Towards Shareable Materials Science: Cloud-Based Data-Driven Modeling for Fatigue Life Prediction in Ti-6Al-4V for Turbine Blade Applications

ABSTRACT

While there have been many efforts directed towards experimental investigations and establishing micromechanical models for fatigue life predictions of titanium alloys for turbine engine applications under low cycle, high cycle, and dwell fatigue conditions, the complexity of the relationship between processing, microstructure, and fatigue life makes generalization of the models challenging. Such efforts would benefit from sharing of information scattered across corporations, academia, and government laboratories, but this represents only a first step. The real value comes from analyzing the information to extract correlations to be used in predicting fatigue life of titanium alloys as a function of the underlying microstructure (texture, morphology, and chemistry) and the preceding thermomechanical processing steps. Both tasks can be achieved using cloud computing and Web 2.0 concepts that emphasize data analytics of user-generated content, usability, and interoperability.

In this report, we attempted to utilize a shared-science approach by building data-driven modeling tools using collected experimental efforts scattered online. We demonstrated such a model using non-linear interpolation models based on Bayesian neural networks (BNN) to predict low cycle and high cycle fatigue performance of turbine engine blade materials, Ti-6Al-4V and Ti6Al4V-ELI as a function of chemistry, heat treatment, microstructure, and waveform of the loading cycle including (i.e., strain-controlled and load-controlled testing). This work is aimed to establishing empirical relations that predict fatigue life as a function of microstructure and TMP which can be easily incorporated into existing FEM design tools. A companion web application was developed to enable online access to the curated data and modeling results with the ability to add new datasets by subscribers for future expansion of the model to increase the scope and accuracy of model predictions.

Model training and validation was accomplished using large amount of experimental data that was published by the National Institute of Material Science (NIMS) between 2002-2010 with different heat treatments, chemistry, and presumably from different suppliers or lots. The data was inserted into an online database that was linked to a microstructure informatics cloud computing (MiCloud). Adding new datasets will be enabled for researchers to upload their data for use by the community.

BACKGROUND

The Metallurgical Problem

To predict fatigue life in titanium alloys (i.e. high-cycle, low-cycle, or dwell), it is crucial to account for the underlying microstructure (including chemistry, morphology, and crystallography) and spatial distribution of constituents which results from standard thermomechanical processing protocols (e.g. MIL-SPECs, AMS, ASTM, BS). However, in 1990s a Scientific Advisory Board of the US Air Force concluded that high-cycle fatigue (HCF) was the single biggest cause of turbine engine failures in military aircraft. Almost 10 years later in 2011, a Ti-Working-Group (TiWG) led by NASA JPL [1] confirmed the presence of non-confirming titanium from various suppliers which had detrimental effect on the fatigue life of spacecraft. While government agencies and large aerospace companies may have private databases for fatigue life in Ti alloys, knowledge extracted from data analytics of this information is limited to the individual agency, leading to duplication and redundancy of work across the fences of interested parties. By leveraging tools and datasets in the open domain, some of this work can be avoided. Despite the wealth of experimental results on both high-cycle and low-cycle fatigue of various turbine engine alloys, there have been limited studies on ways to analyze these findings. Some of the main reasons include: (i) the absence of repositories for all generated data (locally and globally), (ii) lack of a protocols for methodology and prioritization of work based on minimization of duplication of effort, (iii) lack of data analytics tools that enable researchers to analyze published data which causes wasted research money to reproduce published data.

Turbine disks and large forgings made of titanium are known to have spatial variability in microstructure (morphology and crystallography)[2], known as macrozones or microtextured regions (MTRs)[3]. These features are known to have drastic impact on fatigue life in alpha/beta titanium alloys under dwell conditions. While eliminating such heterogeneities is a difficult task in large forgings, small bar material may have less spatial variability in microstructure heterogeneities. Also, selecting the fatigue test samples from the same location in the bar material may also reduce the impact of variability between samples on measured fatigue life. Building a data-driven model on such experimental data will be beneficial for turbine engine blades manufactured from Ti6Al4V. With the above logic in mind, we have selected the NIMS fatigue data for Ti6Al4V that were measured on various bars made of Ti6Al4V under various TMP and loading conditions to enable us building a model that links microstructure, processing, and fatigue life with minimum uncertainties in the starting texture. .

The Data Sharing Problem

Reliable models for predicting fatigue life as a function of microstructure and the associated TMP could greatly improve the efficiency of manufacturing. However, the success of such a model and the associated optimization process is dependent on the assumption that the model developer selected all the features of interest (FOI) in the microstructure that will improve fatigue life. However, such an assumption will be based on the accuracy of the knowledge between the microstructure and fatigue life. Despite the wealth of microstructure-fatigue models in literature for titanium alloys, many are built under the hypothesis that the correlations between the microstructure FOI and the fatigue life are already known. Support for these assumptions is based on limited experiments that can still amount to millions of dollars. If the model fails in certain situations, the need for new FOIs in the microstructure rises to find linkages with the fatigue life which requires repeating many costly experiments and collection of more microstructure data looking for the new FOI.

To make material science and engineering a valuable research product[4], we have embraced an “open science” model by taking two steps: (i) building an infrastructure for Microstructure Informatics Cloud Computing (MiCloud^{TM})[5] that promotes global sharing of processing-microstructure-performance data for material science and engineering based on a detailed ICME workflow[6] and (ii) starting the sharing process by sharing microstructure datasets and eventually the model associated with this report. MiCloud enables the storage, analytics, and sharing of such information while separating this process from the models using them. The unique part of it is the raw data and meta-data are stored on databases that allow various developers to access the databases and train various models. Some of the current data analytics tools on the website includes higher order statistics for microstructure quantification based on n-point correlation functions[6] to extract spatial information about various microstructure features. Such tools are crucial for correlating fatigue life to microstructure features. The fact that most high quality discussions linking microstructure to fatigue life report the spatial location of crack initiation sites (e.g. surface vs sub-surface) and interpretations/descriptions of the fracture surface (ductile fracture, facets, etc) is evidence that correlating fatigue life to microstructure requires knowledge about spatial information of various FOI in the microstructure. With a focus on pattern recognition instead of pattern matching various microstructure-performance correlations can be extracted [6] to train data-driven models. However, the success of these data-driven methods is highly dependent on the availability of the data and the willingness of the scientific community to share it on an infrastructure that embraces sharing and reusability.

While our model will benefit tremendously from government regulations demanding sharing of publically funded research, we are bidding on the desire of researchers to fairly share their work in return for credit, access to data analytics tools, microstructure data, and models.

Data-Driven Modeling

During the last decade data-driven models have become popular in the materials science field to extract complicated correlations between microstructure, processing, and performance[7,8]. These models rely on machine learning, and thus assume the presence of a considerable amount of data capturing the physics of the modeled behavior. The recent push for materials ontology[9] will drastically benefit these “data hungry” or “data-driven” approaches (e.g. machine learning) to extract knowledge for use in materials design. However, accurate predictions from these models require the wide adoption of materials ontology which is faced with overcoming the current paradigm of publishing data in journal articles and storage on private drives. Therefore, we are starting by designing data-driven models in this report to incentivize other researchers to share data and offer additional models to be used by others. We believe if we are to win this challenge and reveal our data-driven model for the community, others will follow which may enable the realization of following materials ontology.

Data driven models can be as simple as linear regression in Excel or as complicated as supervised machine learning using neural networks or deep learning[10]. In this report, we are offering a solution using publically available data on low cycle and high cycle fatigue behavior of titanium alloys by the National Institute for Materials Science (NIMS)[11] to setup a data-driven model for fatigue life prediction. To entice researchers to add their data (future or past), the online use of the model will be offered for researchers who will accept sharing their data via MiCloud[5].

METHODS

Data Curation and Storage in a NoSQL Database

The NIMS fatigue database for Ti6Al4V included information in ten PDF files which are not suitable for any data analytics processes. Therefore, the first step was the extraction of important information from the PDF reports covering the period of 2002 to 2010 into text files, tables with numerical values, images of various microstructures and charts. While the current volume and variety of the data may be handled with standard file system, the current effort is part of a long term project for our group which is planned to grow with more published data as well as uploaded data by MiCloud users. As such, we have adopted various steps following modern web-scale databases by adding the files of various formats into our non-relational database (aka Not only SQL, or NoSQL [http://nosql-database.org/]) which has the characteristic of schema-free, easy replication support, simple API, and "horizontal" scaling to clusters of machines to handle huge amounts of data. Adding all these files from ten reports into a NoSQL database on MiCloud enables future addition of data from other sources (e.g. published research articles or uploaded datasets by MiCloud users).

Building a SQL Database

While the NoSQL database is crucial to handle our unstructured materials-Big-Data, building data-driven models from various heterogeneous databases requires running queries tailored for various data products. As such, we used our cloud-based SQL database and developed schema for the NIMS fatigue data for Ti6Al4V[11].

Data-driven model using Bayesian Neural Network (BNN)

Due to the large number of variables (processing, microstructure, and loading conditions) that affect the fatigue life in titanium alloys, our modeling work follows a two-step process: (i) data-driven models for pattern recognition of phenomenological correlations between inputs and output followed by (ii) targeted micromechanical models guided by the correlations in the first step to establish causations behind some of the correlations. Thus, our first step was to establish a data-driven model using an artificial neural network[12] with Bayesian regularization[8,13] where the processing, microstructure (chemistry and morphology), static properties, and loading conditions were the network inputs and the output was the resulting fatigue life. Regularization is an approach to avoid overfitting (i.e., where the model learns parameters that enable it to exactly reproduce the training data, but yet performs poorly with data outside that training set) where an extra cost function is added during the training process which penalizes complexity in the model. In this work, Bayesian regularization was employed as it has been shown in literature to perform well in the case of sparse data.

We followed three steps during the development of our BNN, namely, training, validation, and testing. Training the model requires calculating the optimum weights by minimizing the cost function (the sum of the squared errors) (E_D) between the training data $\boldsymbol{t_i}$ and the model response $\boldsymbol{a_i}$ (i.e. N_f in this report) using $\boldsymbol{n}$ training data points (we have used 70% of the data for training) :

$$E_D = \sum_{i=1}^{n} (t_i - a_i)^2 \tag{1}$$

Regularization adds another term which enforces smoothness of the model, giving an objective function $F = \beta E_d + \alpha E_w$. Here E_w is the sum of the NN weights (w_i) and α and β are hyperparameters to determine the relative effect of each term. If $\alpha << \beta$ then the model will attempt to match the training weights exactly potentially overfitting, if $\beta << \alpha$ the regularization will favor a smooth and simple model at the expense of a large training error. The primary difficulty in regularization is setting the correct values for the objective function parameters. There has been substantial work in applying Bayesian approaches for NN model selection and comparison, most well-known is the work of David MacKay and colleagues[7,8]. Here we adopt such a framework modeled after the routine of Foresee and Hagan [14].

Consider the set of NN weights as a vector of random variables, $\mathbf{w}$. The probability density of the weights can be updated according to Bayes' rule as:

$$P(\boldsymbol{w}|\, D, \alpha, \beta, M) = \frac{P(D|\, \boldsymbol{w}, \alpha, \beta, M) P(\boldsymbol{w}|\, \alpha, M)}{P(D|\, \alpha, \beta, M)} \tag{2}$$

where D represents the training data, and M is the particular NN model used (considered fixed for our purposes). $P(\mathbf{w}\,|\,\alpha, M)$ is the prior distribution which encodes our prior knowledge of the weights before we start to train the model. $P(D\,|\,\mathbf{w}, \alpha, \beta)$ is the data likelihood function which gives the probability of collecting the training data given the current weights. $P(D\,|\,\alpha, \beta, M)$ is a marginal distribution of the data given the model and the objective function parameters which serves as normalization factor and guarantees a total probability of 1.

Following Foresee and Hagan [14] we assume a Gaussian prior and Gaussian error in the training set data and the following probability densities can be derived:

$$P(D|\, \boldsymbol{w}, \beta, M) = \frac{1}{(\pi/\beta)^{n/2}} exp(-\beta E_D) \tag{3}$$

$$P(\boldsymbol{w}|\, \alpha, M) = \frac{1}{(\pi/\alpha)^{N/2}} \exp(-\alpha E_w) \tag{4}$$

where n is the number of training data points and N is the number of NN weights. By substitution it can be shown, [13,14], that the final updated weights take a very simple form:

$$P(\boldsymbol{w}|\, D, \alpha, \beta, M) = \frac{1}{Z_F(\alpha, \beta)} \exp(-F(w)) \tag{5}$$

where Z_F is a normalization factor.

Assuming uniform prior distributions for α and β , Bayes' rule can be applied to determine optimal values given the data and the model:

$$P(\alpha, \beta | D, M) = \frac{P(D|\alpha, \beta, M) P(\alpha, \beta | M)}{P(D|M)} \tag{6}$$

Maximizing the posterior density is achieved by maximizing $P(D\,|\,\alpha, \beta, M)$. $P(D\,|\,\alpha, \beta, M)$ also turns out to be an explicit function of $Z_F(\alpha, \beta)$. An approximation of Z_F can be determined by considering that at the minimum point of the posterior density of the Hessian matrix of the objective function $Z_F(\alpha, \beta)$ should be approximately quadratic. Thus $Z_F(\alpha, \beta)$ can be estimated as a Taylor series expansion of $F(\mathbf{w})$ about the minimum point of posterior density. $Z_F(\alpha, \beta)$ has an analytically complex but numerically straightforward to compute form that depends solely on the Hessian matrix of the objective function, $H = \beta\nabla^2 E_D + \alpha\nabla^2 E_w$.

For fatigue life predictions, and shows measured fatigue life vs predicted fatigue life on logarithmic scale because the performance in the feature vector is a single scaler (fatigue life in cycles).

Providing a processing-microstructure-fatigue life prediction model using current infrastructure for cloud computing will turn digital research data into valuable assets that can be shared and reused by other researchers. This aligns with the global push for "open science" [reference ward paper in IMMI 2014, 3:22]. While government polices from individual countries may have different requirements for sharing digital data generated from publicly funded research, using incentives for sharing (e.g. the model in this report) will overcome these differences by targeting the researcher needs directly for using an online model to link his work with others or use others data to design new research that does not repeat prior work.

RESULTS AND DISCUSSION

The NIMS database for fatigue of Ti6Al4V [11] includes two grades of alpha/beta titanium alloy, namely, Ti6Al-4V and Ti-6Al-4V ELI. Various bars were supplied for each grade with varying chemistry. Each grade witnessed two heat treatments: (i) solution treating at 930°C for 60min/air cooled followed by aging at 705°C for 120min/air cooled, and (ii) solution treating at 955°C for 60 min/water quenched with aging at 550°C for 240min and air cooling. These heat treatments were designed to produce material in two categories based on ultimate tensile strength of 900 MPa and 1100 MPa. Each combination of heat treated material was tested under two loading conditions, strain control and stress control (Figure 1). These data resulted in 277 fatigue life measurements. Each fatigue life was associated with a feature vector that included processing, microstructure descriptors, chemistry, static properties, and loading condition that was extracted from the SQL database as setup to follow the schema that we designed (Figure 1).

The first step that was done is to find the correlations between fatigue life (N_f) and various input parameters using Spearman's rank correlation[15] for both strain control testing and load control testing (Figure 2). An immediate insight that can be extracted from these figures is the correlation between fatigue life in the high cycle fatigue regime under load control testing and the initiation site compared to the amount of strain under the strain control testing which reveal the difference between crack growth vs crack initiation failures under different loading conditions. Furthermore, the lack of correlation between processing parameters and fatigue life reflect the importance of designing the experimental work to eliminate variability in processing parameters. This highlights the gap in this database that could be filled by other researcher seeking the effect of TMP variability on fatigue life.

The BNN was trained on 70% of the data with 15% for validation, and 15% for testing. Two separate models were used during the training with the first for strain control and the second load control. The validity of the predicted model is demonstrated by comparing the measured and predicted fatigue life (Figure 3). The Pearson's correlations[16] between predicted and measured fatigue life were 0.99 and 0.95 for strain control and load control testing, respectively. The model was better in predicting N_f under strain control conditions compared to load control which can be explained by the absence of more accurate data on crack initiation sites and spatial distribution of microstructure constituents.

A more detailed comparison between model prediction and measurements are revealed by plotting all data for S-N curves for all samples on one plot (Figure 4). While such a busy plot will be of less use to designers, it is important to identify the areas where the model diverted from measured data especially in the high cycle fatigue regime and as expected the run-outs.

Using the BNN model to extract physical insights on the correlation between chemistry and TMP as reflected in UTS and the underlying microstructure are two examples of many that can be extracted from the model. As shown in Figure 5a and c, for both Ti6Al4V (ASTM B348) and Ti6Al4V ELI (ASTM 136) the 1100 MPa heat treatment always resulted in withstanding higher stresses under the same fatigue life or longer fatigue life under the same applied stress than that exhibited by the 900 MPa heat treatment. The effect of chemistry is shown in Figure 5b and d, revealing that Ti6Al4V had longer fatigue life than Ti6Al4V ELI after 900 MPa heat treatment (strain and load controlled loading). However, it had shorter fatigue life under strain control loading (i.e., low-cycle fatigue) if it was heat treated to 1100 MPa heat treatment, though load controlled HCF of Ti6Al4V was higher than Ti6Al4V ELI at the same 1100MPa heat treatment. These results are intriguing that the reduction in interstitial contents in the ELI resulted in longer LCF life when water quenched after solution treatment compared to Ti6Al4V. More detailed microstructure analysis will be needed to investigate the impact of chemistry changes on the thickness of alpha laths and the size of secondary alpha colonies of both materials. Note that the non-smooth curves in Figure 5 can be a result of limited data within the range of S-N curve with kinks. This can be attributed to the need for more experiments in this range that shown multiple applied stresses to result in the same fatigue life between 1.5 E+6 and 1.2 E+7 for Ti6Al4V 900MPa (Figure 5c). This revealed areas for future experiments that will be guided by the BNN results. To enable other researchers to add new data for better predictions by the BNN, a website on MiCloud was setup to allow them to fill the needed data following the NIMS data structure which feeds directly into MiCloud databases. Users can input numerical data manually or upload files, microstructure, and texture (Figure 6).

Conclusions

- Efficient use of data published by other researchers to build a practical data-driven model for material science can be achieved if the available dataset contains well planned experiments to reveal the physics behind the behavior of the material.
- Bayesian neural network (BNN) provides a useful methodology to extract phenomenological correlations between microstructure, processing, and fatigue life.
- Spatial information about various microstructure constituents are needed for better predictions of fatigue life in both low cycle and high cycle fatigue loading conditions.
- Cloud computing offers a practical solution for making material science and engineering data more valuable research product that can be shared and reused.
- Reduction in interstitial contents in Ti6Al4V ELI resulted in enhancement of low cycle fatigue life over Ti6Al4V when they were solution treated at 930°C/60min/AC then aged at 705°C/120min/AC.
- Water quenching from solution temperature followed by aging resulted in increase in fatigue life for both Ti6Al4V and ELI grades under strain control and load control test conditions.

Availability of Supporting Data

All supporting data were curated from [http://smds.nims.go.jp/fatigue/index_en.html] [11], the online database offered by NIMS. The data covers ten reports dating from 2000 to 2011 which can be viewed online via signing for a free account on the NIMS website. It has been previously used in 2013 for demonstrating Materials Ontology by T. Ashino [9] and in 2012 to demonstrate MatDB [17]. To help the judges accessing the data, we hosted it online on MiCloud[5] with a graphical user interface (GUI) to allow the judges fast access to the S-N curves that was published on the NIMS website[11] (Figure 7). To access the dataset on MiCloud, (1) use a web browser (we recommend Google Chrome) (2) navigate to https://micloud.icmrl.com, (3) the User ID: MGIjudge1, and password: MRLBBN0331 and click the "Log In" button. The main application page will open. This page contains icons for all of the applications that are available within MiCloud. The application for the current report is called "Ti-Life". To open this web app, scroll down to the section called Material Models. Click on the icon to open Ti-Life.

.

REFERENCES

[1] JPL Document No D-69922, Report of the Ti Working Group Concerning the Assessment of Non-Conforming Material as It Affects JPL Flight Projects, 2011, .

[2] N. Gey, P. Bocher, E. Uta, L. Germain, M. Humbert, Texture and Microtexture Variations in a near- a Titanium Forged Disk of Bimodal Microstructure, Acta Mater. 60 (2012) 2647–2655.

[3] A.L. Pilchak, K. Nakase, I. Inagaki, Y. Shirai, A.H. Rosenberger, J.C. Williams, The Influence of Microstructure and Microtexture on Fatigue Crack Initiation and Growth in α + β Titanium, n.d., .

[4] C.H. Ward, J.A. Warren, R.J. Hanisch, Making Materials Science and Engineering Data More Valuable Research Products, (2014) 1–17.

[5] Materials Resources LLC, MiCloud:Microstructure Informatics Cloud Computing. URL: https://micloud.icmrl.com/auth/login?from_page=/.

[6] A.A. Salem, J.B. Shaffer, D.P. Satko, S.L. Semiatin, S.R. Kalidindi, Workflow for Integrating Mesoscale Heterogeneities in Materials Structure with Process Simulation of Titanium Alloys, Integr. Mater. Manuf. Innov. 3 (2014) 24.

[7] C.A.L. Bailer-Jones, T.J. Sabin, D.J.C. MacKay, P.J. Withers, Prediction of Deformed and Annealed Microstructures Using Bayesian Neural Networks and Gaussian Processes, (1997) 1–7.

[8] D.J. MacKay, Probable Networks and Plausible Predictions-a Review of Practical Bayesian Methods for Supervised Neural Networks, Netw. Comput. Neural Syst. 6 (1995) 469–505.

[9] T. Ashino, Materials Ontology: An Infrastructure for Exchanging Materials Information and Knowledge, Data Sci. J. 9 (2010) 54–61.

[10] Y. LeCun, Y. Bengio, G. Hinton, Deep Learning, Nature (2015) 436–444.

[11] National Institute for Materials Science, NIMS Fatigue Data Sheet for Ti-6Al-4V. URL: http://smds.nims.go.jp/fatigue/index_en.html.

[12] H.K.D.H. Bhadeshia, H.J. Stone, Neural-Network Modeling, 22 (2009) 435–439.

[13] K.P. Murphy, Dynamic Bayesian Networks: Representation, Inference and Learning. University of California, Berkeley, 2002.

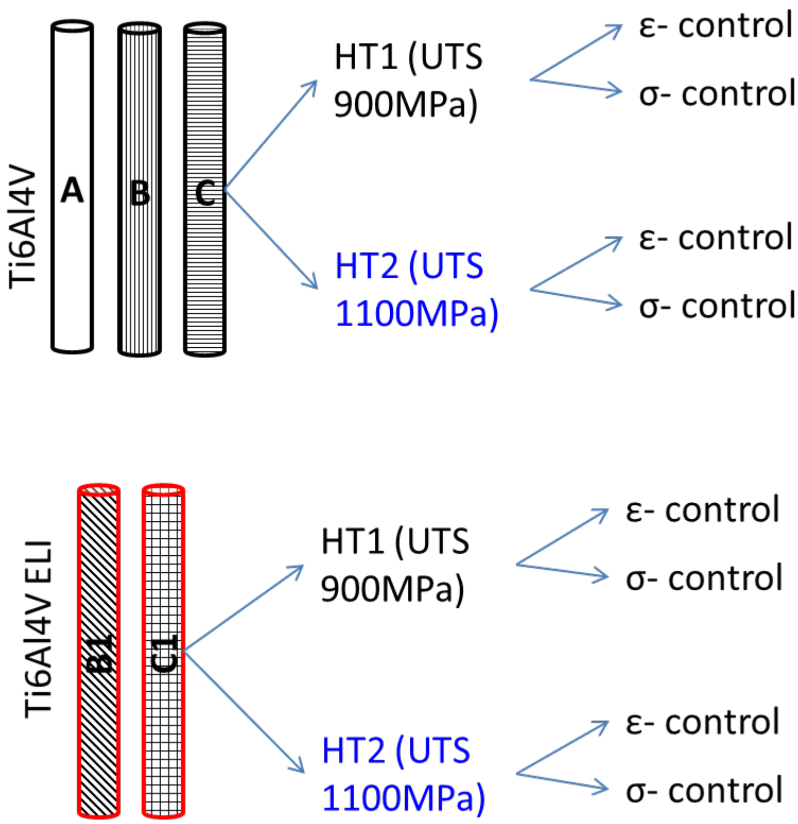
[14] F.D. Foresee, M.T. Hagan, Gauss-Newton Approximation to Bayesian Learning, in: Proc. 1997 Int. Jt. Conf. Neural Networks, vol. 3 Piscataway: IEEE, 1997, pp. 1930–1935.

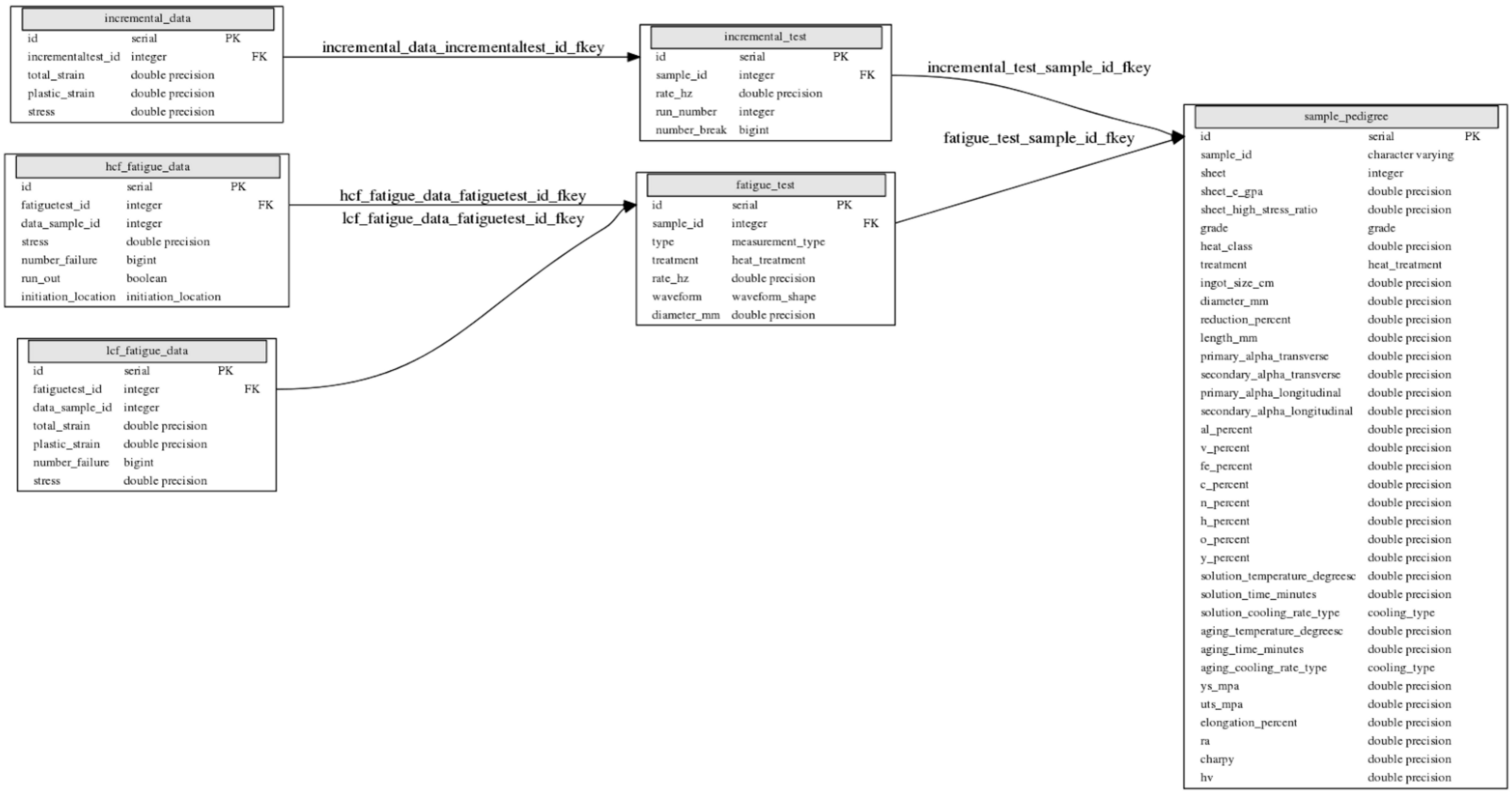
[15] M.G. Kendall, Rank Correlation Methods, 1970, .

[16] R.A. Fisher, Statistical Methods for Research Workers, 1958, .

[17] T.S.P. Austin, H.H. Over, MatDB Online-a Standards-Based System for Preserving, Managing, and Exchanging Engineering Materials Test Data, Data Sci. J. 11 (2012) ASMD11–ASMD16.

FIGURES


Ti6Al4V
A
B
C
HT1 (UTS 900MPa)
HT2 (UTS 1100MPa)
ε- control
σ- control
ε- control
σ- control
Ti6Al4V ELI
B1
C1
HT1 (UTS 900MPa)
HT2 (UTS 1100MPa)
ε- control
σ- control
ε- control
σ- control



incremental_data
id serial PK
incrementaltest_id integer FK
total_strain double precision
plastic_strain double precision
stress double precision
hcf_fatigue_data
id serial PK
fatiguetest_id integer FK
data_sample_id integer
stress double precision
number_failure bigint
run_out boolean
initiation_location initiation_location
lcf_fatigue_data
id serial PK
fatiguetest_id integer FK
data_sample_id integer
total_strain double precision
plastic_strain double precision
number_failure bigint
stress double precision
incremental_data_incrementaltest_id_fkey
hcf_fatigue_data_fatiguetest_id_fkey
lcf_fatigue_data_fatiguetest_id_fkey
incremental_test
id serial PK
sample_id integer FK
rate_hz double precision
run_number integer
number_break bigint
fatigue_test
id serial PK
sample_id integer FK
type measurement_type
treatment heat_treatment
rate_hz double precision
waveform waveform_shape
diameter_mm double precision
incremental_test_sample_id_fkey
fatigue_test_sample_id_fkey
sample_pedigree
id serial PK
sample_id character varying
sheet integer
sheet_e_gpa double precision
sheet_high_stress_ratio double precision
grade grade
heat_class double precision
treatment heat_treatment
ingot_size_cm double precision
diameter_mm double precision
reduction_percent double precision
length_mm double precision
primary_alpha_transverse double precision
secondary_alpha_transverse double precision
primary_alpha_longitudinal double precision
secondary_alpha_longitudinal double precision
al_percent double precision
v_percent double precision
fe_percent double precision
c_percent double precision
n_percent double precision
h_percent double precision
o_percent double precision
y_percent double precision
solution_temperature_degreesc double precision
solution_time_minutes double precision
solution_cooling_rate_type cooling_type
aging_temperature_degreesc double precision
aging_time_minutes double precision
aging_cooling_rate_type cooling_type
ys_mpa double precision
uts_mpa double precision
elongation_percent double precision
ra double precision
charpy double precision
hv double precision

Figure 1 (Top) Summary of test conditions in NIMS fatigue database for Ti6Al4V[11] and the associated Schema for NIMS [11] Ti6Al4V fatigue data using SQL database on Micloud [5]. Load control and strain control were under zero mean stress and strain respectively.

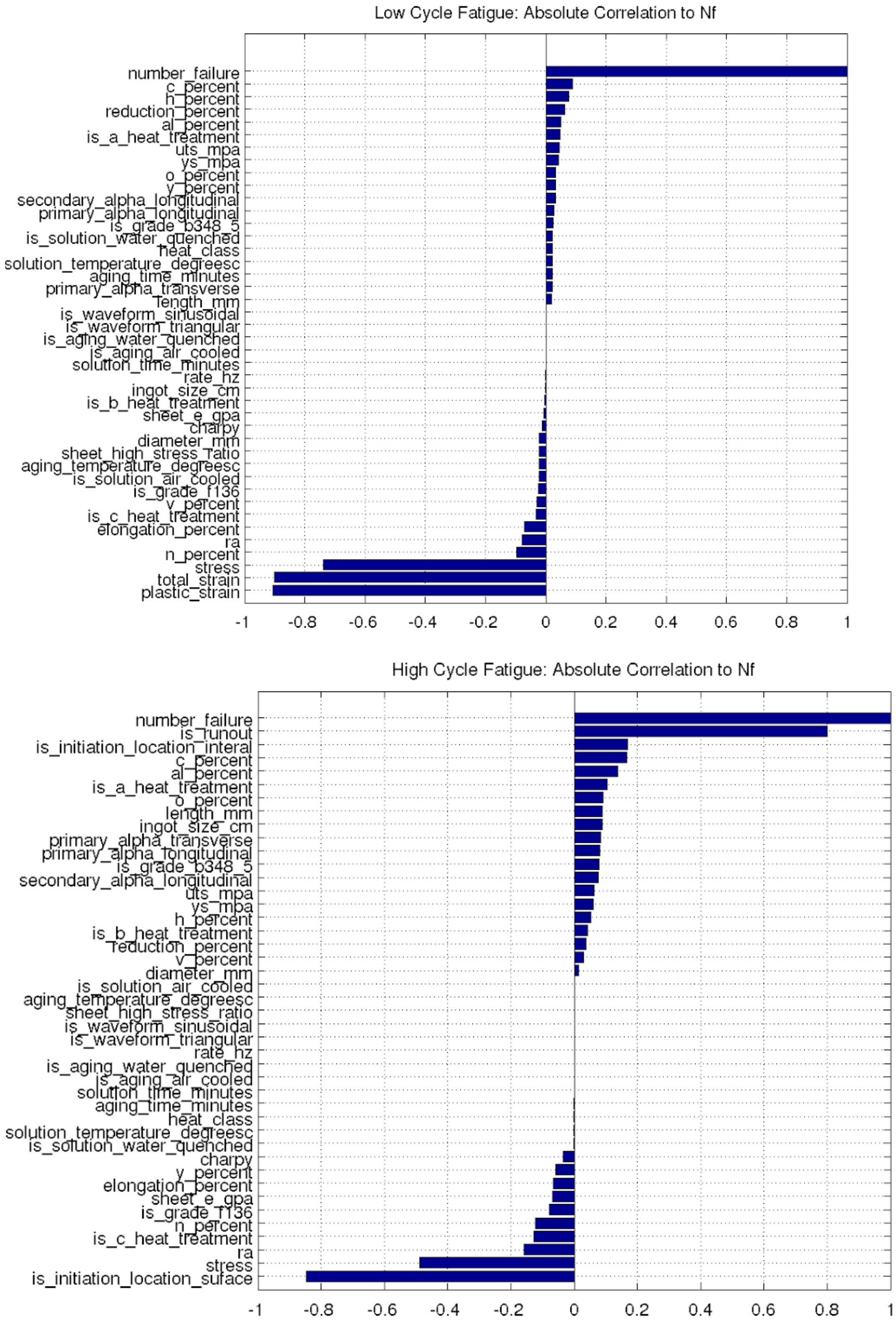


Figure 2 Spearman's rank correlation between fatigue life and processing, microstructure, and test conditions inputs from NIMS fatigue data [11] (top) strain control testing and (bottom) load control testing

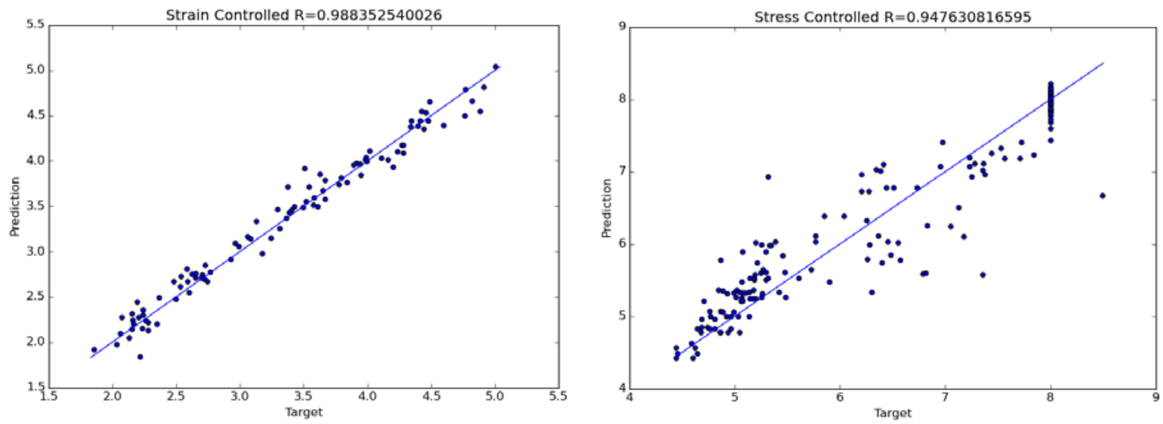


Figure 3 Ti6Al4V fatigue life values measured vs predicted for (left) strain control testing using and (right) load control. Axes are the log N_f..

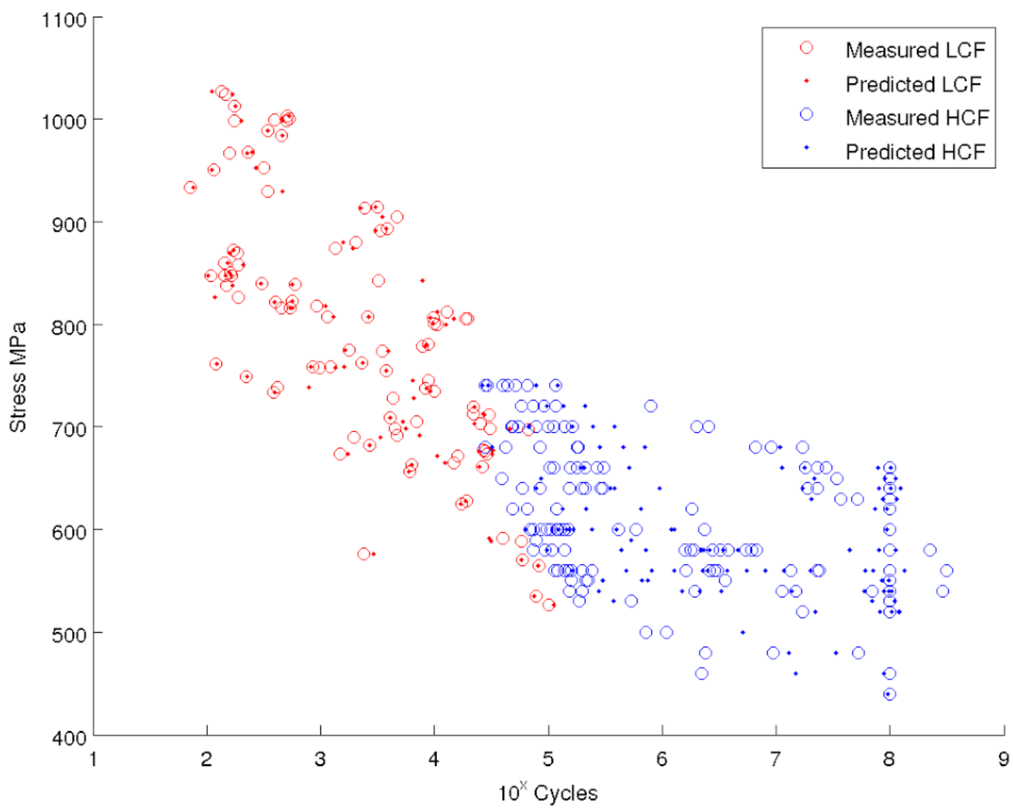


Figure 4 Stress-fatigue life for Ti6Al4V and Ti6Al4V ELI measured (open circles) vs predicted (dots) for strain control (red color) and load control (blue color) using Bayesian neural network.

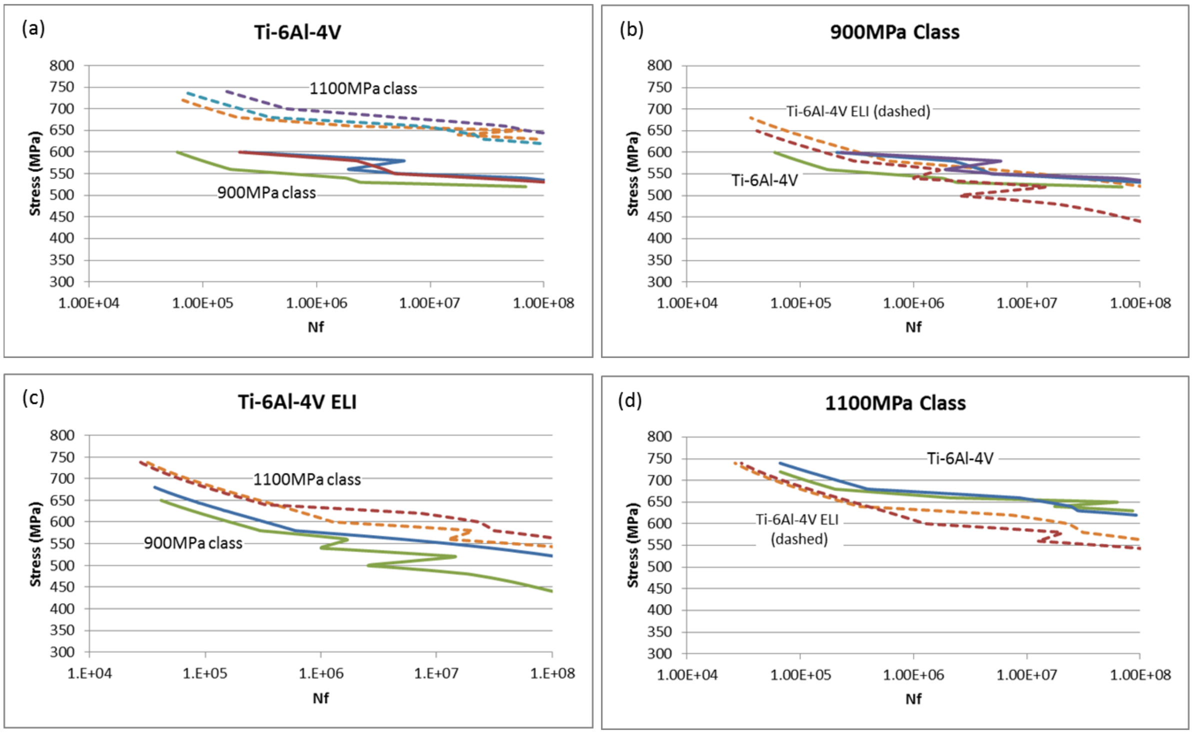


Figure 5 The effect of heat treatments (a,c) and chemical composition (b,d) on fatigue life using BNN predictions for Ti-6Al-4V and Ti-6Al-4VELI [11]**. Different colors reflect different lots of material.**

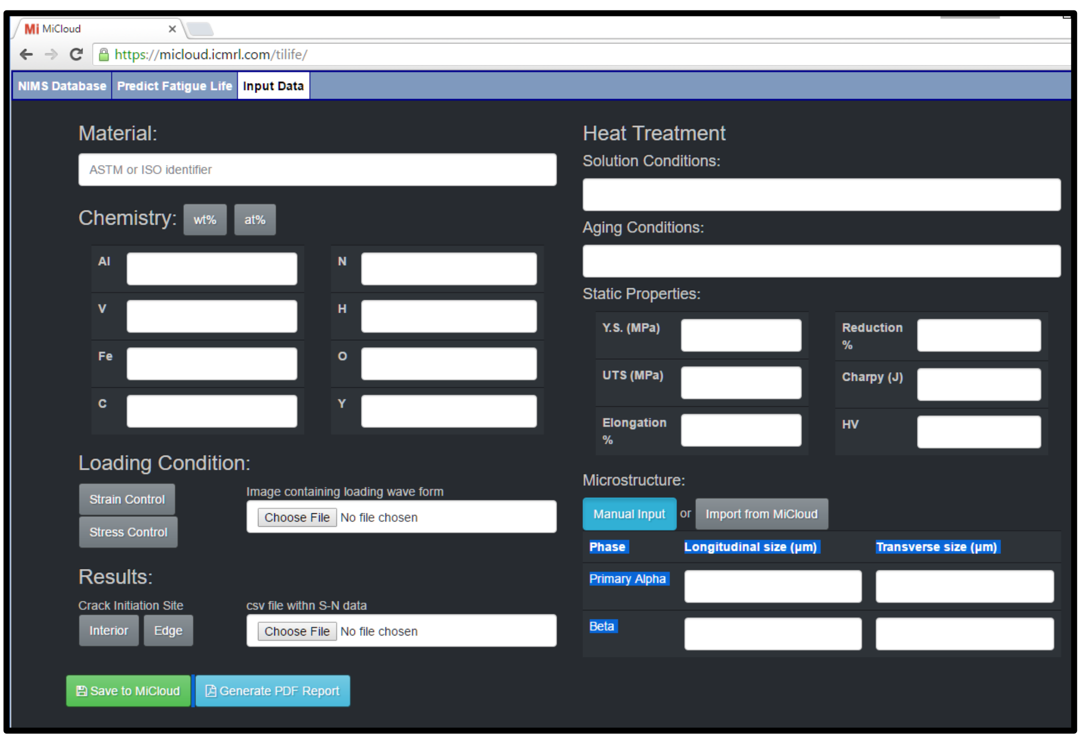


Figure 6. Website on MiCloud [6] **to allow global users sharing data for reuse in the BNN and comparison with current datasets.**

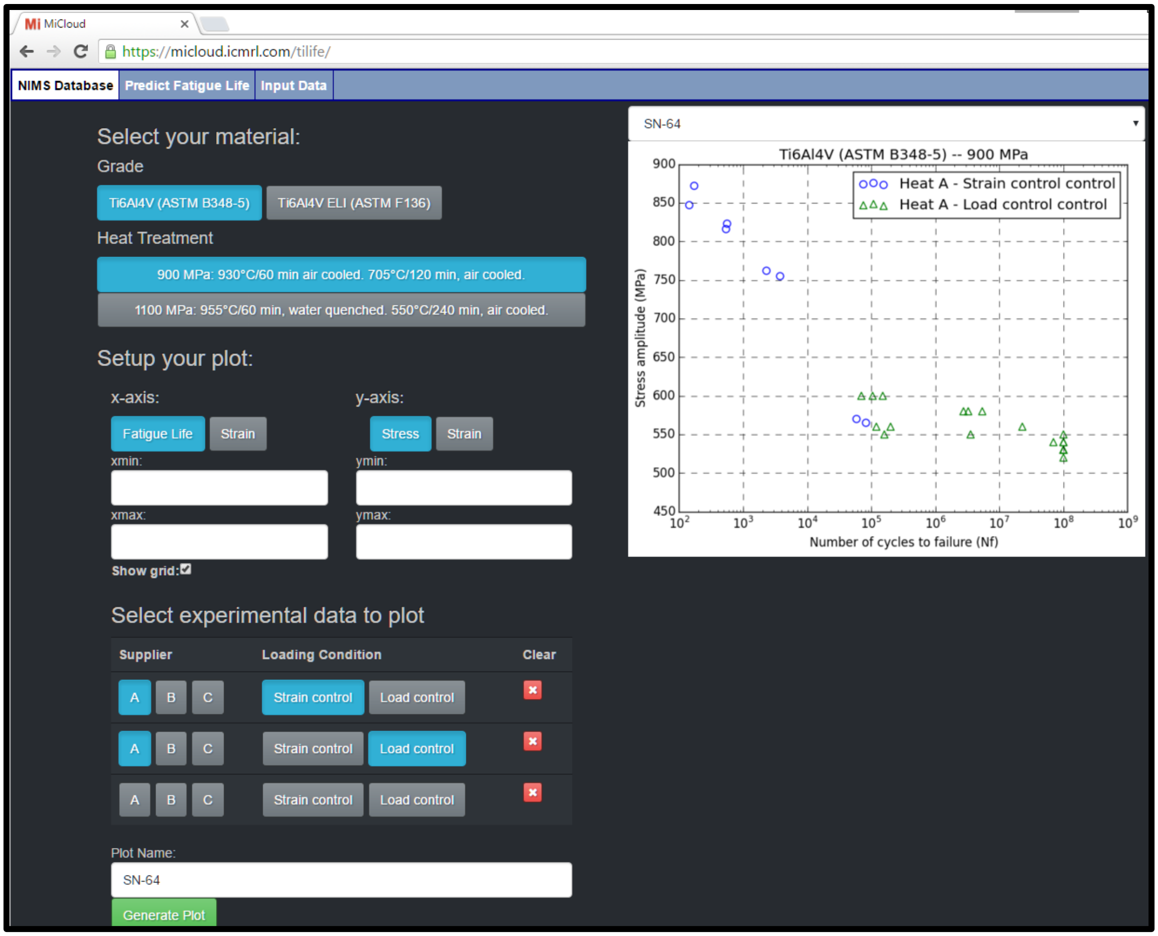
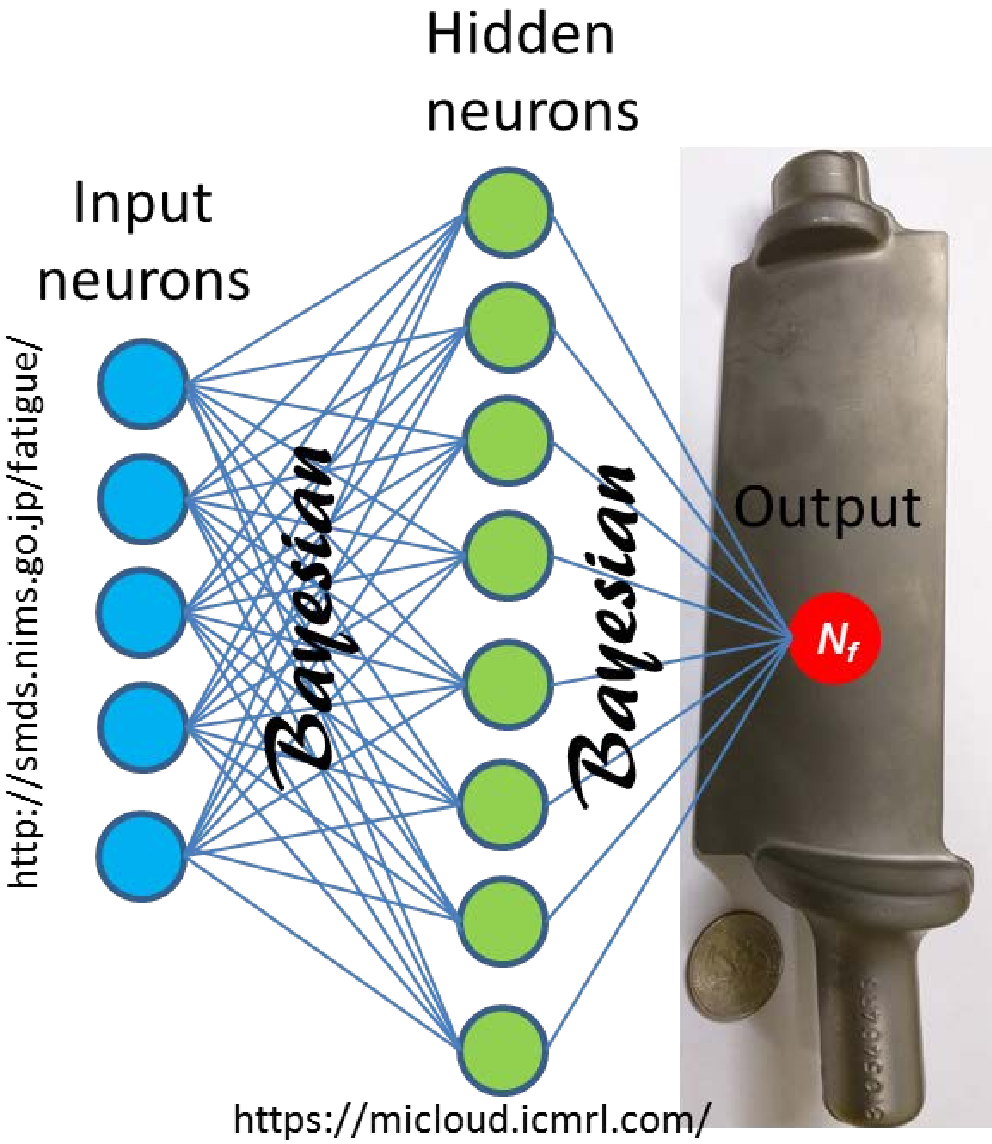


Figure 7 Hosting NIMS[1] **Fatigue dataset on MiClould** [6] **with a GUI to view S-N curves.**

GRAPHICAL ABSTRACT



SIGNIFICANCE OF THE WORK

Tomorrow’s scientific innovation will be fueled by open science run by researchers working together across the globe. While literature is full of material knowledge in 2D charts and text as projections and interpretations of multidimensional data, we are offering tomorrow’s solution by building a data-driven model for predicting fatigue life for turbine blade titanium using publically available. Instead of repeating expensive experiments to look at data limited by cost, we harvested the internet for data covering more than 10 years of research by other countries then we used machine learning for our model that can be shared, reused, and expanded using our cloud computing platform. Gaps in research can be mapped by the model while physical correlations are revealed.