symbolic Regression, Predictive Mode

Abstract. During a heat treatment, a material undergoes microstructural changes that result in an alteration of its hardness. In a two-step heat treatment, the material is first adjusted to an initial hardness via a specified cooling rate. Subsequently, the hardness is reduced through a tempering process, while its ductility is increased. Depending on the tempering duration, tempering temperature, and initial hardness, different resulting hardness values are obtained. The resulting hardness after a chosen heat treatment is thus far been difficult to predict. This work employs symbolic regression to develop a model that predicts the hardness evolution of 42CrMo4 steel as a function of cooling rate, tempering duration, and tempering temperature. By describing the model with few parameters, it has also been demonstrated that cooling rates and tempering temperatures leading to a target hardness can be determined. The overall model achieves a coefficient of determination of $R^2 = 98.50\,\%$ for known experimental data and a combined coefficient of determination of $R^2 = 93.13\,\%$ for previously unknown cooling rates (forward) and previously unattained resulting hardness values (inverse). Our work shows that the resulting hardness of 42CrMo4 can be predicted using a small number of parameters. This work is anticipated to establish a foundation for further research endeavors. For instance, the approach using symbolic regression can be further adapted to identify physically interpretable constants. Furthermore, the model description offers the possibility of coupling with a simulation model to accurately predict the hardness of a component.

Introduction

Objective

This study aims to model the hardness of 42CrMo4 (see Table 1) steel following a two-step heat treatment comprising hardening and tempering. Although various models exist [1, 2, 3, 4], predicting final hardness remains challenging due to the complexity of the underlying hardening mechanisms [2, 3]. Recent advancements in machine learning offer new methodologies to address such problems, even with limited datasets [5, 6].

Table 1: Chemical Composition of 42CrMo4 (wt. %).

	Specified	**Actual**
C	$0.38 - 0.45$	0.42
Mn	$0.60 - 0.90$	0.72
Si	< 0.40	0.40
Cr	$0.90 - 1.20$	1.05
Mo	$0.15 - 0.30$	0.29
Fe	bal.	bal.

Experimental data were obtained in two stages: first, hardening by austenitizing the samples at $850\,°C$ for $30\,\text{min}$ and cooling at four rates ($0.6\,\frac{K}{s}$, $4.0\,\frac{K}{s}$, $12.5\,\frac{K}{s}$, and $76.0\,\frac{K}{s}$), and second, tempering at different temperatures ($400\,°C$, $525\,°C$, $650\,°C$) and durations ($30\,\text{min}$, $60\,\text{min}$, $120\,\text{min}$) (see Figure 1 for an example of the heat treatment procedure). Symbolic Regression [6] was employed to model the hardening and tempering behavior, fitting the experimental data while enforcing parsimony to minimize model complexity and mitigate overfitting. First, a model was developed correlating the untreated hardness and cooling rate with the hardness after the hardening step. A subsequent model was then derived to predict the final hardness based the tempering duration, tempering temperature and this initial hardness. These models were combined into a single, consolidated model consistent with the experimental data. The model's predictive capability was validated by forecasting hardness values for unseen data. Furthermore, an objective function was formulated to optimize the process parameters to achieve a target hardness, thereby validating the model's invertibility using additional experimental data.

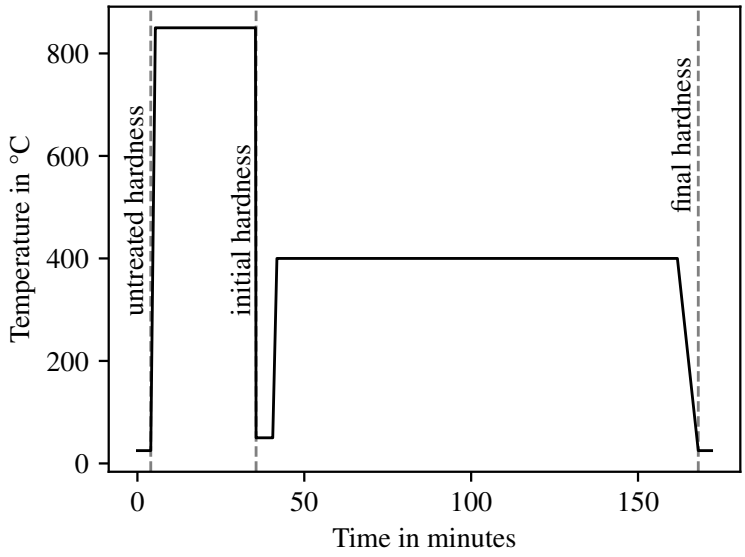

Fig. 1: An exemplary heat treatment of a 42CrMo4 sample involved heating at $10.0\,\frac{K}{s}$ to $850\,°C$, holding for $30\,\text{min}$ for normalization, followed by rapid cooling at $76.0\,\frac{K}{s}$ to $50\,°C$. This established the initial hardness and concluded the hardening step. A subsequent tempering step was performed by heating at $5.0\,\frac{K}{s}$ to $400\,°C$, holding for $120\,\text{min}$, and then cooling at $1.0\,\frac{K}{s}$ to $25\,°C$, which established the final hardness and concluded the tempering step.

Justification

The need for this study arises from the persistent gap in understanding the underlying mechanisms that determine a material's final hardness with respect to its heat treatment [7]. Despite various models in the materials engineering literature [1, 2, 4], the problem of predicting final hardness from a two-step process of hardening and tempering remains unsolved. Addressing these open problems helps in furthering the understanding of hardness evolution in 42CrMo4. Previous research has focused on empirical Jominy hardenability correlations [8] and physics-based models simulating phase transformations [9] and carbide precipitation [10], but still does not solve the problem of creating a robust, generalizable, and inverse model that accurately predicts final hardness [11] across a wide range of combined hardening and tempering parameters. Consequently, this study aims to decrease the gap between high-fidelity modeling and practical application by using techniques from recent developments in machine learning. The relevance of this study is underscored by its potential adaptability to other materials besides the 42CrMo4 investigated here. Predicting the resulting hardness from heat treatment further allows for the optimization of process parameters to yield more precise results with lower

input resources. Furthermore, the present work aligns with the broader objective of further investigating the hardening capabilities of steel alloys. By rigorously investigating the hardening behavior of 42CrMo4, this work helps to further the understanding of hardening mechanisms and opens a discourse for further research.

Related Work

Heat treatment

This work investigates the steel alloy 42CrMo4. 42CrMo4 belongs to the group of tool steels or quenched and tempered steels [12]. Tool steels are particularly suitable for hardening due to their medium carbon content and alloying elements like chromium and molybdenum, which facilitate the formation of a martensitic microstructure upon rapid cooling from the austenitizing temperature [13]. A common heat treatment for 42CrMo4 involves hardening followed by tempering [13]. The initial condition of 42CrMo4 possesses a certain base hardness, which is significantly increased through hardening. This is due to the transformation of the microstructure at elevated temperatures. Upon exceeding a specific temperature (the austenitization temperature), the microstructure transforms into austenite [14]. Subsequent rapid cooling transforms this austenite into martensite, a phase characterized by a significantly increased hardness. Alongside the increase in hardness, however, the ductility decreases, leading to a lower fracture strain [13]. To counteract this embrittlement, the material is subjected to a subsequent tempering step. Depending on the tempering temperature and tempering duration, the hardness decreases with a simultaneous increase in ductility. Tool steels are typically heated to their austenitization temperature, usually within $820\,^{\circ}\mathrm{C}$ to $880\,^{\circ}\mathrm{C}$. In this study, cooling rates from $0.6\,\frac{\mathrm{K}}{\mathrm{s}}$ to $76.0\,\frac{\mathrm{K}}{\mathrm{s}}$ were employed to produce hardness levels of about $480\,\mathrm{HV1}$ to $660\,\mathrm{HV1}$. The tempering temperatures are typically in the range of $300\,^{\circ}\mathrm{C}$ to $600\,^{\circ}\mathrm{C}$ with a tempering duration between 60 minutes and 120 minutes [15]. Upon completion of the heat treatment, the microstructure of the steel alloy is permanently altered, resulting in an increased hardness compared to the initial state.

Material modeling

Modeling a material's behavior is complex due to the multitude of underlying mechanisms [16]. While simplified models exist, they often rely on idealized conditions. The prediction of hardness after a two-step heat treatment (hardening and tempering) exemplifies this challenge. Several models in the literature attempt to predict final hardness based on process parameters. For instance, [17] established a classic tempering parameter (the H-parameter) that relates hardness to tempering temperature and time. However, such models are typically calibrated for a specific initial condition (e.g., a fixed as-quenched hardness) and do not inherently account for variations in the initial hardening step, such as different cooling rates. An alternative approach by [5] uses neural networks to predict hardness, incorporating parameters like tempering temperature and time. While powerful, these data-driven models often require large datasets and function as "black boxes," lacking the parsimony and physical interpretability desired for fundamental understanding. The model developed in this work differs by simultaneously addressing both heat treatment steps. It explicitly links the final hardness to the cooling rate from the hardening step, the resulting initial hardness, the tempering duration, and the tempering temperature. This integrated approach, achieved through Symbolic Regression [6], provides a consolidated and invertible model that captures the combined effect of these key, limited input parameters (cooling rate, tempering duration, tempering temperature), filling a gap left by models that treat the hardening and tempering steps in isolation.

Symbolic regression

To identify a model that approximates the experimental data, the approach of Symbolic Regression is investigated. Symbolic Regression searches the space of mathematical functions for candidates that optimally approximate the prescribed data. The methodology for finding such an optimum follows different strategies. In [18], this optimum is found using genetic programming. Conversely, [19] demon-

strates finding an optimum using a method based on sparse regression. To further constrain the optimization problem, various limitations can be imposed, such as restricting the set of usable symbols or the degree of complexity (the length of the mathematical expression). Ready-made software packages exist for each of these approaches, implementing the different optimization strategies. To evaluate an optimum, it is necessary to define an objective function. Consider the following minimization problem as an example: Given a dataset $\{(x_i, z_i)\}_{i=1}^{n}$ with $x_i \in \mathbb{R}^d$ and $z_i \in \mathbb{R}$ Symbolic Regression seeks a function f that minimizes, for instance, the sum of squared differences.

$$\min_{f \in F} \sum_{i=1}^{n} (f(x_i) - z_i)^2 \tag{1}$$

where F is the space of mathematical functions. This function space can be constrained by permitting only a subset of unary (e.g., $\sqrt{\circ}, log(\circ), \circ^2, |\circ|$) and binary operators (e.g., $+, -, \times, /$). This work employs the approach by [6], which leverages probabilistic and interpretable machine learning to discover symbolic models. The methodology, often implemented using PySR (Python Symbolic Regression) [6], builds upon principles of genetic programming while incorporating techniques for improved robustness and scalability. Further details on the implementation and documentation can be found at the library's official repository [6].

Methodology

Ground truth

A series of heat treatments were first conducted using the DIL805 A/D/T deformation dilatometer from TA Instruments to generate a dataset from which the mathematical model is developed. Cooling rates of $0.6\,\frac{\mathrm{K}}{\mathrm{s}}$, $4.0\,\frac{\mathrm{K}}{\mathrm{s}}$, $12.5\,\frac{\mathrm{K}}{\mathrm{s}}$, and $76.0\,\frac{\mathrm{K}}{\mathrm{s}}$ were employed to achieve a range of initial hardness values. Subsequently, the samples are tempered at different tempering temperatures and for a specified tempering duration. Figure 1 shows the temperature-time curve for a sample cooled at a rate of $76\,\frac{\mathrm{K}}{\mathrm{s}}$. Each sample is subsequently cut, and the hardness is measured using the ZwickRoell DuraScan 70 G5. The hardness values reported in this work were measured according to the Vickers hardness scale using a test force of 1 kilogram $-$ force according to HV1. The Vickers hardness testing method was specifically developed for measuring the hardness of a wide range of materials, including metals, by assessing the resistance to plastic deformation from a diamond pyramid indenter [20]. The measured hardness values were determined at different positions, and their standard deviation is also reported. After evaluating all hardness measurements, the data are summarized in a table to serve as a basis for further analysis. Further quantities are derived from the measured values. Among other things, the hardness difference between the initial hardness and the final hardness is determined for each data point. Additionally, a corrected tempering duration is calculated, which additively comprises the nominal tempering duration and the duration derived from the heating and cooling rates during the tempering process. The quotient of the hardness difference and the corrected tempering duration yields the rate of hardness change, $\frac{dHV}{dt}$.

Parsimony

The history of science shows that many fundamental laws can be described, to a first approximation, by comparatively simple relationships [21]. For example, Hooke formulated his law of elasticity based on simple proportional relationships [22]. Similarly, Newton was able to derive a simple description of universal gravitation building upon the work of Kepler [23, 24, 25]. These are examples that demonstrate how a pure observation of data has led to the discovery of fundamental laws. The parsimony of equations naturally counteracts overfitting to the data. Through this parsimony, one can inherently counteract the inherent noise in the data. The trade-off between model complexity and model accuracy is illustrated in the example of Figure 2. Particularly with limited data, it is crucial to ensure that the model maintains a certain minimum level of parsimony, as a model with too many parameters can easily fit the entire dataset perfectly, leading to overfitting.

Symbolic regression

The software package used to perform the Symbolic Regression is PySR. The core algorithm was implemented in the Julia language and ported to Python. Julia offers the advantage of being a compiled language, where the code is compiled into a kernel that executes more efficiently compared to Python. In this work, one model is identified for the initial hardness and a separate model for the final hardness. Subsequently, both models are executed in series and compared with the experimental data. PySR offers various options for configuring the search criteria for an optimal equation. For instance, the selection of operators can be constrained, as well as the number of search iterations, the maximum model size (a measure of complexity), and the loss function. By default, the mean squared error is used as the loss function, but a user-defined loss function can also be specified as a Julia function. In the subsequent procedure, a custom loss function is defined to incorporate previously known knowledge about the behavior of hardness. The selection of hyperparameters, such as unary operators, binary operators, population size, the number of iterations, and the target loss, is mentioned in terms of their choice and specification, but will not be discussed in further detail.

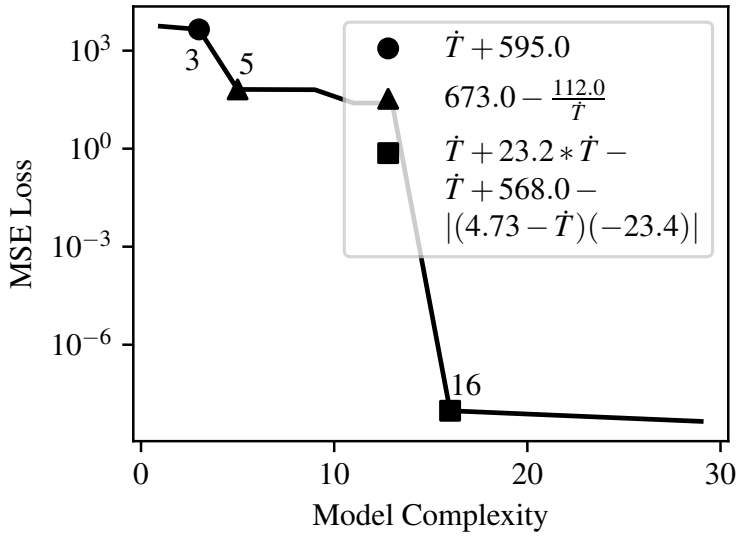

Fig. 2: Pareto front for the search for the first-step hardening model. Symbolic Regression identifies candidate fits through a trade-off between loss and model complexity. While high complexity can reduce the loss significantly—note the logarithmic scale of the vertical axis—it increases the risk of overfitting, particularly on a small dataset.

Experiments

Hardening

To obtain a suitable mathematical description of the resulting hardness, heat treatments were initially performed on cylindrical 42CrMo4 samples with a base hardness of $334.9\,\mathrm{HV1} \pm 5.28\,\mathrm{HV1}$ and dimensions of $10\,\mathrm{mm} \times 2.5\,\mathrm{mm}$ (length × radius). The parameters for the heat treatment are listed in Table 2. For the model description, four different cooling rates were initially selected to achieve an initial hardness. These considered cooling rates are $0.6\,\frac{\mathrm{K}}{\mathrm{s}}$, $4.0\,\frac{\mathrm{K}}{\mathrm{s}}$, $12.0\,\frac{\mathrm{K}}{\mathrm{s}}$, and $76.0\,\frac{\mathrm{K}}{\mathrm{s}}$. Originally, an attempt was made to achieve a maximum cooling rate of $95.0\,\frac{\mathrm{K}}{\mathrm{s}}$; however, the dilatometer used only achieved a cooling rate of approximately $76.0\,\frac{\mathrm{K}}{\mathrm{s}}$. The cooling rate of $76.0\,\frac{\mathrm{K}}{\mathrm{s}}$ is derived from the $t_{8/5}$ time, thus corresponding to the cooling rate from $800\,^{\circ}\mathrm{C}$ to $500\,^{\circ}\mathrm{C}$. After successful heat treatment, the samples were cut along their central axis using a cutting wheel and then mounted. The mounted samples were ground in successive steps down to a grit size appropriate for hardness testing. Subsequently, on the cross-section of each sample, 3 series of measurements, each consisting of 5 points,

were performed to determine the hardness. Measurement series 1 and 3 are located near the outer surface of the sample, while measurement series 2 is located along the central axis of the cylindrical sample. This work considers only the results from measurement series 3. Mean values and a standard deviation were calculated from the 5 measurement points along the individual series. Table 3 shows the initial hardness values for the respective cooling rates. These initial hardness values are plotted in Figure 4, which also depicts the saturation of initial hardness starting from $10\,\frac{\mathrm{K}}{\mathrm{s}}$.

Tempering

The heat treatment investigated in this work involves a second step in addition to hardening, namely tempering. Tempering reduces hardness while simultaneously increasing ductility. The investigations include different tempering durations between $30\,\mathrm{min}$ and $120\,\mathrm{min}$. To study the change in hardness at short times, the hardness values at a cooling rate of $\dot{T} = 12.5\,\mathrm{K/s}$ and at times of $3.75\,\mathrm{min}$ ($225\,\mathrm{s}$), $7.3\,\mathrm{min}$ ($440\,\mathrm{s}$), $11.26\,\mathrm{min}$ ($676\,\mathrm{s}$), and $15\,\mathrm{min}$ were also examined. The tempering temperatures also vary between $400\,^{\circ}\mathrm{C}$ and $650\,^{\circ}\mathrm{C}$. Table 2 lists all results for the hardness values and standard deviations for the investigated parameters.

Augmentation

In addition to the data points from the experiments, further result values were introduced. First, the change in hardness at each time point compared to the initial hardness was determined. Subsequently, the tempering duration was considered as the time difference. This time difference was then corrected for the heating duration and the cooling duration. This correction is based on the duration required to reach the tempering temperatures and to cool down from them, compare Figure 1. The quotient of the hardness change and the time change yields a quantity that is used for the subsequent model formation. In Figure 3, this quotient is plotted against the corrected time. To further constrain the model, artificial data points were introduced. Firstly, for each cooling rate and each tempering temperature, the hardness was kept constant below a tempering temperature of $200\,^{\circ}\mathrm{C}$ This is part of the data augmentation and thus expert knowledge to restrict the model behavior to zero change in hardness below $200\,^{\circ}\mathrm{C}$. It effectively guides the model such that a significant hardness change only occurs beyond a specific temperature threshold. The value of $200\,^{\circ}\mathrm{C}$ was determined based on empirical data. Secondly, for each cooling rate and tempering temperature, an additional data point was introduced to promote a constant hardness change beyond $120\,\mathrm{min}$. This was achieved by duplicating each data row at $120\,\mathrm{min}$ and changing its time to $240\,\mathrm{min}$. Furthermore, for the curve with a cooling rate of $\dot{T} = 76.0\,\frac{\mathrm{K}}{\mathrm{s}}$ at $650\,^{\circ}\mathrm{C}$, additional experiments investigating the short-time behavior were carried out. Each data point is time-corrected by the amount of time required to heat and cool the sample to the desired tempering temperature. Consequently, data points corresponding to a higher tempering temperature were corrected by a larger time increment.

Results

Hardening model

To describe a comprehensive model, two separate models were developed in the following and subsequently executed in two steps. From preliminary considerations, it is assumed that the initial hardness depends solely on the cooling rate, since the homogenization temperature is the same for all experiments. The data basis for modeling the initial hardness essentially consists of four points. These represent the initial hardness at different cooling rates. To find a solution using symbolic regression, various parameters were tested. The parameters used were derived from numerous trials and were selected based on their characteristics, such as asymptotic behavior with increasing time. For the initial hardness model, the unary operators $log(\circ), \circ^2, |\circ|, e^{\circ}$, and the binary operators $+, -, \times, /$ were permitted. The population size was set to a multiple of the number of cores. The machine this computation took place used 16 cores to run the symbolic regression, so a population size of $3 * 16 = 48$ was used. The number of iterations varied between 50 and 500. For the target parsimony, a value of 1×10^{-6} were chosen. Care was taken to ensure that the selected model found a balance between low

Table 2: Final hardness values.

Cooling rate	Temperature	Time	Hardness (HV1)
0.6 K/s	400°C	30 min	394.4 ± 16.1 HV
0.6 K/s	400°C	60 min	423.4 ± 6.2 HV
0.6 K/s	400°C	120 min	424.6 ± 10.2 HV
0.6 K/s	525°C	30 min	363.0 ± 8.0 HV
0.6 K/s	525°C	60 min	358.8 ± 4.2 HV
0.6 K/s	525°C	120 min	352.4 ± 11.0 HV
0.6 K/s	650°C	30 min	283.0 ± 4.0 HV
0.6 K/s	650°C	60 min	278.2 ± 4.1 HV
0.6 K/s	650°C	120 min	264.2 ± 6.8 HV
4.0 K/s	400°C	30 min	493.4 ± 3.9 HV
4.0 K/s	400°C	60 min	506.6 ± 3.3 HV
4.0 K/s	400°C	120 min	493.4 ± 5.2 HV
4.0 K/s	525°C	30 min	411.8 ± 8.2 HV
4.0 K/s	525°C	60 min	405.8 ± 6.5 HV
4.0 K/s	525°C	120 min	395.4 ± 10.3 HV
4.0 K/s	650°C	30 min	322.0 ± 9.7 HV
4.0 K/s	650°C	60 min	314.4 ± 9.7 HV
4.0 K/s	650°C	120 min	315.4 ± 10.9 HV
12.5 K/s	400°C	15 min	514.2 ± 6.1 HV
12.5 K/s	400°C	30 min	500.8 ± 2.9 HV
12.5 K/s	400°C	60 min	507.2 ± 3.8 HV
12.5 K/s	400°C	120 min	499.4 ± 6.3 HV
12.5 K/s	525°C	15 min	424.6 ± 5.1 HV
12.5 K/s	525°C	30 min	420.8 ± 6.7 HV
12.5 K/s	525°C	60 min	409.2 ± 2.5 HV
12.5 K/s	525°C	120 min	401.8 ± 9.3 HV
12.5 K/s	650°C	0 min	379.2 ± 7.9 HV
12.5 K/s	650°C	3.75 min	343.0 ± 6.0 HV
12.5 K/s	650°C	$7.\overline{3}$ min	344.2 ± 6.0 HV
12.5 K/s	650°C	11.25 min	343.0 ± 8.9 HV
12.5 K/s	650°C	15 min	333.2 ± 8.0 HV
12.5 K/s	650°C	30 min	322.2 ± 10.2 HV
12.5 K/s	650°C	60 min	315.4 ± 9.3 HV
12.5 K/s	650°C	120 min	312.4 ± 11.3 HV
76.0 K/s	400°C	30 min	504.4 ± 7.1 HV
76.0 K/s	400°C	60 min	502.4 ± 3.5 HV
76.0 K/s	400°C	120 min	496.8 ± 7.8 HV
76.0 K/s	525°C	30 min	411.8 ± 7.8 HV
76.0 K/s	525°C	60 min	407.8 ± 9.3 HV
76.0 K/s	525°C	120 min	408.0 ± 7.9 HV
76.0 K/s	650°C	30 min	327.8 ± 12.1 HV
76.0 K/s	650°C	60 min	316.8 ± 6.6 HV
76.0 K/s	650°C	120 min	307.2 ± 13.6 HV

Table 3: Initial hardness values.

Cooling rate	Hardness (HV1)
0.6 K/s	485.4 ± 20.3 HV
4.0 K/s	643.6 ± 7.1 HV
12.5 K/s	676.0 ± 7.5 HV
76.0 K/s	661.6 ± 5.5 HV

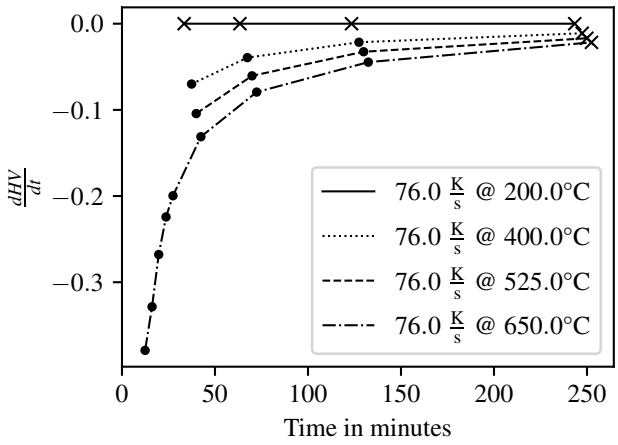

Fig. 3: Extract showing four curves for the rate of change of hardness over time for a cooling rate of $\dot{T} = 76.0\,\frac{\mathrm{K}}{\mathrm{s}}$. Each data point is augmented by calculating the rate of change of hardness with respect to time, i.e., the hardness difference divided by the time difference. Data points marked with 'x' correspond to 'artifically' added data to further restrict the problem.

error and low complexity. As an example, the Pareto front of the hardening model is shown in Figure 2. The model considered hereafter for the first step is shown in Figure 4. The equation describing the model is given by Equation 2.

$$f_1(\dot{T}) = C_1 + \frac{C_2}{\dot{T}} \tag{2}$$

To achieve the accuracy shown in Figure 4, the model parameters are set to $C_1 = 673.0$ and $C_2 = -112.0$.

Tempering model

For the second step, a model for the resulting hardness was developed. To create this model, it was necessary to determine which input parameters would be permitted. From various experimental runs, it became apparent that the resulting hardness primarily depends on three different parameters: the tempering duration t, the tempering temperature T, and the initial hardness HV_{init}. For this step, the experimental data were used for the initial hardness. In contrast to the first model (initial hardness), a simple loss function was insufficient here. For the modeling of the second model, the quotients of hardness change over time change served as the target parameters. To obtain values comparable to the experimental data, the resulting function must be integrated over time. For Symbolic Regression, this means that a simple error function is not adequate. To perform this integration within the optimization, a custom loss function was defined. In addition to the integration, the custom loss function also penalizes deviations from smoothly decaying functions in the time domain and discrepancies from the experimental data. For performing the Symbolic Regression, the same unary and binary operators as in the initial hardness model were permitted. The selected model again represents a compromise between loss and complexity and is given by Equation 3.

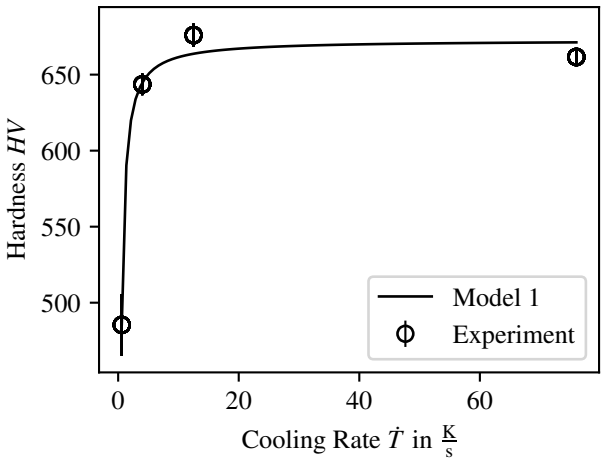

Fig. 4: Graph of the first-step hardening model according to Equation 2. Given the four cooling rates, four different initial hardness values are reached. Each measured initial hardness is plotted with its standard deviation as vertical lines. The model shows a Coefficient of Determination of $R^2 = 98.75\,\%$ and a Root Mean Squared Error of RMSE $= 8.20\,$HV1. The model reaches a plateau past a cooling rate of $\dot{T} \approx 10\,\frac{K}{s}$.

$$f_2(t, T, HV_{init}) = C_3 * t * e^{C_4 * t}(T * HV_{init} + T) \tag{3}$$

$$f_{1+2}(t, T, \dot{T}) = C_3 * t * e^{C_4 * t}(T * f_1(\dot{T}) + T) \tag{4}$$

By selecting the model parameters $C_3 = 1.49 \times 10^{-3}$ and $C_4 = -2.84 \times 10^{-3}$, a model was developed, which is depicted in Figure 5.
In a further figure (Figure 6), the simplification of using the experimental initial hardness was abandoned. The result shows a model which is derived from the combination of the first and second model and is given by Equation 4.

Model validation

Since a mathematically smooth overall model for describing the resulting hardness now exists, it can be used to provide hardness predictions for cooling rates, tempering durations and tempering temperatures not previously considered. Additionally, this overall model can serve as a basis for optimization, whereby the cooling rate, tempering duration and tempering temperature can be determined by specifying a target hardness. Both procedures, forward evaluation and inverse evaluation, are carried out in the following. All validation data are summarized in Table 4. Beginning with the forward evaluation, a series of three additional experiments was conducted. For this purpose, a cooling rate of $1.2\,\frac{K}{s}$ and tempering temperatures of $400\,°C$, $525\,°C$, and $650\,°C$ were selected. To obtain values comparable to previous experiments, the tempering duration was fixed at $30\,$min. The evaluation of the hardness results was performed in the same manner as the previous experiments. The results of the model and the experiments for the forward evaluation are shown in Figure 7. With an $R^2 = 92.59\,\%$ and an RMSE $= 19.48\,$HV1, the validation accuracy is somewhat lower than that of the model with known input parameters (compare Figure 6). As a second step in validating the model, the model was used as the basis for an optimization. It is therefore possible to specify a target hardness, and the task of the optimization is to calculate the optimal cooling rate, tempering duration, and tempering temperature. To test the model's inverse capability, three different hardness values of $\approx 460\,$HV1, $\approx 550\,$HV1, and $\approx 610\,$HV1 were selected, which werent seen by the model before. The tempering duration was again fixed at $30\,$min to obtain comparable results. Thus, only the cooling rate and tempering temperature are freely determinable by the optimization. To achieve the three specified hardness values, cooling

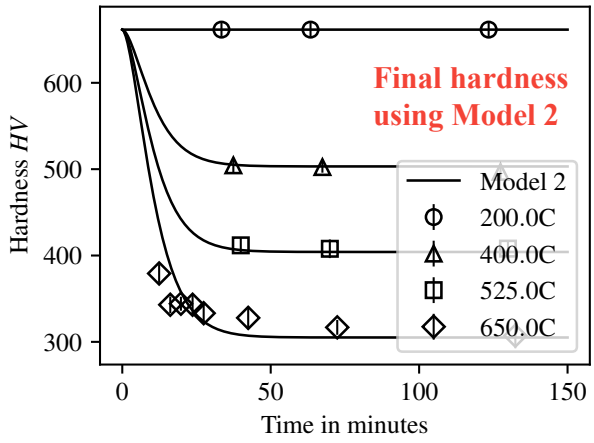

Fig. 5: Graph of the second-step tempering model according to Equation 3. The graph shows the hardness progression for an initial hardness (given from experiment) of $661.6\,\mathrm{HV1} \pm 5.5\,\mathrm{HV1}$ over time. The solid line represents the model and the marker represent the experimental data with the vertical lines showing each measurements uncertainty. Comparing the model, with the given coefficients, to the complete dataset, the model reaches a $R^2 = 98.58\,\%$ and a $\mathrm{RMSE} = 14.49\,\mathrm{HV1}$.

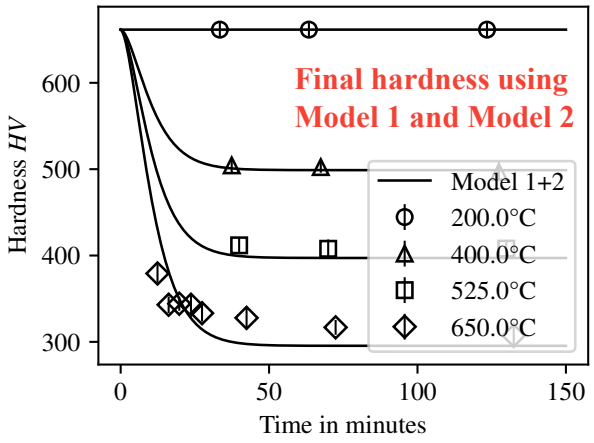

Fig. 6: Graph of the second-step tempering model (Equation 4) with the initial hardness modeled according to the first-step hardening model (Equation 2). Modeling the initial hardness given a cooling rate of $\dot{T} = 76.0\,\frac{\mathrm{K}}{\mathrm{s}}$ results in an initial hardness of $671.5\,\mathrm{HV1}$. Using the approach of predicting the final hardness using both models in series results in a slightly lower $R^2 = 98.50\,\%$ and a slightly increased (worse) $\mathrm{RMSE} = 15.03\,\mathrm{HV1}$.

rates of $35.0\,\frac{\mathrm{K}}{\mathrm{s}}$, $42.0\,\frac{\mathrm{K}}{\mathrm{s}}$, and $45.0\,\frac{\mathrm{K}}{\mathrm{s}}$ were found, which, together with the three found tempering temperatures of $470\,^\circ\mathrm{C}$, $350\,^\circ\mathrm{C}$, and $275\,^\circ\mathrm{C}$, represent the model's prediction. The results of the model and the experiments are shown in Figure 8. With an $R^2 = 71.79\,\%$ and an $\mathrm{RMSE} = 25.72\,\mathrm{HV1}$, the validation accuracy is lower than that of the forward evaluation and the model with known input parameters (compare Figure 6). To further demonstrate the model's accuracy for real-world applications, an additional experimental series was conducted. For actual components, cooling rates are often constrained by the quenching medium. Consequently, only the tempering temperature and tempering duration remain as freely adjustable parameters. For this investigation, target hardness values of $450\,\mathrm{HV1}$ and $550\,\mathrm{HV1}$ were specified. For each target hardness, cooling rates of $8.0\,\frac{\mathrm{K}}{\mathrm{s}}$ and $20.0\,\frac{\mathrm{K}}{\mathrm{s}}$ were fixed. The optimization then determined the required tempering temperature and duration, with all results presented in Figure 9. This validation dataset yields a Coefficient of Determination of $R^2 = 90.97\,\%$ and a Root Mean Squared Error of $\mathrm{RMSE} = 11.96\,\mathrm{HV1}$. While this accuracy is slightly lower than that of the forward validation (compare Figure 7), it is notably higher than the accuracy achieved in the previous inverse validation (compare Figure 8). Considering the entire validation dataset with its 10 data points, an $R^2 = 93.12\,\%$ and $\mathrm{RMSE} = 19.22\,\mathrm{HV1}$, are obtained, which is slightly above the accuracy of the model with known input parameters and shows the models overall capabilities.

Table 4: Validation hardness values.

Cooling rate	Tempering	Experiment (HV1)	Model (HV1)
1.2 K/s	400°C @ 30 min	494.6 ± 35.0 HV	461.40 HV
1.2 K/s	525°C @ 30 min	393.2 ± 12.1 HV	387.49 HV
1.2 K/s	650°C @ 30 min	315.4 ± 17.62 HV	313.57 HV
8.0 K/s	339°C @ 51.10 min	534.4 ± 3.5 HV	550.0 HV
8.0 K/s	467°C @ 53.28 min	453.2 ± 7.4 HV	450.0 HV
20.0 K/s	346°C @ 52.33 min	531.8 ± 5.88 HV	550.0 HV
20.0 K/s	471°C @ 52.88 min	454.8 ± 7.6 HV	450.0 HV
35.0 K/s	470°C @ 30 min	455.4 ± 8.0 HV	459.63 HV
42.0 K/s	350°C @ 30 min	538.6 ± 11.7 HV	553.41 HV
45.0 K/s	275°C @ 30 min	570.2 ± 10.7 HV	612.02 HV

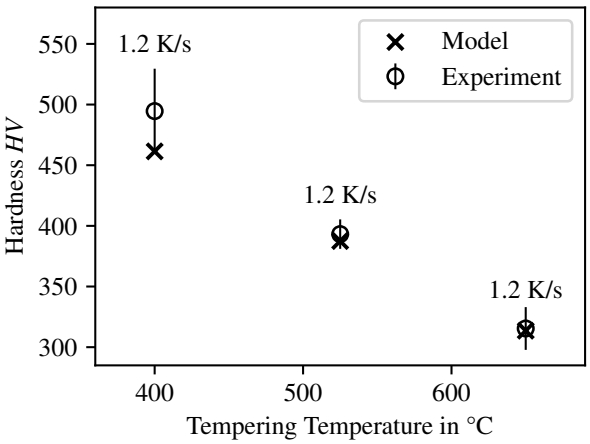

Fig. 7: Forward validation using a cooling rate of $1.2\,\frac{\mathrm{K}}{\mathrm{s}}$ to validate the model on a previously unseen cooling rate. The respective standard error for each experimental data point is displayed using a vertical line. Each plotted data point represents the final hardness after cooling the samples with $1.2\,\frac{\mathrm{K}}{\mathrm{s}}$ and tempering for $30\,\mathrm{min}$. The dataset of these 3 samples yields a $\mathrm{R}^2 = 92\,\%$ and a $\mathrm{RMSE} = 19.48\,\mathrm{HV1}$.

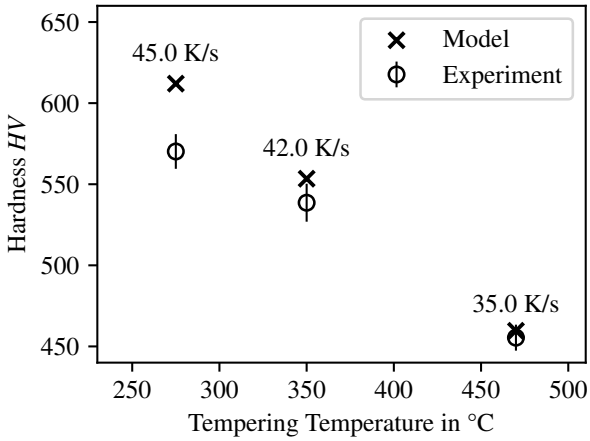

Fig. 8: Inverse validation using a cooling rate of $35.0\,\frac{\mathrm{K}}{\mathrm{s}}$, $42.0\,\frac{\mathrm{K}}{\mathrm{s}}$ and $45.0\,\frac{\mathrm{K}}{\mathrm{s}}$ to validate the model on resulting hardness values not previously attained. The respective standard error for each experimental data point is displayed using a vertical line. Each plotted data point represents the final hardness after cooling the samples with the respective cooling rate and tempering each for $30\,\mathrm{min}$. The dataset of these three samples yields a $\mathrm{R}^2 = 71.79\,\%$ and a $\mathrm{RMSE} = 25.72\,\mathrm{HV1}$.

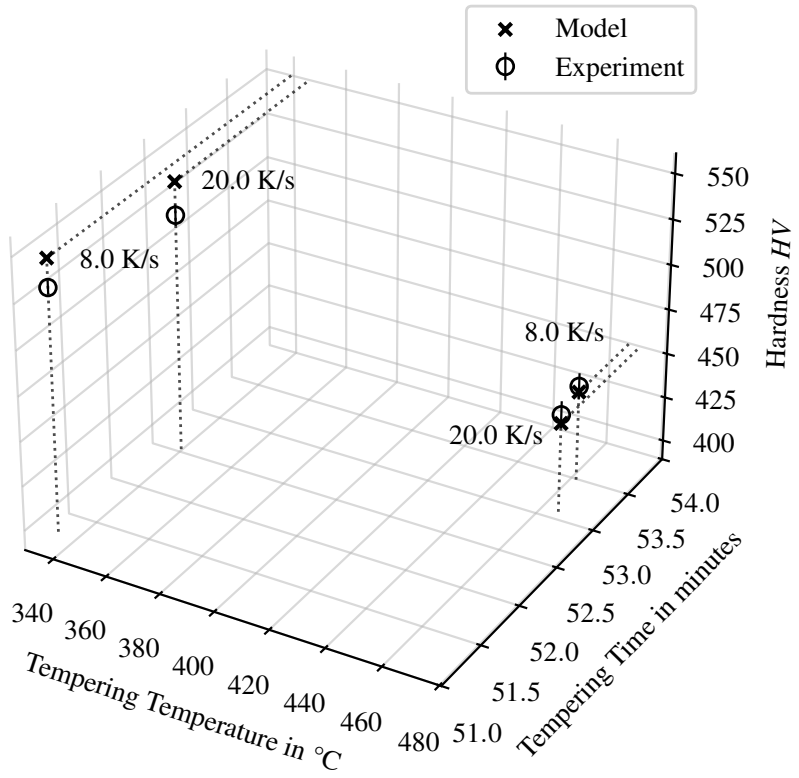

Fig. 9: Second inverse validation using the given cooling rates of $8.0\,\frac{\mathrm{K}}{\mathrm{s}}$ and $20.0\,\frac{\mathrm{K}}{\mathrm{s}}$ and the target hardness values of $450\,\mathrm{HV1}$ and $550\,\mathrm{HV1}$. The respective standard error for each experimental data point is displayed using a vertical line. The second validation dataset, consisting of four samples, yields a $\mathrm{R}^2 = 90.97\,\%$ and $\mathrm{RMSE} = 11.96\,\mathrm{HV1}$.

Discussion

Model fidelity

The following discusses the model accuracy of the individual models and the overall model. First, a model was created to determine the initial hardness based on the cooling rate. A simple model comprising two terms (Equation 2) was identified, which, with the described model parameters, achieves a coefficient of determination of $\mathrm{R}^2 = 98.75\,\%$ and a standard error of regression of $\mathrm{RMSE} = 8.20\,\mathrm{HV1}$. In a second step, a model was developed to determine the resulting hardness based on the tempering duration t, the tempering temperature T, and the initial hardness. The described model is given by Equation 3. This second model is more complex due to the nonlinearities between the input parameters and the time-dependent exponential function. For the second model, a coefficient of determination of $\mathrm{R}^2 = 98.58\,\%$ and a standard error of regression of $\mathrm{RMSE} = 14.49\,\mathrm{HV1}$ were achieved. Considering the second model in isolation reveals its dependency on apriori knowledge of the initial hardness. To establish a comprehensive model, the initial hardness in the second model was replaced with the function for initial hardness (Equation 2). This results in a slightly lower model accuracy, yet still yields a coefficient of determination of $\mathrm{R}^2 = 98.50\,\%$ and a standard error of regression of $\mathrm{RMSE} = 15.03\,\mathrm{HV1}$. To validate the overall model, cooling rates not present in the original experimental data were selected in a first step. This forward validation step shows a coefficient of determination of $\mathrm{R}^2 = 92.95\,\%$ and a standard error of regression of $\mathrm{RMSE} = 19.48\,\mathrm{HV1}$. In a second step, the overall model was used as part of an optimization to determine the cooling rates and tempering temperatures for given target hardness values, and thus revealing the models inverse capabilities. For

this inverse validation step, a coefficient of determination of $R^2 = 71.79\,\%$ and a standard error of regression of RMSE $= 25.72$ HV1 were determined. A possible reason for the comparably lower R^2 lies in the chosen tempering temperature of only $275\,°C$. This tempering temperature might be too close to the lower bound of "no change", which is only based on previous assumptions. In a third step, another optimization with predetermined cooling rates were conducted. For the second inverse validation procedure, a coefficient of determination of $R^2 = 90.97\,\%$ and a standard error of regression of RMSE $= 11.96$ HV1 were determined. For the entire validation dataset, a combined coefficient of determination of $R^2 = 93.13\,\%$ and a standard error of the regression of RMSE $= 19.22$ HV1 are obtained.

Conclusion

Summary and outlook

Both models, for the initial hardness (compare Figure 4) and for the resulting hardness (compare Figure 5), show very good agreement with the experimental data. The use of the overall model (compare Figure 6) also demonstrates very high conformity with the experimental data. However, when evaluating the models on previously unseen data, the overall model exhibits slightly lower accuracy. Consequently, the validation tests yield a coefficient of determination of $R^2 = 93.13\,\%$ and a standard error of regression of RMSE $= 19.22$ HV1.

Table 5: Validating "no changes" below $200\,°C$.

Cooling rate	Tempering	Experiment (HV1)	Model (HV1)
4.0 K/s	200°C @ 30 min	623.2 ± 0.98 HV	645.0 HV
20.0 K/s	200°C @ 30 min	623.8 ± 8.45 HV	671.5 HV

Thus far, only input parameters and target hardness values within the minimum and maximum ranges of the experiments have been investigated. Therefore, the model has only been tested for interpolation within the experimental data. For further development and the identification of a model that more accurately describes the material, the behavior outside the range of the experimental data should also be examined. Additionally, a portion of the validation dataset falls within the $200\,°C$ to $400\,°C$ temperature range. While this is technically an interpolation, it relies on the assumption that no hardness change occurs below $200\,°C$ and that the hardness change only stops at $200\,°C$. Future work is required to critically assess the validity of this assumption, as preliminary experimental data presented in Table 5 already indicate its shortcomings. Furthermore, the physical correctness of the presented models is questionable. The reason for this is the lack of interpretability of the constants in both models. Both models must yield a result with the unit HV1. For the first model (Equation 2), the units of the constants can be assumed as $[C_1] = \text{HV1}$ and $[C_2] = \frac{\text{K}}{\text{s}}$. However, the interpretation of the constant C_2 remains unresolved. For the second model (Equation 3), the units $[C_3] = \frac{1}{\text{K s}}$ and $[C_4] = \frac{1}{\text{s}}$ can be assumed. Here, too, the interpretation of the constants C_3 and C_4 is unclear. Additionally, for future work, the heat distribution throughout an entire component could be considered and used in a more complex simulation setup. This would enable the assessment and adjustment of hardness across the entire component.

Acknowledgements

We acknowledge the contributions of the developers of the free software packages utilized in this work, including, but not limited to, PySR, Matplotlib, NumPy, SciPy, and Pandas. We also acknowledge the role of various large language models (LLMs), such as ChatGPT, Claude, Gemini, and

DeepSeek, in providing suggestions that sparked new ideas and prompted code modifications. Furthermore, we acknowledge the assistance of LLMs in translating parts of this work from its original German manuscript.

References

[1] C. Ulrich, S. Günther, N. Becker, V. Schubert, B. Vetter, C. Leyens, and B. Schlecht: *High strength tempered 42CrMo4 for shafts in drive technology* (Forschung im Ingenieurwesen/Engineering Research, Vol. 89, p. 59) (2025)

[2] A. Çalık, O. Dokuzlar and N. Uçar: *The effect of heat treatment on mechanical properties of 42CrMo4 steel* (Journal of Achievements in Materials and Manufacturing Engineering, Vol. 98, p. 5-10) (2020)

[3] O. Sehyeok and K. Hyungson: *Deep learning model for predicting hardness distribution in laser heat treatment of AISI H13 tool steel* (Applied Thermal Engineering, Vol. 153, p. 583-595) (2019)

[4] B. L. Ferguson, Z. Li and A. M. Freborg: *Modeling heat treatment of steel parts* (Computational Materials Science, Vol. 34, p. 274-281) (2005)

[5] H. A. Mahmoud, G. Shanmugasundar, S. Vyavahare, R. Kumar, R. Cep, S. Salunkhe, S. Gawade and E. S. Abouel Nasr: *Prediction of machine learning-based hardness for the polycarbonate using additive manufacturing* (Frontiers in Materials, Vol. 11) (2024)

[6] M. Cranmer: *Interpretable Machine Learning for Science with PySR and SymbolicRegression.jl* (`https://arxiv.org/pdf/2305.01582`) (2023)

[7] M. Szala, G. Winiarski, Ł. Wójcik and T. Bulzak: *Effect of Annealing Time and Temperature Parameters on the Microstructure, Hardness, and Strain-Hardening Coefficients of 42CrMo4 Steel* (Materials, Vol. 13, p. 2022) (2020)

[8] M. Maniruzzaman, R. J. Pickerill and M. A. Pershing: *Prediction of Tempering Effect on Jominy Hardenability Curve* (Heat Treat 2019, Vol. 2, p. 87-95) (2019)

[9] P. Retzl, Y. V. Shan, E. Sobotka, M. Vogric, W. Wei, E. Povoden-Karadeniz and E. Kozeschnik: *Progress of Physics-based Mean-field Modeling and Simulation of Steel* (BHM Berg- und Hüttenmännische Monatshefte, Vol. 167, p. 15-22) (2022)

[10] T. Wang, J. Du and F. Liu: *Modeling competitive precipitations among iron carbides during low-temperature tempering of martensitic carbon steel* (Materialia, Vol. 12) (2020)

[11] A. Alsheikh and A. Fischer: *Teacher-Student Guided Inverse Modeling for Steel Final Hardness Estimation* (`https://arxiv.org/pdf/2510.05402v1`) (2025)

[12] Information on `https://www.astmsteel.com/product/42crmo4-alloy-steel/` (December 2025)

[13] Information on `https://metalzenith.com/blogs/steel-properties/42crmo4-steel-properties-and-key-applications/` (December 2025)

[14] J. Feng, T. Frankenbach and M. Wettlaufer: *Strengthening 42CrMo4 steel by isothermal transformation below martensite start temperature* (Materials Science and Engineering, Vol. 683, p. 110-115) (2017)

[15] Information on https://songshunsteel.com/the-basic-information-you-need-to- know-about-42crmo4-steel/ (December 2025)

[16] Information on https://escholarship.org/uc/item/3c06175z/ (December 2025)

[17] J. H. Hollomon and L. D. Jaffe: *Time-Temperatures Relations in Tempering Steel* (Transactions of the American Institute of Mining and Metallurgical Engineers, Vol. 162, p. 223-249) (1945)

[18] M. Schmidt and H. Lipson: *Distilling free-form natural laws from experimental data* (Science, Vol. 324, p. 81-85) (2009)

[19] S. L. Brunton, J. L. Proctor and J. N. Kutz: *Discovering governing equations from data: Sparse identification of nonlinear dynamical systems* (Proceedings of the National Academy of Sciences of the United States of America, Vol. 113, p. 3932-3937) (2015)

[20] Information on https://en.wikipedia.org/wiki/Vickers_hardness_test/ (December 2025)

[21] J. N. Kutz and S. L. Brunton: *Parsimony as the ultimate regularizer for physics-informed machine learning* (Nonlinear Dynamics, Vol. 107, p. 1801-1817) (2022)

[22] R. Hooke: *Lectures de Potentia Restitutiva, or of Spring* (1678)

[23] I. Newton: *Philosophiae Naturalis Principia Mathematica* (1687)

[24] J. Kepler: *Astronomia Nova* (1609)

[25] J. Kepler: *Harmonices Mundi* (1619)