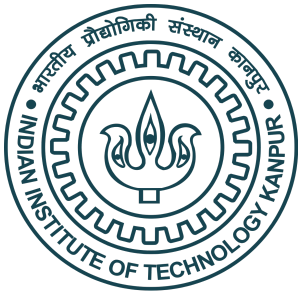
INDIAN INSTITUTE OF TECHNOLOGY KANPUR



SURGE 2022

PROJECT REPORT

"Estimating Creep Curves of High-Temperature Alloys Using Machine Learning and Computational Techniques"

Submitted By

Naman Jain

SURGE Roll No. 2230766

Department of Material Science and Engineering

Indian Institute of Technology Kanpur

Under the guidance of

Professor Niraj Mohan Chawake

Department of Material Science and Engineering

Indian Institute of Technology Kanpur

Kanpur (UP) - 208016, India

CERTIFICATE

This is to certify that the project entitled **"Estimating Creep Curves of High-Temperature Alloys using Machine Learning and Computational Techniques"** submitted by **Naman Jain** as a part of Summer Undergraduate Research and Graduate Excellence 2022 offered by the Indian Institute of Technology, Kanpur, is a bonafide record of the work done by her under my guidance and supervision at the Indian Institute of Technology, Kanpur from 15th May 2022 to 31st July 2022.

Dr. Niraj Mohan Chawake

Associate Professor

Department of Materials Science Engineering

Indian Institute of Technology Kanpur

Acknowledgement

I take this opportunity to convey my sincere regards to my mentor, **Dr. Niraj Chawake** for his outstanding support and invaluable advice which have been instrumental in the successful completion of this work.

Further, I am highly grateful to **Mr. Abhishek Kumar** (Ph.D. student) of IIT Kanpur to act as a co-mentor in project work and for providing me with constant support throughout the SURGE program. I would like to thank IIT Kanpur for choosing me for the SURGE program.

I am thankful to the Department of Materials Science and Engineering, IIT Kanpur. Finally, I would like to thank my friends and family for their constant cheerful support and blessings.

Naman Jain

Student

Department of Material Science and Engineering

Indian Institute of Technology Kanpur

Abstract

Characteristics like oxidation resistance, fatigue strength, corrosion resistance, and creep life are crucial in high-temperature alloys. In our research, we try to study the influence of temperature and applied stress over Ti alloy creep curve behavior by making predictive models that find a creep curve plot at some new environmental conditions using machine learning. Although experimental methods can determine the creep life of superalloy, the measurements are costly and time-consuming due to long-term creep testing and expensive alloy manufacturing. Research on the determination of just the creep lifetime already exists using machine learning models and artificial neural networks (ANNs). Our study tries to predict an alloy's whole strain v/s time creep plot at some new temperature.

Machine learning, which is based on data science, can make the best use of existing raw material data by analyzing internal relationships between data. Relationship models could be built to predict the properties of materials and support the design of new alloys. Data-driven materials development is the core idea of “Material Informatics."

In addition, the stability of microstructure and mechanical properties of materials used in power equipment components in the creep process cannot be ignored, which must be considered when introducing a new steel grade in power engineering. Since creep test usually takes years to complete, the total development process is very time-consuming. There is a need to develop an accurate creep property prediction model to accelerate the development process.

Contents

Introduction

Alloys are used for a huge variety of applications, from the Aircraft industry and automobile manufacturing to furniture and plumbing solutions. And after the installation of material for any application, the industry tries to find ways to earn profit. They may try to reduce the weight, increase their strength, increase flexibility, reduce the effect of temperature on the material's geometry and many more according to the need of the property of any material in different applications. **High-Temperature alloys** find their great use in industries like the aircraft industry in jet engine blade and chasey manufacturing, **Nuclear Reactor components**, jet engines, rocket propulsion systems, turbine blades, pipes in steam turbine power plants, bulb filaments and natural phenomena like the movement of Glaciers.

Creep is the deformation of material on application of continuous constant stress at high-temperature. All materials break at high stress but even when the system can support a load at the instance of its application, it may still break at a later time by creep rupture. Accumulation of microstructural damage may be associated with the thermally activated crossing of energy barriers. Local damage accumulation reduces the energy barriers for future damage activation, thus promoting a tendency to localization. Overall, creep deformation is generally known to have three temporal regimes or stages in a creep life. First, we observe a decelerating strain rate regime associated with (statistical) hardening or aging effects as the weakest elements of the microstructure deform first and become consequentially inactivated by internal back stresses. The decelerating regime is followed by an intermediate regime of constant strain rate and a final accelerating strain rate regime, associated with damage accumulation and strain localization and leading to catastrophic breakdown. The steady-state region, the secondary stage, is desired the most, as it involves a constant strain rate over time, so the material is geometrically stable.

For obvious reasons, understanding the creep behavior and the creep failure dynamics of material is very crucial for the stability analysis of structures. The

prediction of creep curves for a particular alloy at new temperature could be very useful in manufacturing machines or tools that have very low geometrical tolerances. The deformation strain rate for such material can help in precise manufacturing and so avoid future failure of machines due to high temperature creep deformation.

For a particular class of materials, predicting some critical points or regions in the creep curve for a different temperature other than that fed would be highly beneficial in scientific and applied engineering. Machine learning models involve training the model with a significant data set that could be directly experimental data or from the available literature. Initially, the data set is sorted for different parameters and identified essential features. When training happens, we try to increase and optimize the model's accuracy by finding different mathematical equations under which our real experimental curve lies and reducing the error of predictions made.

Literature Review

The study of the creep curve for alloys working at high temperatures is crucial in predicting the estimated stress-rupture time and the deformation rate (or creep rate). Continued slow plastic deformation in materials due to constant stress with time may result in catastrophic failures. Traditionally, the **Larson-Miller parametric method and the θ projection method** are widely used to predict creep life. The Larson-Miller parametric method expresses the relationship between temperature and creep life as a function of stress, and extrapolates to long-term life of alloys by fitting the data in the short term creep life region. The θ projection method and its modified model, are based on the assumption that the creep process is composed of creep strengthening and softening.

Owing to the small data amount, the precision of the model which was constructed with the alloy composition, creep temperature, and creep stress as inputs and used to predict the creep life is low, but the precision of the model which was constructed with the alloy composition, creep temperature, and creep life as inputs and used to predict the creep stress is relatively high. The high-precision model was applied to judge and adjust the prediction results of the low-precision model, thus effectively improving the accuracy of the overall creep life prediction. These properties are temperature- and time-dependent and generally do not exhibit a linear relationship between the parameters and may also degrade at high temperatures. The burger's model and the Findley power law were utilized to model the creep behavior. Several creep models like Norton, Bailey, Bartsch, Garofalo, Theta, Omega, and many more are used to represent deformation behavior. The general equation to represent creep is as follows:

$$\frac{d\epsilon}{dt} = \frac{C\sigma^{m}}{d^{b}} e^{\frac{-Q}{kT}}$$

That relates to creep strain, activation energy, applied stress, grain size (*d*), absolute temperature and constants like *C* that is dependent on material, *m* and *b* are exponents dependent on the creep mechanism.

The simplest method of presenting creep data is a plot of creep strain versus time as is obtained directly from the test data. If sufficient strain-time curves are available for different temperatures and stress levels, the data can be cross-plotted to yield other types of curves, such as isochronous curves.
A curve is required for each condition of temperature and stress. These curves do not permit interpolation or extrapolation. Several methods of correlating creep data stemming from the Arrhenius rate equation:

$$k = \frac{\partial \epsilon}{\partial t} = A\sigma^{n} e^{\frac{-Q}{RT}}$$

Depending on the temperature and stress, different deformation mechanisms are activated. Though there are generally many deformation mechanisms active at all times, usually one mechanism is dominant, accounting for almost all deformation.

Various mechanisms are:

- Bulk diffusion (Nabarro–Herring creep)
- Grain boundary diffusion (Coble creep)
- Glide-controlled dislocation creep: dislocations move via glide and climb, and the speed of glide is the dominant factor on strain rate
- Climb-controlled dislocation creep: dislocations move via glide and climb, and the speed of climb is the dominant factor on strain rate
- Harper–Dorn creep: a low-stress creep mechanism in some pure materials

At low temperatures and low stress, creep is essentially nonexistent, and all strain is elastic. Materials experience plastic deformation rather than creep at low temperatures and high stress. At high temperatures and low stress, diffusional creep tends to be dominant, while dislocation creep tends to be dominant at high temperatures and high stress.

There are two broad mechanisms by which steady state creep takes place: diffusion creep and dislocation creep.

Creep mechanisms:

Diffusion Creep:

Diffusion creep occurs by transport of material via diffusion of atoms within a grain. Like all diffusional processes, it is driven by a gradient of free energy (chemical potential), created in this case by the applied stress. For example, an applied tensile stress creates regions of high hydrostatic tension at the extremities of each grain, along the loading direction. The hydrostatic stress is lower in what might be termed the "equatorial" regions of the grain. Since atoms have lower free energy in these "polar" regions of high hydrostatic stress (ie "low pressure" regions), they will tend to diffuse towards such regions, and this motion will lead to elongation of the grain along the loading direction. Since this occurs on the scale of the individual grains, diffusion distances are shorter in fine-grained materials, which thus tend to be more susceptible to creep. There are two types of diffusion creep, depending on whether the diffusion paths are predominantly through the grain boundaries, termed Coble creep (favored at lower temperatures) or through the grains themselves, termed Nabarro-Herring creep (favored at higher temperatures).

Dislocation Creep:

Dislocation creep is a mechanism involving motion of dislocations. This mechanism of creep tends to dominate at high stresses and relatively low temperatures. Dislocations can move by gliding in a slip plane, a process requiring little thermal activation.

Machine learning prediction methods have been used many times before as well to predict the time of rupture due to creep and the finding the best possible alloy design composition in certain conditions of stress and temperature. Methods like the regression models were used well before also and even using Neural Networks that require a huge dataset for training and building the neural networks were used.

Methodology

Input Data

We have conducted our experimental tests on **Ti-6-Al-7-Mo-4-Zr-3-Sn**, an **alpha-beta** titanium-based alloy with higher strength, low density and good corrosion resistance. This Ti6-Al4 alloy is being studied for a bigger ISRO research project. Until now, we obtained the datasets at six different temperatures at the same constant stress, but a considerable number of datasets are required to train the machine learning model. To predict the whole strain versus time creep curve we needed the strain data at small regular intervals of time for a sufficient number of samples at different temperatures, keeping the stress constant.

Data Processing and Analysis

There was an insufficient amount of strain vs. time data for different temperatures to train any machine learning model. If we used only **6 sample data** then even the simplest Linear regression model also just finds and draws a single linear line that fits the feeded data with the minimum squared sum error possible. With the Literature review, I found very little data on any particular alloy and varying stress. In our study, we strictly required the data for a particular alloy at different constant temperatures and at constant stress for all the samples. We wanted to generate dummy data and use them as datasets for training our machine learning model that could later be replaced with real world experimental datasets for any alloy. Now, after looking at the crucial components of any creep curve, in our case the Titanium alloys creep curves, we see that the primary stage almost resembles to a vertical straight line and then after reaching a point of strain after which the transition of creep curve from primary to secondary stage occurs. Let's call that strain as the **Transition strain**. In the secondary stage, the graph is almost linear with a much less slope as compared to the primary stage. After the secondary stage, we expect a sharp increase in the graph's slope and then finally, the creep rupture happens, but what we observed was after the steady state ended, the material got ruptured soo quickly that it almost looked like it happened just after the secondary stage ended. So, we consider three major data to capture for predicting any creep curve:

1. Transition Strain
2. Steady State Slope
3. Creep Rupture Time

Let's first consider the determination of steady-state slope. We know there can be many different deformation mechanisms and forces acting on the material, but there is a **dominant mechanism** applicable for a temperature range. So, the activation energy needed for that mechanism to take place must also be very nearly the same at all the temperatures in that range. Now, using the well-established Arrhenius type equation for determining the minimum strain rate that is the steady state:

$$\frac{\partial \varepsilon_{ss}}{\partial t} = C\, exp(\frac{-Q}{RT})$$

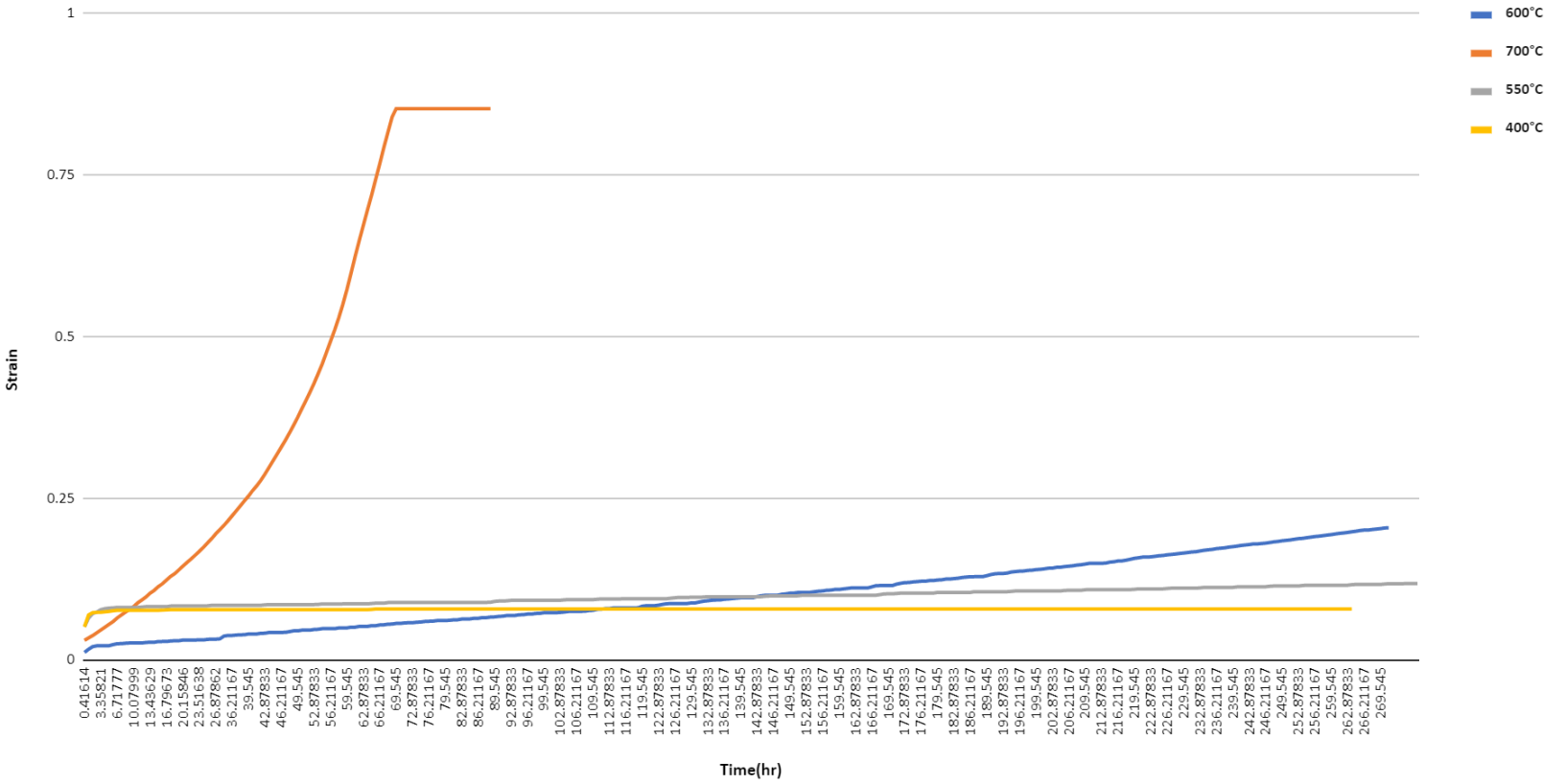
Here, we each temperature range, we have the strain rate and the **activation energy (Q)** as the variables. By solving this equation using available data, we get the activation energy for every temperature range. This is programmed in MATLAB if we just feed the strain and corresponding time value. So, our second component of finding the creep curve is determined.

For the transition strain, we only need to consider the strain value at each sample we have, after which the transition is observed. These points can easily be determined by observing the slope change and then we used Linear regression to find the transition strain for other temperatures, by assuming that the transition strain varies linearly upwards with increase in temperature.

Note that the increase in the transition strain with increase in temperature is not mandatory and alter a little. We also observed that for some cases the steady state graph for a high temperature sample may be below a low temperature sample for some time initially. But then as the slope of higher temperature sample must be bigger, so finally the graph moves up as compared to the lower temperature graph. This approximation and the error it may cause in our prediction results was considered and yet to be resolved to make the predictions more accurate. In this

way, we get the first component of the creep curve prediction as well. The third component that is the creep rupture time was not in our area of interest and so we will not claim about the creep lifetime for any new temperature. Studies and researches have already been held on the creep lifetime prediction.

Using the above methodology, a function to predict the creep curve for any temperature was already programmed in the **MATLAB** software. This function was then used to generate dummy data that was used in place of experimental data in the various ML models that we trained and evaluated.



Model Selection and Evaluation

The first step of building the prediction model is crucial because it forms the basis on which the prediction models will be trained and useful predictions will be made. Dataset is the elemental component of Machine Learning, it is the information/data

organised in form of rows and columns, where each row is an instance of data and each column represents a feature or output. These datasets as we will see in the next section are processed to gain some insights regarding any regularities or patterns that may prove helpful for building an efficient model. Here, we used our experimental creep data and the generated **dummy data** from the MATLAB function made, to train the ML models and make better predictions expectedly.

The performance of the prediction model depends greatly on the quality of the dataset. The raw dataset built from data collection is not cleaned and has noises, redundant information, etc., which if not eliminated may decrease the accuracy of the model in making predictions. Therefore, the step of data pre-processing is necessary to make the dataset viable for the ML model.

The composition space of the dataset was normalised using the defined function of min-max normaliser which linearly rescales every feature (element in composition space) to the [0,1] interval. The transformation values applied to the values of the column is:

$$z = \frac{x - min(x))}{max(x) - min(x)))}$$

where, x: any value in the column

min(x): minimum value of that column

max(x): maximum value of that column

The dataset was split into the training set and the test set in the ratio of 4:1 using *train_test_split* from *sklearn* library.

All **266 samples** were divided into several clusters by clustering. In each cluster, the alloy samples were divided into the training set and testing set through 10-fold cross-validation to evaluate the generalization ability (*fitness* value) of five candidate prediction models for the cluster (SVR, RF, GPR, LR, and RR). In the **10-fold cross-validation** on alloy samples of each cluster, 9 folds of samples were used as training dataset *T* for model training and 1 fold as testing dataset *S* for

model testing. The overall performance of the candidate model for each cluster is the average of the *fitness* values over all 10 iterations.

The workflow includes the following four steps:
(i) Principal component analysis (PCA) was employed to reduce the dimension of input features, and K-means clustering was used to evaluate the visible clustering of the reduced dimension data. Shapiro–Wilk test was introduced to detect the data distribution of all the target features
(ii) Nine different regression methods, consisting of linear regression (LR), stochastic gradient descent (SGD), random forest (RF), decision tree (DT), multi-layer perceptron (SGD), support vector regression (SVR), k-nearest neighbor algorithm (KNN), kernel ridge regression (KR) and AdaBoost (ABR), were utilized to predict four creep target features
(iii) K-fold cross validation and three metrics, namely coefficient of determination (R2), normalized mean absolute error (MAEc), normalized root mean square error (RMSEc), were used to evaluate and select the best model among all the regression models
(iv) Feature correlation analysis was used to explain the internal relationship between input features and target features using Pearson correlation coefficient (PCC) and Spearman correlation coefficient (SCC).

It is evident that this method can predict the properties of compositions in the coverage of the dataset, help to capture the optimal points of elemental interactions, and it can be used to do a high-throughput evaluation of the alloying effects during alloy design.

Conclusion and Future Scope

High Temperature alloys are often cast as a single crystal—while grain boundaries may provide strength at low temperatures, they decrease creep resistance. The primary application for such alloys is in aerospace and marine turbine engines. Creep is typically the lifetime-limiting factor in gas turbine blades. The machine learning explorations conducted on the creep dataset demonstrate the potential of the approach to achieve higher prediction accuracy with ***RMSE, MAPE*** **and** ***R*****2 of 0.3839, 0.0003 and 0.9176** than five alternative state-of-the-art machine learning models. We develop a **divide-and-conquer self-adaptive (DCSA) learning method** incorporating multiple material descriptors for rational and accelerated prediction of the creep rupture life. We characterize a high-quality creep dataset of 266 alloy samples with such features as alloy composition, test temperature, test stress, and heat treatment process.

For a particular class of materials, predicting some critical points or regions in the creep curve for a different temperature other than that fed would be highly beneficial in scientific and applied engineering. Machine learning models involve training the model with a significant data set that can be directly experimental data or from the available literature. Initially, the data set is sorted for different parameters and identified essential features. When training happens, we try to increase and optimize the model's accuracy by finding different mathematical equations under which our real experimental curve lies and reducing the error of predictions made.

Thus, our study promises the use a machine learning as a **reliable approach** to predict the creep behavior of materials for various operating conditions. Plausibly, this study will help reduce a great deal of labor, time, and experimental cost by efficiently predicting mechanical properties and suggesting the structural design of various high-temperature alloys. In future, the project will continue to build the ML model to predict the creep curve data with carrying temperature as well as stress, that would be **at par with the experimental results**, with feature variables as material properties, and finally deduce the composition of alloy for desired strength and ductility in the alloy.

References

1. Prediction of creep failure time using machine learning
2. A machine-learning approach to predict creep properties of Cr–Mo steel with time-temperature parameters
3. Prediction of creep failure time using machine learning
4. Prediction of creep curve of HP40Nb steel using artificial neural network
5. Recent advances and applications of machine learning in solid-state materials science