



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

Thesis for the degree of Master of Mechanical Engineering

Flow Behavior Modeling of Al-Mg-Si Alloy Based on Machine Learning Strategies

Chen Wenning

Department of Mechanical Engineering

The Graduate School

Jeju National University

August 2023

Flow Behavior Modeling of Al-Mg-Si Alloy Based on Machine Learning Strategies

A Thesis submitted to the graduate school of
Jeju National University in partial fulfillment of
the requirements for the degree of Master of Mechanical Engineering
under the supervision of Prof. Dong Won Jung

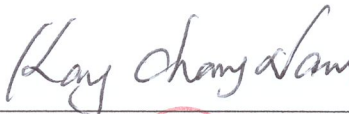

The thesis for the degree of Master of Mechanical Engineering

by Chen Wenning

has been approved by the dissertation committee.

June 15, 2023



Chair

Member



Member

Flow Behavior Modeling of Al-Mg-Si Alloy Based on Machine Learning Strategies

Chen Wenning

Department of Mechanical Engineering

The Graduate School

Jeju National University

Abstract

Aluminum (Al) and its alloys play an essential role in various application scenarios and occupy a large part of current industrial production and manufacturing. As an aviation alloy, the Al-Mg-Si 6xxx series alloys represented by AA6061 alloy are designed with excellent mechanical properties, which have been used in many high-strength applications. To guarantee the safety of the AA6061 alloy in manufacturing and application processes, the alloy's mechanical properties must be studied first. Flow behavior is a fundamental concept to describe the mechanical properties of materials, which can evaluate the strength, formability, fracture, service life, etc. In order to study the flow behavior of alloys quantitatively, constitutive modeling is a way generally adopted. Traditional modeling methods, such as phenomenological and physical models, describe flow behavior by establishing equations to build the response based on materials' deformation conditions.

However, these strategies are used to be inaccurate and complex.

Machine learning (ML), usually known as artificial intelligence (AI), sheds a new light on flow behavior modeling as its strong regression ability. By importing the number of experimental sample data with enough features to ML for training and learning, the models with high accuracy can be established to predict material constitutive behavior.

In the present work, taking AA6061-T6 alloy as a subject, the description performance under different designed deformation conditions of different ML strategies was studied and compared. Compression tests under temperatures 300 °C, 370 °C, 440 °C, 510 °C and strain rates 0.001 s⁻¹, 0.01 s⁻¹, 0.1 s⁻¹, 1 s⁻¹ were conducted on Gleeble-3800 thermal simulator respectively. Thermal processing maps were established, and the desired processing regions were determined as 300 °C and 0.001 s⁻¹, 300 °C and 1 s⁻¹, 440~510 °C, and 0.1~1 s⁻¹.

ML strategies like multiple linear (MLL), traditional ML nonlinear regression methods generalized additive model (GAM), regression tree (RT), random forest (RF), support vector regression (SVR), and multilayer perceptron (MLP) were selected to model the received flow stress data. After that, the performances of the established regression models were evaluated by statistical methods such as correlation coefficient (R), relative error (δ), average absolute relative error ($AARE$), and residual error (RE). Ultimately, the analysis reveals that the RT model shows the highest R (0.99995), lowest $AARE$ (0.00708), most minor δ distribution, and RE range, namely, the best description and prediction performance. Moreover, the performance of the MLP model is slightly worse than the RT model, but the model structure is simpler. This work checks the AA6061-T6 alloy constitutive modeling

abilities of different ML algorithms, which provides a reference for other materials' modeling works that need ML algorithms selection. Moreover, the deformation mechanism was analyzed by thermal process map and flow stress curves, providing instruction for Al-Mg-Si alloy industrial process.

기계 학습 정책 기반 Al-Mg-Si 합금 흐름 동작 모델링

진문녕

제주대학교 대학원 기계공학 과

요약

알루미늄과 관련된 두랄루민은 여러 가지 응용 상황에서 가장 중요한 역할을 하고 있으며 현재의 공업 생산과 제조 작업에서 아주 큰 공간을 차지하고 있다. AA6061 두랄루민을 대표로 된 Al-Mg-Si 6xxx 시리즈의 두랄루민은 항공용 두랄루민으로 훌륭한 기계적 성능을 가지고 있으므로 많은 고강도 응용 분야에 사용되었다. 제조와 응용 과정에서 AA6061 두랄루민의 안전성을 확보하기 위해서 반드시 두랄루민 역학 성능을 미리 연구해 봐야 한다. 유동적 행위는 재료 역학 성능을 설명할 때의 기본적인 개념으로 볼 수 있으며 재료의 강도, 성형성, 파손, 그리고 사용 수명 등을 평가하는 데 사용할 수 있다. 기본 구조 모델은 항상 열로이의 유동적 행위를 정기적으로 연구하는데 사용된다. 현상 모델과 물리 모델과 같은 전통적인 모델링 방법은 방정식 구축을 통해서 유동적인 행위를 설명하고 이를 통해 재료 변형 조건을 기반으로 된 호응을 구축하겠다. 그러나 이러한 방법들은 보통 정확성이 낮고 복잡하다.

항상 인공 지능(AI)으로도 불리는 기계 학습(ML)은 아주 강한 회귀 능력으로 인해 유동적 행위 모델링에 새로운 가능성을 가져온다. 충분한 특징을 가지고 있는 실험 샘플

데이터를 ML 에 많이 입력하여 훈련 및 학습을 통해 고정밀 모델을 구축하여 재료의 기본구조 행위를 예측하려고 한다.

본 연구는 AA6061-T6 알루미늄을 대상으로 전개하였으며 다양한 ML 전략이 서로 다른 디자인 변형 조건에서의 설명 성능을 비교 및 분석하였다. Gleeble-3800 열시뮬레이터에서 각각 300 °C, 370 °C, 440 °C, 510 °C 및 0.001 s⁻¹, 0.01 s⁻¹, 0.1 s⁻¹, 1 s⁻¹ 응변속율에서의 압축 실험을 실시했다. 열가공 그래프를 구축하고 이상적인 가공 구역을 300 °C 및 0.001 s⁻¹, 300°C 및 1 s⁻¹, 그리고 440~510 °C 및 0.1~1 s⁻¹ 로 확정하였다.

본 연구에서는 다원적 선성(MLL), 전통적인 비선형 회귀 방법 광의 가성모델(GAM), 회귀 트리(RT), 랜덤 포레스트(RF), 서포트 벡터 회귀(SVR), 그리고 다층적 퍼셉트론(MLP) 등과 같은 ML 전략을 이용하여 얻는 유동적인 응력 데이터에 대해 모델링을 하였다. 다음으로 상관 계수(R), 상대 오차(δ), 평균 절대 상대 오차(AARE), 그리고 잔차(RE) 등의 통계 방법을 사용하여 구축한 회귀 모델의 성능에 대해 평가하였다. 결국, 분석 결과에 따라 RT 모델이 드러난 가장 높은 상관 계수(R, 0.99995), 가장 낮은 AARE, 가장 좁은 δ 의 분포 그리고 RE 구간은 최상의 설명과 예측 성능임을 알 수 있다. 또한, 또한, MLP 모델의 성능은 RT 모델보다 약간 떨어지지만, 모델 구조는 더 간단하다. 본 연구는 서로 다른 ML 산법이 AA6061-T6 알루미늄에 대한 기본 구조 능력을 검증하였다. 이 외에 열가공 그래프와 유동적 응력 곡선을 통해 알루미늄 마그네슘 실리콘 알루미늄의 변형 메커니즘을 분석하였으며 이를 통해 알루미늄 마그네슘 알루미늄의 공업화 생산에 대해 지도적인 방안을 제공한다.

ACKNOWLEDGMENT

May this world peace.

Starting from the golden autumn of 2021 and ending in the midsummer of 2023, looking back on the two years, all that one can see is memories.

Thanks to supervisor Prof. Dongwon Jung. The professor's academic guidance has benefited me greatly. In addition, Professor Jung is kind to others, respects students' perspectives, and motivates us to practice according to our ideas. This environment motivates me to explore unknown fields and maintain my curiosity about academia, and I will maintain this positive attitude in my future research.

Thanks to my course teachers Prof. Ko Jin Hwan, Prof. Kim Nam Jin, Prof. Anil Kumar Khambampati, Prof. Parthiban Pazhamalai, and Prof. Muhammad Muqet Rehman. The high-quality courses provided by the professors have taught me much practical professional knowledge.

Thanks to my lab mates and good friends Krishna Singh Bhandari, Shahid Aziz, and Kosimov Nodirbek. I am very grateful for the mutual support from my colleagues during my two-year master's program. We are progressing together on the path of exploration and knowledge. Despite coming from different countries, we have forged precious friendships daily. We have gone through many things together, which will be a precious asset in my life.

Thanks to my friend Mohammed Awais Farooqi. Although we were not in the same

laboratory, Mr. Farooqi brought us joy and many happy moments.

Thanks to Mr. Chen Guangjin (my father), Hu Yinhuan (my mother), and my family. A thousand feet tall tree always remembers its deep roots in fertile soil. Thanks to my family for their meticulous care and support over the past 20 years, providing me with infinite love and warmth, allowing me to stand on their shoulders and see the wider world. The kindness of nurturing cannot be repaid, and only by constantly striving and becoming their pride, wishing my father, mother, and family good health, peace, and joy.

Thanks to my Li Sijia. On the evening of October 2, 2020, two lucky people began their story and journey. I am fortunate to be able to share my happiness, worries, thoughts, and daily life with a lively and outgoing girl. We study, go abroad, work, and travel together. We share the same joy and sadness. I am an intelligent and romantic person. Thank you, Miss Li, for your continuous encouragement and tolerance. I hope you are always young and beautiful, with no acne. I hope you do not gain weight after overeating. I hope you always have the money to buy what you want. Furthermore, hope our story is infinitely exciting, and our journey is infinitely long...

CONTENTS

LIST OF ABBREVIATIONS & SYMBOLS.....	XII
Abbreviations.....	XII
Symbols.....	XV
LIST OF FIGURES	XIX
LIST OF TABLES.....	XXIII
1 INTRODUCTION	1
1.1 Introduction of Al -Mg-Si 6xxx series Al alloy.....	1
1.2 Introduction of AA6061 alloy.....	5
1.3 Research about flow behavior constitutive modeling.....	7
1.3.1 Flow behavior research method	8
1.3.2 Research status of Al alloy thermal deformation flow behavior.....	11
1.4 Machine learning (ML).....	16
1.4.1 Development history	16
1.4.2 Research status of ML.....	18
1.4.3 Classification of Machine Learning	22
1.4.4 Process of ML	23
1.5 Research purpose and significance	24

2	EXPERIMENT PROCEDUR	27
2.1	Introduction.....	27
2.2	Material preparation.....	27
2.3	Cylinder uniaxial thermal compression test.....	28
2.4	OM observation	29
2.5	Flow stress of Al-Mg-Si AA6061-T6 alloy	31
2.6	Summary of the chapter.....	36
3	THERMAL PROCESSING MAP	38
3.1	Introduction of thermal processing map	38
3.2	Establishment of thermal processing map	38
3.2.1	Establishment of power dissipation map.....	39
3.2.2	Establishment of instability map.....	46
3.2.3	Establishment of thermal processing map.....	48
3.3	Summary of the chapter.....	51
4	LINEAR REGRESSION MODEL	52
4.1	MLL regression.....	52
4.2	Regression verification	55
5	GAM REGRESSION MODEL.....	58

5.1	GAM.....	58
5.1.1	Form of the GAM.....	58
5.1.2	Calculation principle of the AM: back-fitting algorithm	59
5.1.3	Calculation principle of the GAM: local scoring algorithm.....	61
5.2	Regression verification	63
6	SVR MODEL.....	65
6.1	SVM.....	65
6.1.1	Linear separability.....	65
6.1.2	Loss function.....	68
6.2	SVR.....	70
6.3	Regression verification	73
7	RT MODEL	77
7.1	DT.....	78
7.2	RT	80
7.3	Regression verification	82
8	RF MODEL.....	88
8.1	RF	89
8.2	Regression verification	91

9	MLP MODEL	94
9.1	MLP	95
9.1.1	Training of MLP.....	96
9.2	Regression verification	101
10	DISCUSSION.....	105
10.1	Performances evaluation of models by R	105
10.2	Performances evaluation of models by RE	107
10.3	Performances evaluation of models by δ distribution.....	109
10.4	Performances evaluation of models by $AARE$	111
10.5	Summary of the chapter	113
11	CONCLUSION	115
	REFERENCES.....	117
	Publications during the master's degree period.....	124

LIST OF ABBREVIATIONS & SYMBOLS

Abbreviations

Al	Aluminum
ML	Machine learning
AI	Artificial intelligence
MLL	Multiple linear
GAM	Generalized additive model
RT	Regression tree
RF	Random forest
SVR	Support vector regression
MLP	Multilayer perceptron
R	Correlation coefficient
$AARE$	Average absolute relative error
RE	Residual error
Cu	Copper
Mn	Manganese
Si	Silicon
Mg	Magnesium
Zn	Zinc
Cr	Chromium
Ti	Titanium
Fe	Ferrum
Pb	Lead

Bi	Bismuth
Zr	Zirconium
MJ-C	Modified Johnson-Cook model
A-T	Arrhenius-type model
MZ-A	Modified Zerilli-Armstrong model
OM	Optical microscope
SEM	Scanning electron microscope
TEM	Transmission electron microscope
BL	Bayesian learning
ANN	Artificial neural network
DAG	Directed acyclic graph
SLP	Single-layer perceptron
AL	Adaptive linear
BP	Back propagation
DNN	Deep learning neural network
CNN	Convolutional neural network
RNN	Recurrent neural network
DT	Decision tree
CART	Classification and regression tree
SLIQ	Supervised learning in quest
SVM	Support vector machine
SVC	Support vector clustering
DH	Dynamic hardening
DS	Dynamic softening
DRX	Dynamic recrystallization
DRV	Dynamic recovery
CDRX	Continuous dynamic recrystallization

DMM	Dynamic material modeling
AM	Additive model
RSS	Root sum square
KTT	Karush-Kuhn-Tucker

Symbols

*Some symbols appear more than once, their specific meaning follows from their context.

δ	Relative error
T	Temperature
σ	True stress
ε	True strain
$\dot{\varepsilon}$	Strain rate
C	Chemical composition
M	Initial microstructure
σ_e	Engineering stress
ε_e	Engineering strain
P	Input energy
G	Dissipation covariance
J	Dissipation covariance
m	Strain rate sensitivity exponent
J_{max}	Maximum dissipation covariance
η	Power dissipation factor
D	Dissipation function
ξ	Instability coefficient
\hat{y}	Target variable in MLL regression model
w_0	Intercept in MLL regression model
w_i	Regression coefficient in MLL regression model
x_i	Sample feature in MLL regression model
X	Sample feature matrix in MLL regression model
w	Column matrix of weights in MLL regression model

ϵ_i	Error in MLL regression model
$p(\epsilon_i)$	Error probability Gaussian distribution in MLL regression model
μ	Average error in MLL regression model
σ^2	Variance error in MLL regression model
$L(w)$	Likelihood function in MLL model
$J(w)$	Loss function in MLL model
α	Intercept in GAM model
$f_j(x_j)$	Smooth function in GAM model
$g(\mu)$	Link function in GAM model
s_i	Nonparametric smooth function in GAM model
n	Linear prediction value in GAM model
m	Iteration times in GAM model
ϵ	Threshold value in GAM model
V_i^{m-1}	variance of y at $\hat{\mu}_i$ in GAM model
X	Input data in SVM model
y	Learning objective in SVM model
X_i	Feature space in SVM model
w	Normal vector of the hyperplane in SVM model
b	Intercept of the hyperplane in SVM model
L	Lagrange equation in SVM model
$l_{0/1}$	0/1 loss function in SVM model
C	Constant in optimization function in SVM model
$l_{hinge}(z)$	Hinge loss function in SVM model
α_i	Lagrange multiplier
μ_i	Lagrange multiplier
C	Regularization constant in SVR model

l_ϵ	Loss function in SVR model
ξ_i	Slack variable in SVR model
$\hat{\mu}_i$	Lagrange multiplier
$\hat{\alpha}$	Lagrange multiplier
X	Processed data for training in standardization procedure
X_{input}	Import experimental data in standardization procedure
y	Mapping function in DT model
x	Sample point in DT model
c	Class in DT model
R	Sample space in DT model
k	Class after classification in DT model
t	Sample point in DT model
p	Class posterior probability in DT model
j	Classification feature in DT model
s	Threshold in DT model
L	Loss function in DT model
$p_{R_m^-}$	Posterior probability after binary classification in DT model
$p_{R_m^+}$	Posterior probability after binary classification in DT model
$E(R_m^-)$	Sample space after binary classification in DT model
$E(R_m^+)$	Sample space after binary classification in DT model
$E(R_m)$	Sample space before the classification in DT model
p	Final predicted values in <i>MSE</i> equation
y	Expected output data in <i>MSE</i> equation
y	Scalar output in MLP model
f	Activation function in MLP model
w	Weight in MLP model

x	Input in MLP model
b	Basis in MLP model
Δp	Gradient in gradient descend algorithm
η	learning rate in MLP model
k	Hidden neuron number in the first hidden layer of MLP model
m	Input neuron number of MLP model
l	Hidden neuron number in the second hidden layer of MLP model
s	Number of the input data
$L(w, b)$	Loss function in MLP model
c	Add value before activation function operation in MLP model
$Cov(Y, P)$	Covariance of Y and P
Y	Set of the experimental flow stress data in R equation
P	Set of the predicted flow stress data in R equation
$Var Y $	Variance of Y in R equation
$Var P $	Variance of P in R equation
y_0	Constant in Gaussian distribution equation
A	Constant in Gaussian distribution equation
ω	Standard deviation in Gaussian distribution equation
μ	Average value in Gaussian distribution equation
y	Probability in Gaussian distribution equation

LIST OF FIGURES

Fig.1. 1. Equilibrium phase diagrams of Al–Mg–Si alloys: (a) Al–xMg–2.4Si alloy, (b) Al–7.5Mg–xSi alloy.	3
Fig.1. 2. Industry applications of metal flow behavior.....	7
Fig.1. 3. Procedure from experimental analysis to production (tensile test as example). 8	
Fig.1. 4. Factors and interrelationships affecting the flow stress of metal thermal deformation.....	12
Fig.1. 5. ML classified by different classification standards.....	23
Fig.1. 6. Data processing and evaluation in ML.	24
Fig.2. 1. (a) Gleeble-3800 simulator, (b) Visual Work Window, (c) Samples after compression, (d) Processing map.	29
Fig.2. 2. Geometrical shape deformation and corresponding pictures.	31
Fig.2. 3. Flow stress curves under experimental deformation conditions: (a) $\dot{\epsilon}= 0.001 \text{ s}^{-1}$, (b) $\dot{\epsilon}= 0.01 \text{ s}^{-1}$, (c) $\dot{\epsilon}= 0.1 \text{ s}^{-1}$, (d) $\dot{\epsilon}= 1 \text{ s}^{-1}$	32
Fig.3. 1. $\ln\sigma - \ln\dot{\epsilon}$ plots and cubic polynomial function fitting results under each strain: (a) 0.1; (b) 0.2; (c) 0.3; (d) 0.4; (e) 0.5; (f) 0.6; (g) 0.7; (h) 0.8.....	41
Fig.3. 2. Dissipation maps under experimental conditions divided by true strain: (a) 0.1; (b) 0.2; (c) 0.3; (d) 0.4; (e) 0.5; (f) 0.6; (g) 0.7; (h) 0.8.	43
Fig.3. 3. 3D power dissipation maps of AA6061-T6 alloy: (a) along temperature; (b) along strain rate; (3) along strain.....	44

Fig.3. 4. Instability maps under experimental conditions divided by true strain: (a) 0.1; (b) 0.2; (c) 0.3; (d) 0.4; (e) 0.5; (f) 0.6; (g) 0.7; (h) 0.8.....	48
Fig.3. 5. Thermal processing maps under experimental conditions divided by true strain: (a) 0.1; (b) 0.2; (c) 0.3; (d) 0.4; (e) 0.5; (f) 0.6; (g) 0.7; (h) 0.8.....	49
Fig.3. 6. 3D thermal processing maps of AA6061-T6 alloy: (a) along temperature; (b) along strain rate; (3) along strain.....	50
Fig.4. 1. Operation mechanism flow chart of the MLL regression algorithm.....	55
Fig.4. 2. Regression verification of the MLL model at different experimental deformation conditions, divided by strain: (a) 0.001 s^{-1} , (b) 0.01 s^{-1} , (c) 0.1 s^{-1} , and (d) 1 s^{-1}	57
Fig.5. 1. Operation mechanism flow chart of the GAM regression algorithm.....	63
Fig.5. 2. Regression verification of the GAM model at different experimental deformation conditions, divided by strain: (a) 0.001 s^{-1} , (b) 0.01 s^{-1} , (c) 0.1 s^{-1} , and (d) 1 s^{-1}	64
Fig.6. 1. The difference between SVM and SVR.....	71
Fig.6. 2. Operation mechanism flow chart of the SVR algorithm.....	73
Fig.6. 3. Regression verification of the SVR model at different experimental deformation conditions, divided by strain: (a) 0.001 s^{-1} , (b) 0.01 s^{-1} , (c) 0.1 s^{-1} , and (d) 1 s^{-1}	76
Fig.7. 1. Types of DT.....	77
Fig.7. 2. Classification of a binary tree.	79

Fig.7. 3. Regression process of RT (three times growing).....	81
Fig.7. 4. Operation mechanism flow chart of DT (RT).....	82
Fig.7. 5. The <i>MSE</i> variation with the split times.....	83
Fig.7. 6. The structure of the RT model.	84
Fig.7. 7. Regression verification of the RT model at different experimental deformation conditions, divided by strain: (a) 0.001 s ⁻¹ , (b) 0.01 s ⁻¹ , (c) 0.1 s ⁻¹ , and (d) 1 s ⁻¹	87
Fig.8. 1. Data random selection and feature random selection in RF.....	89
Fig.8. 2. Operation mechanism flow chart of RF.	91
Fig.8. 3. The structure of the first RT and the structure of the second RT	92
Fig.8. 4. Regression verification of the RT model at different experimental deformation conditions, divided by strain: (a) 0.001 s ⁻¹ , (b) 0.01 s ⁻¹ , (c) 0.1 s ⁻¹ , and (d) 1 s ⁻¹	93
Fig.9. 1. Perceptron (b), inspired by neurocyte (a).....	94
Fig.9. 2. Topological structure of a two hidden layer MLP.	97
Fig.9. 3. Operation mechanism flow chart of back propagation MLP.	100
Fig.9. 4. The <i>MSE</i> and <i>R</i> variation along with the hidden layer neuron number.....	101
Fig.9. 5. <i>MSE</i> variation during the optimization process.	102
Fig.9. 6. The fitting performances of the training and test data groups.....	102
Fig.9. 7. Regression verification of the MLP model at different experimental deformation conditions, divided by strain: (a) 0.001 s ⁻¹ , (b) 0.01 s ⁻¹ , (c) 0.1 s ⁻¹ , and (d) 1 s ⁻¹	103
Fig.10. 1. Correlation plots of established models: (a) MLL, (b) GAM, (c) SVR, (d) RT,	

(e) RF, (f) MLP.....	106
Fig.10. 2. Residual error plots of established models: (a) MLL, (b) GAM, (c) SVR, (d) RT, (e) RF, (f) MLP.....	108
Fig.10. 3. δ distribution plots of established models: (a) MLL, (b) GAM, (c) SVR, (d) RT, (e) RF, (f) MLP.....	110
Fig.10. 4. <i>AAREs</i> of established models (a) in different temperatures, and (b) in different strain rates.....	112

LIST OF TABLES

Table 1-1 Added elements, characteristics and applications of different series Al alloys.	1
Table 1-2 Types, conditions and applications of 6xxx Al alloy.....	3
Table 1-3 The regulation chemical composition of AA6061 alloy (wt. %).	5
Table 1-4 The regulation mechanical properties of AA6061 alloy (MPa).	5
Table 1-5 Phenomenological models adopted most widely.	13
Table 1-6 Models and their performances in previous works.	13
Table 1-7 Model performances from previous literatures.	15
Table 2-1 Chemical compositions of cylinder samples.....	28
Table 2-2 Mechanical properties of cylinder samples.....	28
Table 2-3 Automatic polishing process.	30
Table 2-4 Dominate deformation mechanisms under experimental conditions.	35
Table 3-1 m values and η values under different conditions.	42
Table 3-2 $\xi\varepsilon$ values under different conditions.	46
Table 5-1 Typical link functions in GAM.	61
Table 6-1 α values in SVM model.	74
Table 6-2 Support vectors of the SVM model.	74
Table 7-1 The parameters of the RT model.	84
Table 10-1 <i>AAREs</i> of different established models.	111

1 INTRODUCTION

1.1 Introduction of Al -Mg-Si 6xxx series Al alloy

Al and its alloys are regarded as the most economical and practical material in various application fields. Al alloys are instructed by strengthening Al as structure material and are usually created by adding alloy elements to the Al matrix and applying heat treatment. Based on retaining the lightweight, Al and Al alloys receive the characteristics of high specific strength, high specific stiffness, and excellent casting and deformation properties, which contribute to the application potential as structure material. Consequently, Al alloys are widely used in the fields of transportation, aerospace, civil engineering, mechanical engineering, electromechanical, and so on. The processing methods of Al can be divided into two types: casting processing and deformation processing. Specifically, the casting Al alloys are produced by iron mold, sand mold, investment mold, die-casting method, etc., melting the gradients in proportion. The first step of producing deformation alloy is casting. After that, the received as-casted billet is processed into plastic products by plastic methods like tension, extrusion, forge, rolling and so on. There are eight series of Al alloys according to different element additions. See **Table 1-1**.

Table 1-1 Added elements, characteristics and applications of different series Al alloys.

Series	Added elements	Characteristics	Applications
1xxx	Pure Al (>99%)	High formability, corrosion-resist	Hardware parts, conductor
2xxx	Cu	High stiffness, high strength	Aviation industry

3xxx	Mn	Excellent antirust performance	Damp environment like auto bottom, refrigerator, air-conditioner
4xxx	Si	Corrasion-resist, low melt point	Material for weld, forge, mechanical parts, architecture
5xxx	Mg	High elongation ratio; Low density	Decorative items, aircraft fuel tank
6xxx	Mg, Si	Corrasion, oxidation-resist, excellent interface and process performances	Valves, pistons, connectors, camera parts, aircraft parts
7xxx	Zn	High strength, high stiffness, high wear resistance	Aircraft structural components, future manufacturing material
8xxx	Beyond the above	High thermal conductivity	Heat dissipation components, bottle lips

As shown in the table, 6xxx Al alloy is also known as Al-Mg-Si alloy, which uses Mg₂Si as the strengthening phase. Al-Mg-Si Al alloys with different mechanical properties will receive by adding different amounts of Mg and Si elements during the casting process. **Fig. 1. 1** displays the phase diagrams of die-cast Al-Mg-Si alloys [1]. The eutectic temperature of the alloy is around 550 °C. The Mg₂Si strengthening phase immediately formed with the addition of elements (second phase strengthening).

Moreover, with a large amount of Mg element or a small amount of Si element, AlMg will precipitate at low temperatures as a solid-solution strengthening. Due to the atomic size of Mg (3.20Å) is bigger than Al (2.86Å), the solid-solution strengthening effect is very strong. In addition, 6xxx alloys belong to heat treatment strengthened Al alloy. The stiffness and strength can be optimized by aging hardening treatment However, the aging hardening sequence of 6xxx Al alloy is complex and varies with the heat treatment process and alloy composition. Still, most of them follow [2]:

SSSS → solute clusters → GP zones → metastable β'' → metastable β', U1, U2, and B' → β and Si.

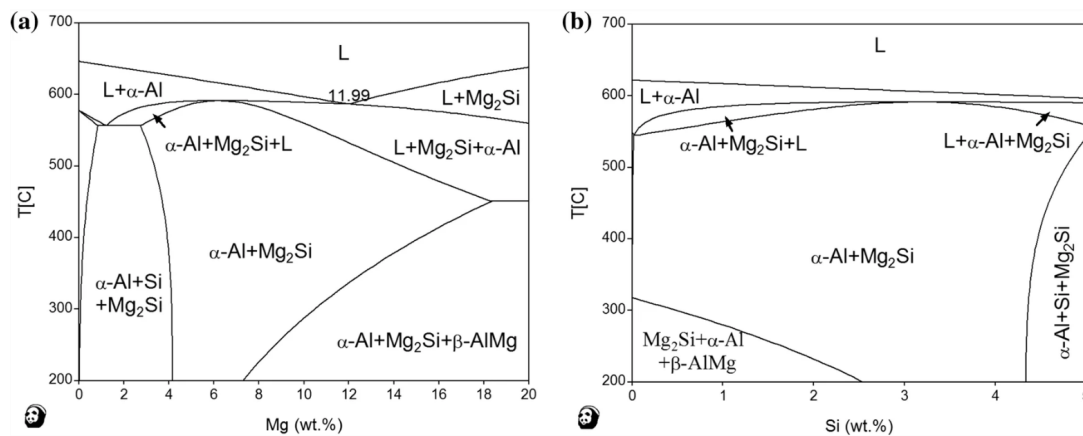


Fig.1. 1. Equilibrium phase diagrams of Al–Mg–Si alloys: (a) Al–xMg–2.4Si alloy, (b) Al–7.5Mg–xSi alloy.

The characteristics of corrosion-resist, high strength, excellent welding properties, and good homogeneity guarantee the alloys' high market share, especially in engineering applications. However, according to the proportion of added elements, 6xxx alloys are divided into different types. **Table 1-2** shows the different types of 6xxx alloys and their typical conditions and applications.

Table 1-2 Types, conditions and applications of 6xxx Al alloy.

Name	Types	Process conditions*	Applications
6005	Wire stock, bar, extruded tube	T1, T5	Structure components that require higher strength than 6063 alloy such as TV antenna, ladder,
6009 6010	Plate	T4, T6	Car body panels
	Plate	O, T4, T6	
	Heavy-gauge sheeting	O, T451, T651	
	Stretching tube	O, T4, T6, T4510, T4511	
6061	Extrusion line, bar, shape, and tube	T51, T6, T6510, T6511	Generally used as various industrial structure components that require certain strength, high corrosion resistance, good weldability in manufacturing rolling stock, tram, ship, tower architecture, lorry, hardware, etc.
	Tube	T6	
	Rolled or extruded structural shape	T6	
	Cold processed bar	O, H13, T4, T541, T6, T651	
	Cold processed	O, H13, T4, T6,	

	line	T89, T913, T94	
	Rivet wire	T6	
	Forged piece	F, T6, T652	
6063	Stretching tube	O, T4, T6, T83, T831, T832	Architectural shapes, irrigation pipes, extruded materials for vehicles, fences, elevators, furniture, platforms, etc., as well as decorative components of different colors used in light industry departments, ships, airplanes, buildings, etc.
	Extrusion line, bar, shape, and tube	O, T1, T4, T5, T52, T6	
	Tube	T6	
6066	Stretching tube	O, T4, T42, T6, T62	Forgings and extruded materials for welded structures.
	Extrusion line, bar, shape, and tube	O, T4, T4510, T4511, T42, T6, T6510, T6511, T62	
	Forged piece	F, T6	
6070	Extrusion line, bar, shape, and tube	O, T4, 4511, T6, T6511, T62, F, T6	Extruded materials and machine parts for heavy-duty welded structures and the automotive industry, including conduits, marine components, cable towers, bridges, pipes, etc.
	Forged piece	F, T6	
6101	Extrusion line, bar, shape, and tube	T6, T61, T63, T64, T65, H111	High strength bars, heat dissipation devices, conductive materials, high-strength busbars, etc. for buses.
	Tube	T6, T61, T63, T64, T65, H111	
	Rolled or extruded structural shape	T6, T61, T63, T64, T65, H111	
6151	Forged piece	F, T6, T652	Used for machine parts and components, forging crankshaft parts, for both high strength, good malleability, and good corrosion resistance.
6201	Cold processed line	T81	Used for high strength conductivity line and bar.
6205	Plate	T1, T5	High impact extruded parts, thick plates, and pedals.
	Extruded material	T1, T5	
6262	Stretching tube	T2, T6, T62, T9	High stress threaded mechanical components that require higher corrosion resistance than 2011 and 2017 alloy, and with excellent machinability.
	Extrusion line, bar, shape, and tube	T6, T6510, T6511, T62	
	Cold processed line	T6, T651, T62, T9	
	Cold processed bar	T6, T9	
6351	Extrusion line, bar, shape, and tube	T1, T4, T5, T51, T54, T6	Extruded shapes, transportation pipelines for oil, water, etc., and extruded structural components for vehicles.
6463	Extrusion line, shape, and tube	T1, T5, T6, T62	Automotive decorative parts with bright surfaces after anodizing treatment, as well as various instruments and building shapes.
6A02	Plate	O, T4, T6	Mechanical parts that require high

Heavy-gauge sheeting	O, T4, T451, T6, T651	corrosion resistance and high plasticity, complex shaped die forgings and forgings, aircraft engine parts.
Tube, bar, shape	O, T4, T451, T6, T651	
Forged piece	F, T6	

*The definition of process conditions refers to [3].

1.2 Introduction of AA6061 alloy

AA6061 is one of the most widely used Al-Mg-Si alloys, a high-quality Al alloy that has undergone heat treatment and a pre-stretching process. Except for the applications the **Table 2** shown, the alloy has been adopted in communication, electrical fixtures, aerospace fixtures, and precision machining. **Table 1-3** and **Table 1-4** show the chemical compositions and mechanical properties regulation of AA6061 alloy.

Table 1-1 The regulation chemical composition of AA6061 alloy (wt. %).

Cu	Mn	Mg	Zn	Cr	Ti	Si	Fe	Al
0.15-0.4	0.15	0.8-1.2	0.25	0.04-0.35	0.15	0.4-0.8	≤0.7	Balance

Table 1-2 The regulation mechanical properties of AA6061 alloy (MPa).

Tensile strength	Compression strength	Elastic modulus (GPa)	Ultimate bending strength	Bending yield strength
205	55.2	68.9	228	103

In addition to the main alloying elements Mg and Si, other elements can also be added to 6061 alloys to improve different mechanical properties: the machinability can be improved by adding Pb or Bi; Zr and Ti can be used to refine grain and control recrystallization structure; the Ti and Fe negative effect on conductivity can be offset by adding a small amount of Cu; without reducing its corrosion resistance the alloy's strength can improve through adding the amount of Zn or Cu; if there is some Cr or Mn as elements, the negative

effect of Fe will be neutralized.

Due to the broad application prospects, the research of AA6061 alloy has always been one of the hotspots in materials science. First, Lu et al. [4] achieved the alloy design of AA6061 alloy by adjusting the content and proportion of alloying elements to improve its mechanical properties and corrosion resistance. Researchers also study the processing technology of AA6061 alloy and improve its deformation capacity and plasticity by optimizing process parameters such as alloy casting, hot working, and cold working to obtain higher mechanical properties [5]. In addition, functional surface modification is a hot topic as well. Researchers improve the surface properties of AA6061 alloy, such as corrosion resistance, wear resistance, and lubricity, by surface modification techniques such as surface spraying and chemical treatment [6]. AA6061 alloy sheet is always adopted as important layers for composites because of its friendly cost and excellent mechanical properties. Researchers studied the composite properties of AA6061 alloy with fiber-reinforced composite materials to obtain higher mechanical properties and strength, such as preparing AA6061 alloy composite wheels with higher impact resistance, deformation resistance, and corrosion resistance [7], [8].

As a maturely studied alloy, a significant part of the research direction of AA6061 alloy is about the thermal deformation mechanism, especially on constitutive models during alloy deformation, combined with the emerging numerical simulation technology. However, with the continuous exploration of methods and techniques, constitutive modeling can be further subdivided into many methods.

1.3 Research about flow behavior constitutive modeling

Flow behavior is a standard method to study the mechanical properties of materials, which is expressed as the curve made in the coordinate system with stress as the ordinate and strain as the abscissa. As basic research for material mechanical properties, flow behavior modeling is the foundation for all industries (see Fig.1. 2).

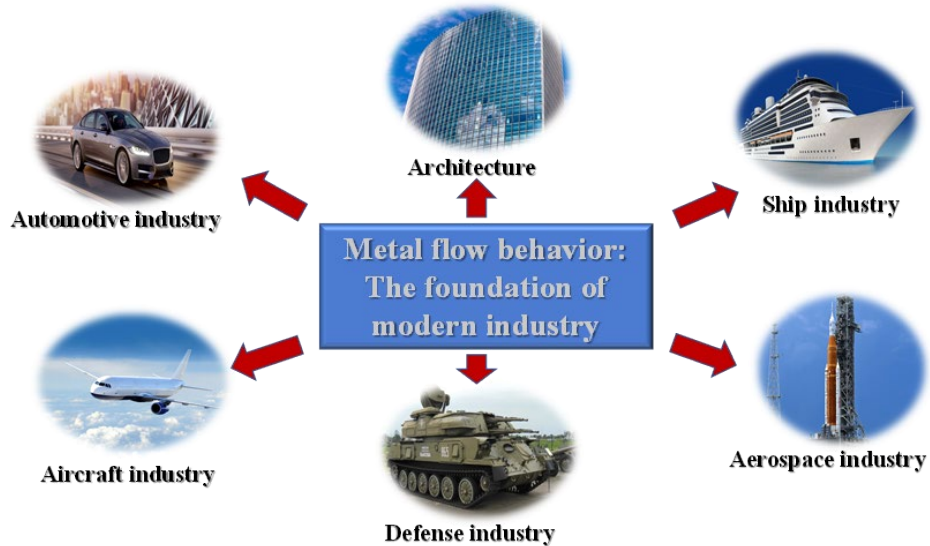


Fig.1. 2. Industry applications of metal flow behavior.

Flow behavior plays a vital role in engineering applications, as it can be used to evaluate the materials' plasticity and fatigue performance [9], [10]. In the process of material selection and design, the materials' suitability and service life need to be determined based on the magnitude and type of flow behavior. In addition, flow behavior is also an essential parameter in mechanical processing and manufacturing processes, as it can affect the material's processing performance and the dimensional accuracy of the workpiece. Additionally, with the prosperity of numerical simulation, constitutive models are

programmed widely to predict material deformation characteristics. In industrial processes, the flow behavior analysis, modeling, simulation, and production consist of a whole procedure set, as seen in Fig.1. 3.

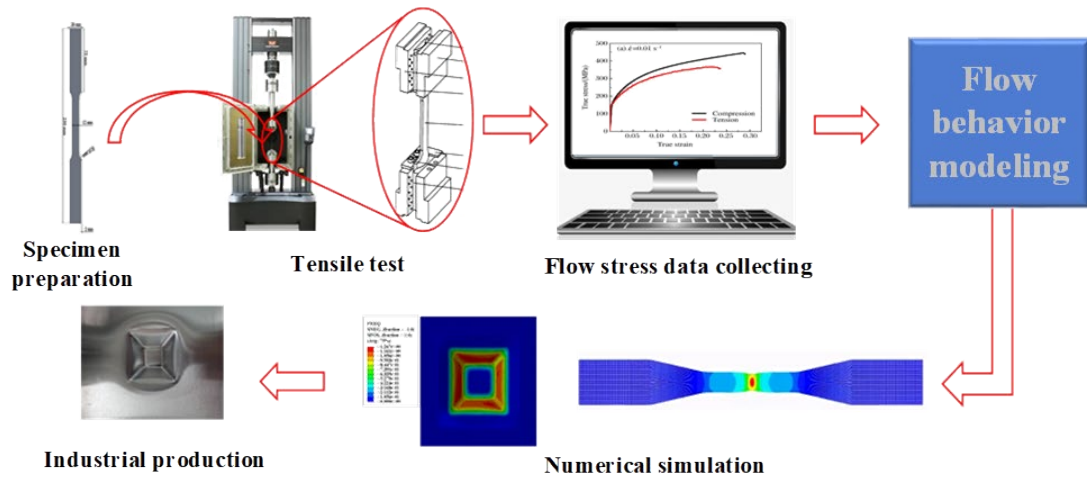


Fig.1. 3. Procedure from experimental analysis to production (tensile test as example).

1.3.1 Flow behavior research method

It is necessary to utilize some basic experimental methods when studying materials' flow behavior. Three methods are mainly adopted: uniaxial compression, tension, and torsion. These methods provide researchers thermal deformation behavior of materials under different deformation conditions before the optimization design and implementation of all kinds of metal forming processes like forging, drawing, extrusion, rolling, and so on [11], [12]. When plastic forming researchers work on new material development and classification, the relevant formability indicator and materials' thermal deformation characteristics can be established with the help of these methods. The flow stress–true strain data (the essence of the model) obtained from the above experiments can be used to analyze practical material

deformation behavior, supporting the formulation of processing technology in industrial production and exploration of the essence.

1.3.1.1 Tensile test

In most cases, the tensile test is utilized to simulate extrusion and tension. The tensile flow behavior directly relates to the ultimate pressure during the extrusion, which can evaluate the deformation temperature rise effect, the load, and life of die, friction effect, etc. The area reduction reflects the true high-temperature plasticity when the material is under simple stress conditions. However, some drawbacks exist: 1. Compared with other methods, the strain rate of the tensile test is relatively low. The strain range is usually at $10^{-6} \text{ s}^{-1} \sim 10 \text{ s}^{-1}$, and the strain rate of 10^3 s^{-1} and above only can be achieved in some special tensile equipment. When the deformation reaches a certain level, a necking phenomenon will occur. The equivalent strain rate increases sharply with the appearance of necking, resulting in the abnormal deformation of flow stress. Accordingly, the strain effect on flow stress is hard to capture; 2. Due to the necking effect, deformation and softening mechanism, and relationship among deformation condition, performance, and microstructure become more challenging to study; 3. The strain value under necking is far below the strain under industrial hot processing.

1.3.1.2 Torsion test

Compared with other tests, the torsion test is able to guarantee material deformation at an extensive strain range and constant strain rate but with no instability. There is no geometrical

softening and hydrostatic pressure during deformation. The torsion test is widely adopted in measuring the material's formability and flow stress under large strain deformation conditions. Additionally, the material deformation keeps uniform deformation strain rate and torsion strain along the axial direction and is limited in specimen gauge dimension range. No obvious geometrical deformation exists, namely, no ununiform unstable deformation phenomenon like drum and necking. There are also two shortcomings of the torsion test: 1. The strain, strain rate, and stress show linear characteristics along the axial direction, which leads to difficulty in data explanation; 2. To equipoise the geometrical change resulting from the anisotropy and texture change during material torsion, a little axial load is required to add at the fixed end. As a consequence, the stress condition will be more complicated.

1.3.1.3 Compression test

Compression tests can be divided into isothermal and non-isothermal compression according to temperature deformation conditions. Based on strain types, compression can be divided into plane strain compression and uniaxial compression. The non-isothermal compression test is the closest situation to the actual forging process. However, the flow stress is usually higher than the actual value because of the external influence. Plane strain compression suits isotropic material and plate hot rolling simulation work. However, because of the external influence, the flow stress is usually higher than the true value. Besides, the geometrical softening phenomenon and the uncertainty of drum shape and friction condition make this test unsuitable for thermal constitutive relationship research. In this work, uniaxial

isothermal compression tests were adopted. This kind of test can measure materials' true stress-strain relationship during thermal deformation in a large strain rate range directly, whatever strain rate sensitive or insensitive material. However, due to friction between the specimen and the anvil, ununiform deformation will occur when the strain exceeds a particular value.

As a consequence, the deformation rate will slow down, and the original uniaxial stress state will also become a complex triaxial stress state. Under this condition, side surface cracks will grow prematurely because of tensile stress. Fortunately, this problem can be settled by improving the surface lubrication condition, and the lateral drum phenomenon can be greatly eliminated when the lubrication effect is good.

1.3.2 Research status of Al alloy thermal deformation flow behavior

Thermal deformation flow behavior is one of the essential performances of metal materials under high temperatures. It is influenced by chemical composition, strain rate, deformation degree, temperature, and interior microstructure evolution. No matter whether in metal plastic deformation theory research or formulation of a reasonable hot processing process, thermal deformation flow behavior matters significantly. The flow stress (σ) of Al and its alloys entirely depends on deformation condition (temperature T , strain ε , and strain rate $\dot{\varepsilon}$), chemical composition C , and initial microstructure M . The interrelationships of these parameters are displayed in **Fig.1. 4.** and can be expressed by:

$$\sigma = f(T, \varepsilon, \dot{\varepsilon}, C, M) \quad (1-1)$$

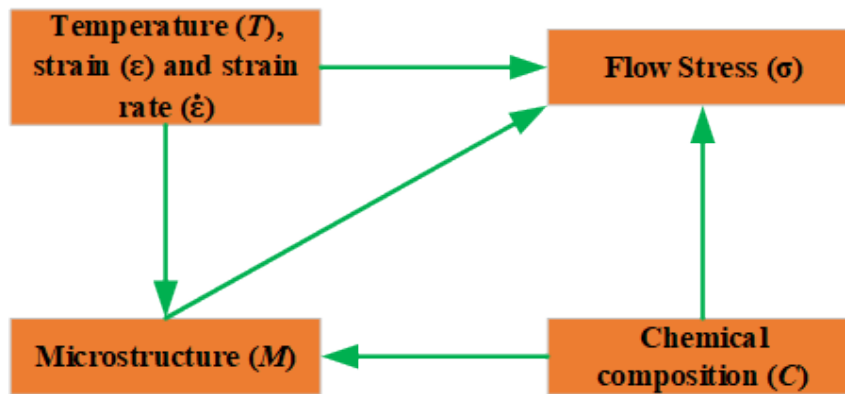


Fig.1. 4. Factors and interrelationships affecting the flow stress of metal thermal deformation.

Up to now, various models have been derived to describe the flow behavior of alloys. The traditional way uses constitutive equations that comprehensively take influential factors as independent variables to predict σ . In general, constitutive equations can be divided into two types: physical models and phenomenological models.

1.3.2.1 Phenomenological model

Generally, deformation conditions and chemical composition also affect the evolution of microstructure. Consequently, the flow behavior of Al alloy is a complex system. In the practical deformation process, the chemical composition keeps constant, which can be substituted by the material coefficient. Moreover, the microstructure during deformation is influenced by the deformation condition. The **Eq. (1-1)** can be simplified to the following equation:

$$\sigma = f(T, \varepsilon, \dot{\varepsilon},) = f_1(T)f_2(\varepsilon)f_3(\dot{\varepsilon}) \quad (1-2)$$

After years of development, many phenomenological constitutive models and their

modifications have been proposed. The following table gives some phenomenological model samples people have studied most widely.

Table 1-3 Phenomenological models adopted most widely.

Phenomenological models	Equations
Modified Johnson-Cook (MJ-C)	$\sigma = (A_1 + B_1\varepsilon + B_2\varepsilon^2)(1 + C_1 \ln \dot{\varepsilon}^*) \exp[(\lambda_1 + \lambda_2 \ln \dot{\varepsilon}^*)T^*]$
Arrhenius-type (A-T)	$\left\{ \begin{array}{l} Z = \dot{\varepsilon} \exp\left(\frac{Q}{RT}\right) \\ \dot{\varepsilon} = AF(\sigma) \exp\left(-\frac{Q}{RT}\right) \\ f(x) = \begin{cases} \sigma^{n_1}, & \alpha\sigma < 0.8 \\ x, & \alpha\sigma > 1.2 \\ [\sinh(\alpha\sigma)]^n, & \text{for all } \sigma \end{cases} \end{array} \right.$
Modified Zerilli-Armstrong (MZ-A)	$\sigma = (C_1 + C_2\varepsilon^n) \exp[-(C_3 + C_4\varepsilon)T^* + (C_5 + C_6T^*) \ln \dot{\varepsilon}^*]$
Modified Fields-Backofen (MF-B)	$\sigma = C\varepsilon^n \dot{\varepsilon}^m \exp(bT + s\varepsilon)$
Kobayashi-Dodd	$\sigma = \sigma_0 \varepsilon^n \dot{\varepsilon}^m (1 - \beta\Delta T)$
Wang-Jiang	$\sigma = (A + B\varepsilon^n) \left[1 - CT \ln\left(\frac{\dot{\varepsilon}}{\dot{\varepsilon}_0}\right) \right]$
SK-Paul	$\sigma = \sigma_0 e^{A \ln\left(\frac{\dot{\varepsilon}}{\dot{\varepsilon}_0}\right) - k(T - T_a)}$ $+ [B\varepsilon + C(1 - e^{-\beta\varepsilon})] \left(1 - H \ln\left(\frac{\dot{\varepsilon}}{\dot{\varepsilon}_0}\right) \right) (1 - G(T - T_a))$
Modified Hensel-Spittel	$\sinh(\alpha\sigma) = A'_{HS} e^{m_1 T} \varepsilon^{m_2} e^{\frac{m_4}{\varepsilon}} (1 + \varepsilon)^{m_5 T} e^{m_7 \varepsilon} \left(\frac{\dot{\varepsilon}}{s-1}\right)^{m_3}$

Taking AA6061 as an example, **Table 1-6** shows the models and their accuracy performances in the previous researches. *R* and *AARE* were adopted as evaluation indicators.

Table 1-4 Models and their performances in previous works.

Models	Test types	Temperatures (°C)	Strain rates (s ⁻¹)	<i>R</i>	<i>AARE</i> (%)
MJ-C	Compression	400, 450, 500, 550	0.1, 1, 10	0.9916	5.71 [13]
	Tensile	Room temperature	0.0004, 0.004, 0.04, 0.2, 0.4	-	2.34 [14]
A-T	Compression	400, 450, 500, 550	0.1, 1, 10	0.9982	2.85 [13]
	Compression	400, 450, 500	0.01, 0.1, 1	-	7.33 [15]

	Compression	350, 390, 430, 470, 510	0.01, 1	0.991	2.33 [16]
Modified Hensel-Spittel	Compression	400, 450, 500, 550	0.1, 1, 10	0.9862	6.68 [13]
MF-B	Compression	350, 390, 430, 470, 510	0.01, 1	0.992	2.15 [16]

1.3.2.2 Physical model

Unlike phenomenological models, physical models consider microstructures such as dislocation density, grain size, grain density, etc. Consequently, physical models have the merit of describing internal structural changes. However, to start physical modeling work, high resolution microscopes like optical microscope (OM), scanning electron microscope (SEM), and transmission electron microscope (TEM) are necessary to obtain micro information before analysis and modeling. Wang et al. [17] combined the phenomenological Arrhenius-type and physical models to study the microstructure evolution of 30CrMnSiNi2A alloy. With R 0.972 and $AARE$ 5.753%, the physical model owns a slightly lower accuracy than the Arrhenius-type model (R 0.979 and $AARE$ 5.283%). Integrating dynamic atom interactions of dislocation and solute, Lin et al. [18] proposed a physical constitutive model that fits the flow stress of nickel-based alloy. Wen et al. [19] put forward dislocation-based and improved dislocation-based physical models and compared with them. The improved model achieves higher $AARE$ and lower R . Considering dislocation density, Zhang et al. [20] built a two-stage physical model to study the CrMnSiNi high-strength alloy hardening mechanism and tensile properties, and good results were received. Lin and Yang [21] conducted early work on intelligence optimization algorithms that optimized physical constitutive modeling. By implying the multi-objective optimization for parameter

determination, higher precision was received. Furthermore, some physical models were created based on phenomenological models by considering microstructures, e.g., Ashtiani and Shayanpoor [22] modified the original J-C model by adding grain size into the equation. The following table shows the physical model performances from the above kinds of literature.

Table 1-5 Model performances from previous literatures.

	[17]	[18]	[19]	[20]	[21]	[22]
Alloys	30CrMnSiNi2A	Nickel-based	Ultrahigh strength	CrMnSiNi	Superplastic	AA1070
<i>R</i>	0.972	0.995	0.991	0.96774	-	0.9626
<i>AARE</i> (%)	5.753	4.51	5.814	-	-	10.797

1.3.2.3 Research trend of flow behavior constitutive model

Generally speaking, phenomenological models have higher accuracy than physical models, as can be seen in section 1.3.2.1 and section 1.3.2.2. However, errors exist in both models. For phenomenological models, *AAREs* are around 7% [8, 10]. Even worse, the *AARE* of one physical model is more than 10% [22]. Two strategies were proposed to address the accuracy-lose problem: 1. Adopting optimization algorithms (e.g., the gradient descent algorithm, genetic algorithm [23], particle swarm algorithm [24], grey wolf algorithm [25] and so on) to optimize constitutive equations, namely, optimizing interior material parameters of the models [26]; 2. Choosing ML methods to fit and predict the materials' flow behavior [27]–[29].

Due to the limit in equation expression, significant errors still exist in the optimization works (although large errors can be reduced). On the other hand, the new path opened up by ML has profound future significance. The models established through ML with the ability to

achieve minimized error in the flow stress data. However, a comprehensive study about ML on material flow behavior modeling was rarely reported. So, the work on comprehensive ML constitutive modeling is destined to have great significance.

1.4 Machine learning (ML)

ML is a subset of AI, belonging to computational science. ML is specialized in analyzing and explaining the pattern and structure of the data. Without human interaction, ML algorithms are able to finish the tasks of learning, reasoning, and decision-making. Clearly speaking, ML is when users input a large amount of data to computers, and then computers analyze the data and output data-driven recommendations and decisions according to the import data. ML will integrate new information to improve future decisions if there is any modification of algorithm recognition. In short, ML is the core of AI and the fundamental approach to making computers intelligent. ML involves a broad algorithm knowledge like statistics, probability theory, approximation theory, etc.

1.4.1 Development history

The Origin of ML can date to the 17th century. The Markov chain and the derivation of the least squares method consist of the fundamental and tool for widely used ML. From 1950 to 2000, ML can be divided into four stages based on research targets and methods.

1.4.1.1 1955-1965

In this stage, people mainly focused on studying the execution ability of systems. The data fed back by systems can be detected by changing the machine environment and the corresponding performance parameters. For instance, by changing the free space effects of the program inside the system, the system will change the self-organization to choose the optimal environment to survive. However, the research results, like the chess program achieved in this stage, are far from human requirements.

1.4.1.2 1965-1975

The main task in this stage was implanting the knowledge of all kinds of fields into systems to achieve the goal of using machines to simulate the human learning process. At the same time, the logical structure and graph structure were adopted for system description. In this period, various symbols are generally utilized as machine languages. However, researchers realized that learning is a long process, and machines cannot learn more profound knowledge. Hence, bits of knowledges from many scholars were put into the systems. Based on this, a certain achievement was made.

1.4.1.3 1975-1985

In this period, ML gained rapid growth. Multi-conception learning expanded from single-conception learning, and different learning methods and strategies were explored. Besides, in this stage, ML has been combined with a variety of applications and received big success. Simultaneously, the development and research were greatly stimulated by the requirements

of expert systems in knowledge acquisition. After the first expert learning system, automatic knowledge acquisition and example, inductive learning systems became the research target of machine learning applications. In the 1980s, ML has begun to rise worldwide.

1.4.1.4 1980s

This is the newest stage of ML research. ML has become a new subject that integrates computational science, automation, mathematics, neurophysiology, biology, psychology, etc. ML absorbs diverse learning methods, and integrated learning system research with various forms is on the rise. Moreover, a unified view of all kinds of fundamental problems about AI and ML is emerging. Machine learning is becoming increasingly widespread, and many have been translated into practical benefits.

1.4.2 Research status of ML

ML refers to the science of studying how to simulate and realize human being learning activity by computers, one of the most cutting-edge and intelligent fields among AI. As the approach to realizing AI, ML has been one of the crucial subjects of AI and has gone through rapid development. Traditional ML builds learning mechanisms by simulating human beings' learning mechanisms, including Bayesian learning (BL), artificial neural network (ANN), DT, RF, SVM, etc.

1.4.2.1 BL

BL stems from British mathematician Bayes, who proved a special case about Bayesian theory [30]. BL is one of the early ML research directions. Bayesian statistics was established around the 1950s, as an important part of statistics [31]. The Bayesian network was proposed by Pearl [32] and has now been one of the most popular research topics. Bayesian network is an uncertainty processing model that simulates the causal relationship in the human reasoning process. Its network topology is a directed acyclic graph (DAG).

1.4.2.2 ANN

As an algorithm with the nonlinear adaptive information process ability, ANN can overcome the traditional AI flaws in unstructured information processing, speech recognition, and pattern. ANN drew attention in the 1940s (perceptron) [33] and received fast development. ANN can be divided into a forward network and a feedback network.

A feedback network refers to feedback between neurons, which a complete undirected graph can represent. The network has no feedback and can be represented by a directed acyclic graph. This network realizes the transformation of signal from input space to output space, and its information processing ability comes from multiple combinations of simple nonlinear functions. In general, forward networks include single-layer perceptron (SLP) [34], MLP [35], adaptive linear (AL), backpropagation (BP) [36] and so on.

A feedback network refers to feedback exists between neurons, which can be represented by an undirected complete graph. The information processing of this neural network is a transformation of states, which can be processed using dynamic system theory. The stability

of the system is closely related to associative memory function. Representative networks include Hamming [37], Hopfield [38], and BAM [39].

Furthermore, with decades of development, deep learning neural networks (DNN) [40], such as convolutional neural networks (CNN) [41] based on image recognition processing and recurrent neural networks (RNN) [42] based on language processing have become research hotspots in ML in recent years and maintain strong development momentum.

1.4.2.3 Decision tree (DT)

DT is a method commonly used in ML. In the late 20th century, J. Ross Quinlan [43] introduced Shannon information theory to DT, then created the ID3 algorithm. Many improved decision tree algorithms were proposed by ML researchers. After that, on the basis of the decision tree pruning method, Quinlan improved the ID3 algorithm to the C4.5 algorithm, which avoids the overfitting phenomenon to some extent [44]. Throughout, DT has been continuously improved, such as the classification and regression tree (CART) algorithm from Leo Breiman [45], the supervised learning in quest (SLIQ) algorithm from Manish Mehta [46], the rough set-based optimization algorithm proposed by Zhang [47], and extreme learning tree-based algorithm proposed by Wang [48]. In recent years, fuzzy decision trees have also flourished [49]. Researchers have proposed hierarchical regression algorithms [50], constrained hierarchical induction algorithms [51], and functional tree algorithms [52], considering the correlation between attributes. These three algorithms are all based on decision tree algorithms combined with multiple classifiers. They have conducted

some experiments and studies on the possible correlations between attributes, but these studies still need to comprehensively explain how the correlation between attributes affects decision tree performance.

1.4.2.4 RF

RF is a supervised machine learning algorithm. Due to its accuracy, simplicity, and flexibility, it has become the most commonly used algorithm. It can be used for classification and regression tasks, coupled with its non-linear characteristics, making it highly adaptable to various data and situations. RF was first proposed by Ho in 1995 [53]. This year, he developed a prediction creation equation created by random data. 2001 Brightman and Cutler expanded the algorithm by adding bootstrap aggregating, and the RF was created [54].

RF is developed on the basis of DT, which is an algorithm using multi-DT for prediction and classification. Although a single DT only has one result and a narrow range of groups, forests can ensure more groups and decisions, thus obtaining more accurate results. In addition to this, RF adds randomness to the model by finding the best feature in the random feature subset. Overall, RF is a model with broad diversity.

1.4.2.5 SVR

Support vector machine (SVM) was proposed by V.N. Vapnik et al. in 1964 [55], which has a wide application on pattern recognition problems like text classification, face recognition, bioinformatics, handwritten character recognition, etc. After years of evolution, a lot of

modification SVM algorithms were developed: improved algorithm for skewed data, Platt's probabilistic outputs [56], multiple class SVM [57], least square SVM (LS-SVM) [58], structured SVM [59], multiple kernel SVM [60], support vector regression (SVR) [61], support vector clustering (SVC) [62], Semi-supervised SVM [63] and so on.

SVR has good sparsity and robustness. Due to the decision boundary being decided by the support vector (other sample points do not participate in empirical risk minimization), SVR owns excellent sparsity. The structure risk minimization and empirical risk are considered when using SVM to optimize problems. Therefore, SVM has good stability.

1.4.3 Classification of Machine Learning

According to different classification standards, ML can be classified into different categories. The learning method is the most commonly used classification standard, on which ML can be divided into supervised, unsupervised, and semi-supervised weak supervised ML strategies. Apart from the learning method, model function, model complexity, and learning objectives should also be division methods. **Fig.1. 5** introduces ML classified by different standards. Flow behavior modeling is a typical supervised ML problem as data with labels like temperature, strain rate, and strain are imported into algorithms for training, testing and validation.

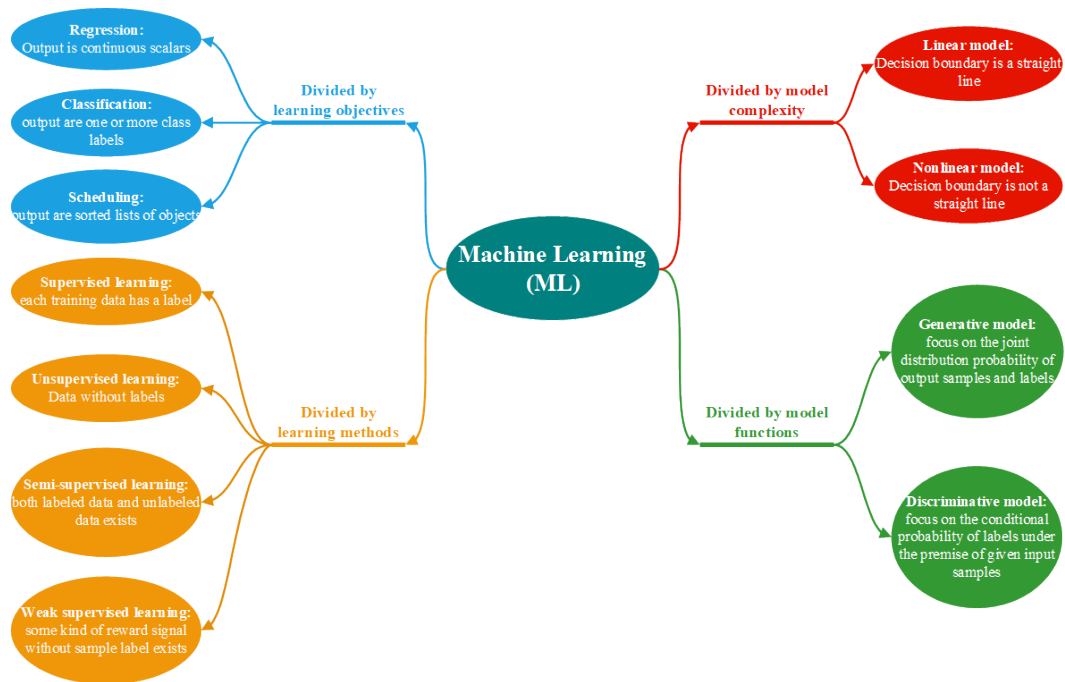


Fig.1. 5. ML classified by different classification standards.

1.4.4 Process of ML

Data collection, process, and validation are essential steps of ML. In most cases, data for training are obtained from experiments or databases. Databases are pretty different for different fields. For material mechanics science, multiple complete large-scale databases have been established [64], such as MakeItForm, MatMatch, MatWeb, MATDAT, Materials Project, AFLOW, etc. Nowadays, due to the advancement of data mining, text mining, intention mining and so on are receiving enthusiastic attention from people. Besides, learning, validation, and test parts are essential for the accuracy and performance of ML algorithms. Cross-validation and many statistical methods are used as evaluation indicators for ML, see Fig.1. 6.

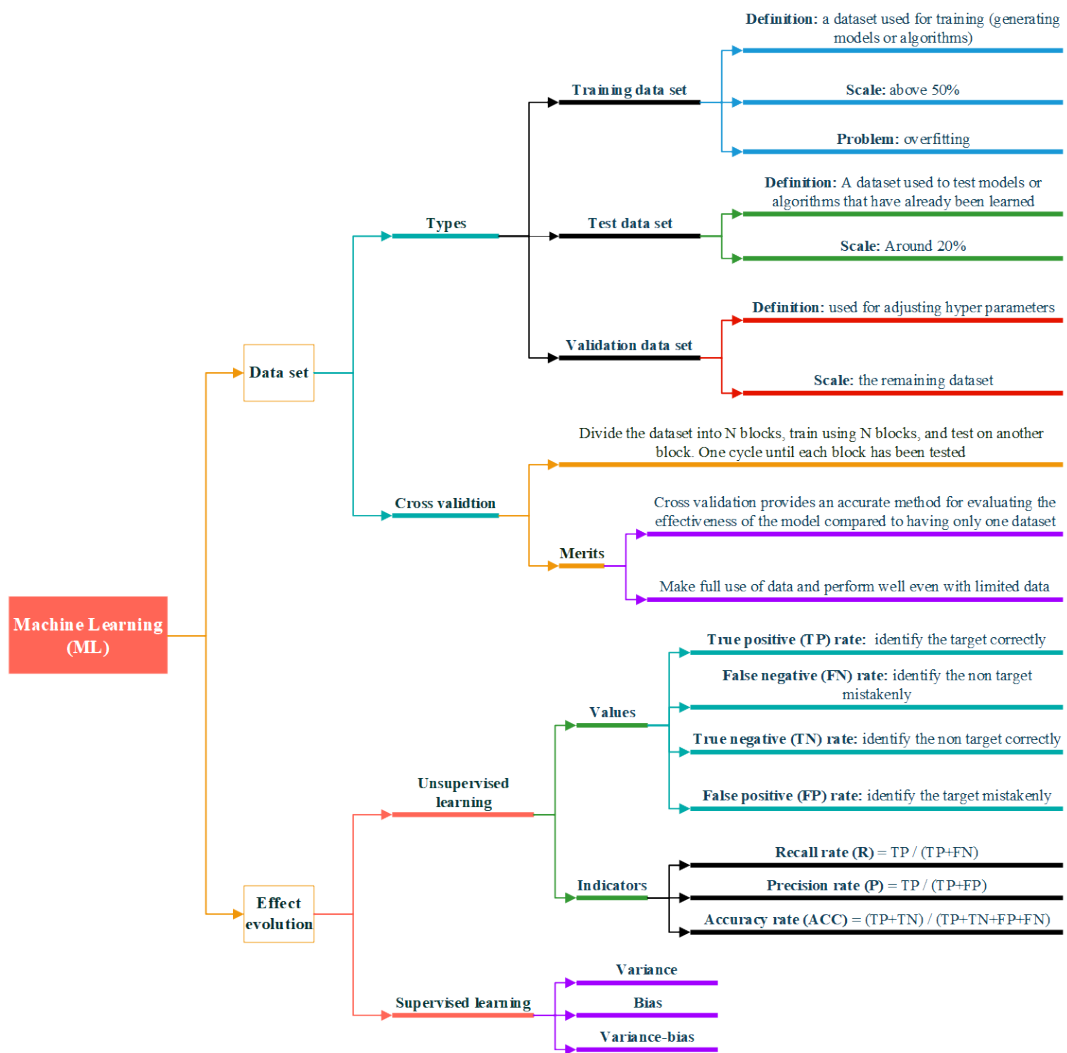


Fig.1. 6. Data processing and evaluation in ML.

1.5 Research purpose and significance

Al-Mg-Si alloy is a commonly used high-strength structural material in aviation, automobile, and other industries. In the design and manufacturing process of materials, the prediction and control of flow stress are critical, as it directly affects the deformation, strength, and durability of the material. Therefore, establishing a flow stress model for Al-Mg-Si alloy is

an important research task.

ML is a method that can automatically learn from data and make predictions, and it has been widely used in the field of materials science. The research objective of establishing a flow stress model for Al-Mg-Si alloy is to use ML algorithms to learn the flow stress law of the alloy from existing experimental data and to establish an accurate prediction model. This model can be used to guide the material design and manufacturing process, as well as to optimize material properties.

Specifically, the significance of this research includes the following:

1. Improving the efficiency and quality of material design and manufacturing: By establishing an accurate flow stress model for Al-Mg-Si alloy, more accurate predictions and controls can be made in the material design and manufacturing process, thereby improving production efficiency and product quality;
2. Exploring the inherent rules of alloy flow stress: By using machine learning algorithms, the inherent rules and characteristics of alloy flow stress can be learned from a large amount of experimental data, revealing the relationship between alloy flow stress and its deformation conditions;
3. Enriching the comprehensive research on the modeling of ML-assisted flow constitutive models: This research provides a reference for selecting ML algorithms for other material modeling and promotes the application of machine learning in computational material mechanics;
4. Promoting the development of materials science: This research uses machine

learning algorithms to explore the relationship between material properties, providing a new way of thinking and method for the development of materials science.

The present thesis mainly conducts the following research work:

1. The thermal cylinder compression tests were finished on Gleeble-3800 thermal simulator, and the flow stress data were collected to study the constitutive relationship of Al-Mg-Si AA6061-T4 alloy;
2. The DMM-based thermal processing maps were drawn to determine the desired deformation conditions under experimental conditions. The power dissipation and instability regions were discussed in detail. The ideal processing conditions were given in this work;
3. ML strategies like linear regression algorithms (MLL) and nonlinear regression algorithms (GAM, SVR, RT, RF, MLP) were adopted to describe the flow behavior;
4. To determine the performances of traditional regression methods and ML methods, the accuracy of received models was analyzed and compared by statistical methods, such as R , δ , $AARE$, and RE .

2 EXPERIMENT PROCEDUR

2.1 Introduction

Flow behavior is a primary parameter characterizing the deformation of metals and alloys. During the plastic deformation, the flow stress determines the amount of applied load and dissipated energy. On the one hand, the magnitude of flow stress is always regarded as the basis for process equipment selection. On the other hand, flow stress provides a reference for die and process design. Furthermore, in some extent, the change of flow behavior reflects the evolution of the material microstructure. Meanwhile, to simulate metal plastic deformation process by numerical analysis methods, the functional relationship between flow behavior and macro thermodynamic parameters (the media connecting dynamic response and thermodynamic parameters) needs to be studied first. All in all, establishing material flow behavior during plastic deformation has excellent engineering and academic significance [65]–[67].

Flow behavior constitutive models are usually built based on test data. Flow stress data are collected from thermal simulators at specified temperature range, strain range, and strain rate range. In this section, the flow stress–true strain data received through cylinder compression tests, and the experiment process are introduced.

2.2 Material preparation

The present work selects the extruded Al-Mg-Si alloy AA6061-T6 produced by Alnan Aluminum Inc. as experimental material. T6 means the alloy condition that the alloy is treated by solid solution heat treatment and artificial aging. The production standard refers to Q/ ALNAN 31-2019. The cylinder sample is Rasategaev-type compression type, with geometrical dimensions of Ø10 mm diameter and 15 mm high. Both chemical compositions and mechanical properties satisfy standard requirement. Check **Table 2-1** and **Table 2-2**.

Table 2-1 Chemical compositions of cylinder samples.

Cu	Mn	Mg	Zn	Cr	Ti	Si	Fe	Al
0.33	0.12	0.9	0.05	0.28	0.02	0.68	0.5	Balance

Table 2-2 Mechanical properties of cylinder samples.

Tensile strength (MPa)	Yield strength (MPa)	Elongation (%)	Hardness
373	350	13	90-106 HB

2.3 Cylinder uniaxial thermal compression test

Compression tests were conducted on the Gleeble-3800 simulator, and the experiment was achieved by putting the samples between two flat anvils for upsetting. Sixteen samples were prepared for 16 different deformation conditions: 0.001 s⁻¹ and 300 °C, 0.001 s⁻¹ and 370 °C, 0.001 s⁻¹ and 440 °C, 0.001 s⁻¹ and 510 °C, 0.01 s⁻¹ and 300 °C, 0.01 s⁻¹ and 370 °C, 0.01 s⁻¹ and 440 °C, 0.01 s⁻¹ and 510 °C, 0.1 s⁻¹ and 300 °C, 0.1 s⁻¹ and 370 °C, 0.1 s⁻¹ and 440 °C, 0.1 s⁻¹ and 510 °C, 1 s⁻¹ and 300 °C, 1 s⁻¹ and 370 °C, 1 s⁻¹ and 440 °C, 1 s⁻¹ and 510 °C. As discussed in section 1.3.1.3, drum shape resulting from friction has a significant negative

effect on flow stress calibration. Hence, graphite was put at both cylinder bottoms as a lubricant.

The samples were heated to the designated temperatures with a heat rate of 10 °C/s inside the simulator. Following by keeping for 4 minutes at the designed temperatures. After that, the samples were compressed until the strain was 0.8. After compression, the samples were quenched with water immediately to keep the microstructure constant. The experiment details can be seen in Fig.2. 1.

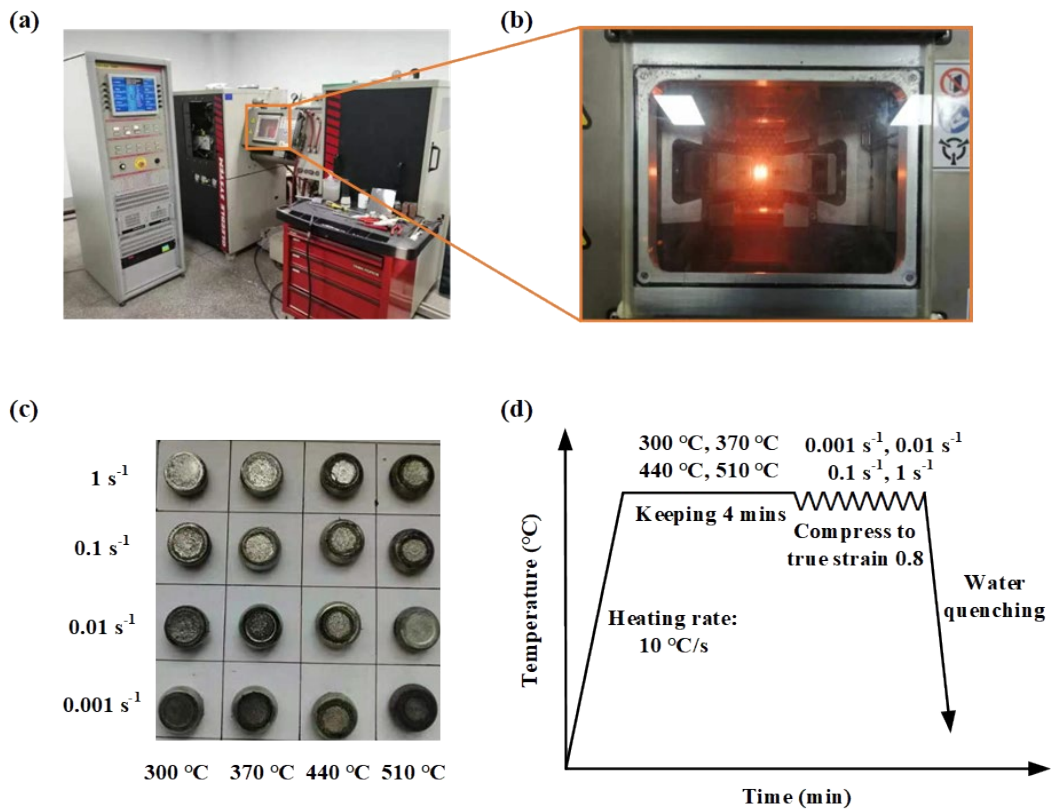


Fig.2. 1. (a) Gleeble-3800 simulator, (b) Visual Work Window, (c) Samples after compression, (d) Processing map.

2.4 OM observation

One sample before deformation and one sample after compression at 0.001 s^{-1} and $510 \text{ }^\circ\text{C}$ were cut by a metallographic cutting machine along the axis to check the microstructure of the natural planes of the samples. Polishing and etching are two processes required to conduct before observation. Three steps from rough to thin were designed in the automatic polishing process, see **Table 2-3**.

Table 2-3 Automatic polishing process.

	Step 1	Step 2	Step 3
Grinding, polishing	Si ₂ C sand paper	SC-JP polishing cloth	ZN-JP polishing cloth
abrasive	P400-P2500	3 μm MD-W Diamond polishing solution	0.05 μm SiO ₂ polishing solution
Lubricant	Water	PL-W polishing lubricant	-
Pressure (N)	15	20	15
Rotation speed	Up 80/down 200	Up 80/down 150	Up 80/down 120
Rotation direction	Single	Single	Single

The initial microstructure was etched by Graff and Sargent's etchant with a composition of 3.0 gm CrO₃, 0.5 ml HF (concentrated), 15.5 ml HNO₃, and 84 ml H₂O [68]. Since the thermal deformation leads to the performances of intra-grain and grain boundaries tending to be consistent, the etching etchant cannot distinguish them. The microstructure, after deformation, was etched by hydrofluoric acid corrosion solvent to check the flow lines. The solution ratio is HF: H₂SO₄: H₂O = 1:2:17. Finally, the microstructures were observed by OLYMPUS PEM3 metallurgical microscope. Meanwhile, the initial microstructure checking was enlarged 50x, and the deformation microstructure was enlarged 25x, see **Fig.2. 2**.

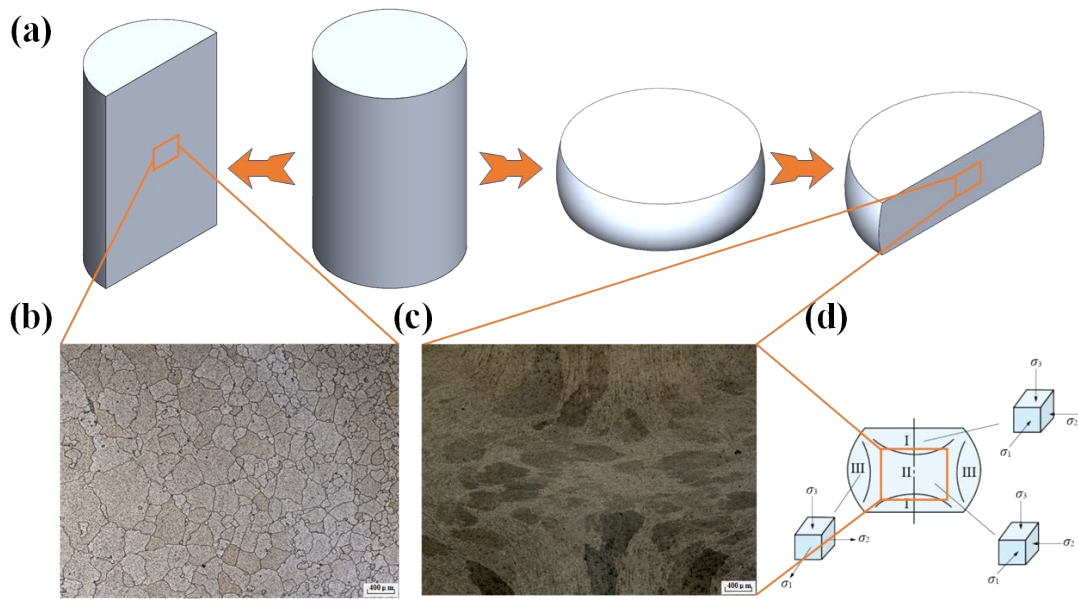


Fig.2. 2. Geometrical shape deformation and corresponding pictures.

As can be checked in **Fig.2 .2 (b)**, equiaxed grain distributed in the internal extruded Al-Mg-Si alloy uniformly, providing goodly deformation conditions. As shown in **Fig.2 .2 (d)**, during upsetting deformation, three areas inside the cylinder can be divided based on the deformation degree: I. Hard deformation area; II. Large deformation area; III. Small deformation area. It is worth noting that areas I and II are under the triaxial compressive stress condition, while area III is under the stress condition of biaxial compression and uniaxial tension. **Fig.2 .2 (c)** displays the area in area II, the direction of flow lines is consistent with the stress condition of area II. The flow lines from up and down spread along the transverse midplane.

2.5 Flow stress of Al-Mg-Si AA6061-T6 alloy

The flow stress–true strain data was exported directly from the Gleeble-3800 simulator. Following equations are embedded in the simulator, with the function of transferring the collected engineering stress (σ_e)-strain (ε_e) to the true stress-strain that is needed in the analysis:

$$\varepsilon = \ln(1 + \varepsilon_e) \quad (2-1)$$

$$\sigma = \sigma_e(1 + \varepsilon_e) \quad (2-2)$$

Eventually, the received data were plotted to flow stress curves, which are integrated in **Fig.2**.

3.

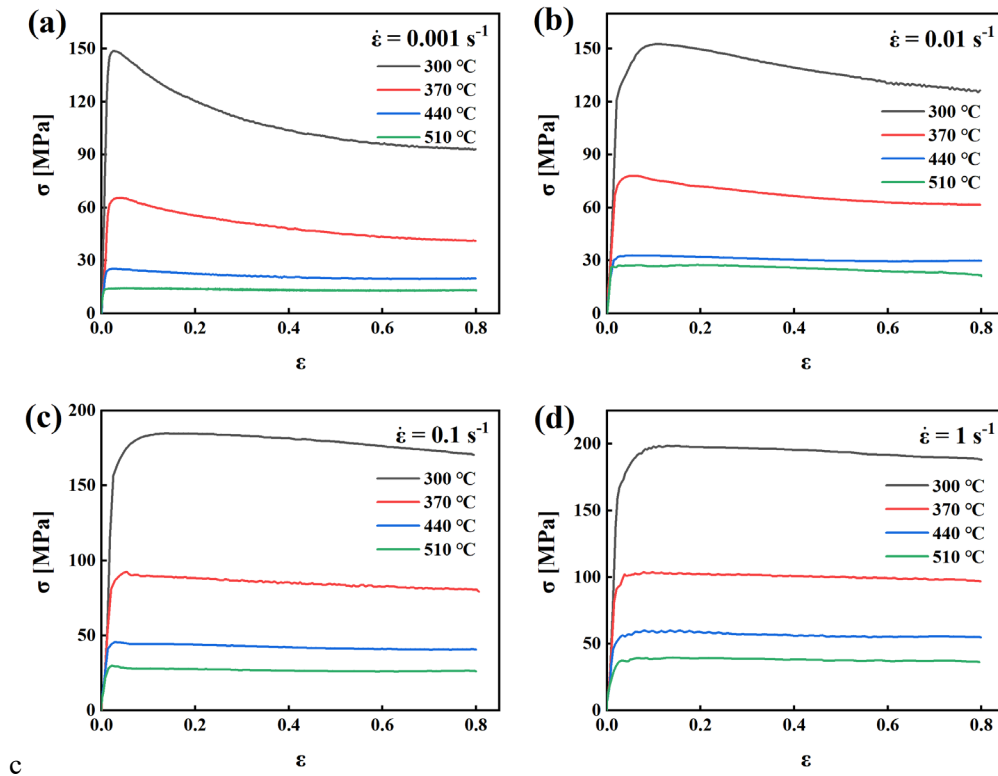


Fig.2. 3. Flow stress curves under experimental deformation conditions: (a) $\dot{\varepsilon} = 0.001 \text{ s}^{-1}$, (b) $\dot{\varepsilon} = 0.01 \text{ s}^{-1}$, (c) $\dot{\varepsilon} = 0.1 \text{ s}^{-1}$, (d) $\dot{\varepsilon} = 1 \text{ s}^{-1}$.

As **Fig.2. 3** shows, when the strain rate is the same, flow stress decreases with the increase in temperature. This is because when the temperature rises, the activity of atoms is enhanced,

and the movement of dislocations becomes easier, reducing the load required for the deformation of AA6061-T6 alloy. Whereas when the temperature keeps the same, flow stress increase with the increase of strain rate, revealing the positive strain rate sensitivity characteristic. The increase in strain rate leads to a sharp increase in the deformation of the alloy at the same time, which requires it to simultaneously increase the number of slip dislocations and accelerate its movement speed during deformation, resulting in a rapid increase in peak stress.

As mentioned in section 1.3.2, the variation of flow stress is influenced by a coupling effect of temperature, strain rate, strain, chemical composition, and microstructure. The Al alloys' deformation mechanism has been illustrated in many literatures [69]–[71]. In general, Al deformation is the comprehensive embodiment of internal dynamic hardening (DH) and dynamic softening (DS). Both DH and DS are essentially movements of dislocations. At the initial deformation stage, the enormous kinetic energy exerted externally drives the generation and movement of a large number of dislocations. The generated dislocations interact with each other or are pinned by impurity atoms, resulting in a lot of entanglements, which is represented externally as a sharp increase in stress. As seen in **Fig.2. 3**, deformation hardening occurred in all deformation conditions.

On the micro-scale, DS is mainly results from dynamic recrystallization (DRX) and continuous dynamic recovery (DRV). Due to Al alloy's the high stacking fault energy for Al alloy, continuous dynamic recrystallization (CDRX) contributes to the primary recrystallization mechanism [23]. Sub-grain boundary continuously absorbs dislocations

during the process of thermal deformation, increasing the angle. Finally, the low-angle grain boundary becomes the large-angle grain boundary. Namely, sub-grain becomes true grain. Unlike the DRX that eliminates sub-grains and dislocations through the movement of large-angle grain boundaries, the transition from low-angle grain boundaries to large-angle grain boundaries also consumes a large amount of dislocation density, leading to the apparent softening effect. Only at the critical deformation and a higher deformation temperature will DRX occur. Moreover, the time required for DRX decreases with increasing temperature. DRV is that dislocations undergo cross slip and climb under the action of thermodynamic activation energy, thereby annihilating each other and rearranging. During this process, DRV consumes dislocations and causes alloy softening. Generally, CDRX consumes much more dislocations than DRV, bringing more obvious softening effects.

Consequently, DRV needs less energy to carry out, which means DRV happens easier than CDRX. On the flow stress curves, the mechanism dominated by dynamic recrystallization often exhibits noticeable stress reduction and softening characteristics. Besides, at low strain rates, after a sharp increase, the flow stress curve remains in the form of a wavy line at a certain stress value, which is also a manifestation of DRX. This is because dislocation density increases slowly, dynamic recrystallization cannot balance with work hardening and metal work hardening makes the curve rise. When the dislocation density increases enough to make dynamic recrystallization dominant, the curve decreases again. This process is cyclic, but the amplitude of the fluctuation gradually decreases.

Furthermore, it should be aware that DH, DRX, and DRV are not a single process but always happen simultaneously. At different stages, the dominant mechanism will change. Based on the above statement, **Table 2-4** lists the different dominant deformation mechanisms under experimental conditions.

Table 2-4 Dominate deformation mechanisms under experimental conditions.

$\dot{\epsilon}$ \ T	300 °C	370 °C	440 °C	510 °C
0.001 s ⁻¹	DRX	DRX	DRX	DRV
0.01 s ⁻¹	DRX	DRX	DRV	DRV
0.1 s ⁻¹	DRX	DRX	DRV	DRV
1 s ⁻¹	DRX	DRX	DRX	DRX

Under deformation conditions 300°C and 0.001 s⁻¹, 300°C and 0.01 s⁻¹, 300°C and 0.1 s⁻¹, 370°C and 0.001 s⁻¹, and 370°C and 0.01 s⁻¹, DS beats WH. The obvious peak phenomenon shows the typical DRX characteristic. The flow stress curves under deformation conditions 370°C and 0.1 s⁻¹, 300°C and 1 s⁻¹, 370°C and 1 s⁻¹, 440°C and 1 s⁻¹, 510°C and 1 s⁻¹ show typical DRX characteristic under low strain rates. At low temperatures, the atoms in the material structure are stable, making the alloy hard to deform. As a consequence, more load is required for deformation. A large amount of load offers big deformation energy. When the strain rate is low, DRX has enough time to realize. Hence obvious softening was achieved.

On the contrary, when the strain rate rises, the time is too short to carry out CDRX. Consequently, DRV domains. At conditions 440 °C and 0.01 s⁻¹, 440 °C and 0.1 s⁻¹, 510 °C and 0.001 s⁻¹, 510 °C and 0.01 s⁻¹, 510 °C and 0.1 s⁻¹, the flow stress curves show that flow stress comes to nearly stable after the increase, which means the WH and DS strike to

balance after WH period. This phenomenon is consistent with the plastic deformation characteristics dominated by DRV. As the strain increases, the density of dislocations increases through proliferation, leading to dislocation entanglements and cellular substructures. However, due to the high temperature of thermal deformation, it provides thermal activation conditions for the recovery process. The density of dislocations is continuously reduced through climbing edge-shaped dislocations, the cross slip of screw-shaped dislocations, detachment of dislocation nodes, and subsequent offsetting of dissimilar dislocations on new slip surfaces. When the growth rate and disappearance rate of dislocations reach a balance, hardening does not occur, and the stress-strain curve transitions to a steady-state flow stage when it is horizontal.

Since the kinetic energy of atoms increases at high temperatures, DRV is easier to perform. Before CDRX could start, DRV had already consumed the vast majority of deformation energy. Therefore, When the strain rate is 1 s^{-1} , the simulator will exert enormous kinetic energy on the material in a short period of time, and the alloy has sufficient conditions for CDRX. This is also the reason why the alloy tends to lean towards CDRX at the high strain rate. The level of strain rate here is defined relative to the conditions of experimental design.

2.6 Summary of the chapter

The Gleeble-3800 simulator was utilized to conduct axial thermal compression tests for Al-Mg-Si AA6061-T6 alloy. The experiment sets, and process are introduced in detail. Besides,

the initial and deformation microstructure were observed to guarantee the goodly deformation conditions. The deformation mechanism of the alloy was discussed comprehensively: As the deformation temperature decreases, and strain rate increases, AA6061-T6 alloy will undergo a softening process dominated by CDRX; As the deformation temperature increases and strain rate decreases, AA6061-T6 alloy will undergo a softening process dominated by DRV.

3 THERMAL PROCESSING MAP

3.1 Introduction of thermal processing map

Thermal processing map refers to the map characterizing the intrinsic workability of the material, which has many applications: 1. Selecting deformation processing parameters and improving the material's workability; 2. Controlling the structure formed in the deformation process and analyzing the evolution law and deformation mechanism; 3. Avoiding infections by analyzing the plastic instability.

Based on irreversible thermodynamics theory, physical system simulation, and large deformation continuous media mechanics theory, Gegel and Prasad [72] established dynamic material modeling (DMM), which is the basis of the DMM-based thermal processing map. The salient feature of DMM is that the model combines the energy consumed by material plastic deformation and the energy received from the external system (energy from external load). As a consequence, the DMM-based thermal processing map is able to explain how the plastic deformation of the workpiece dissipates energy from the external environment.

3.2 Establishment of thermal processing map

The DMM-based thermal processing map consists of a power dissipation map and an instability map. To achieve a wide angle of view, 3D thermal processing maps gain more

favors these days [73]–[76]. This section shows the whole process of building the thermal processing map for AA6061-T6 alloy.

3.2.1 Establishment of power dissipation map

Gegel and Prasad regard the workpiece, die, and equipment as a thermodynamically closed system.

The input energy can be expressed as:

$$P = \sigma \dot{\epsilon} \quad (3-1)$$

Where P is the input energy, σ is the flow stress (MPa), $\dot{\epsilon}$ is the strain rate (s^{-1}).

P can be dissipated in two terms: 1. Dissipation (G): the energy consumed by a material during plastic deformation, most of which is converted into thermal energy, and a small portion is stored in the form of crystal defect energy; 2. Dissipation covariance (J): energy consumed by structure evolution during material deformation. From the perspective of atomic motion, the P can be divided into kinetic energy and potential energy: 1. The kinetic energy is related to the motion of atoms, that is, to the motion of dislocations. The conversion of kinetic energy is dissipated in the form of thermal energy, thus corresponding to dissipation; 2. The potential energy is related to the relative position between atoms, and changes in microstructure will inevitably cause changes in atomic potential energy, thus corresponding to the dissipation covariance. See **Eq. (3-2)**:

$$P = \sigma \dot{\epsilon} = G + J = \int_0^{\dot{\epsilon}} \sigma d\dot{\epsilon} + \int_0^{\sigma} \dot{\epsilon} d\sigma \quad (3-2)$$

According to the DMM theory, the relationship between flow stress and strain rate can be represented by an exponential function form:

$$\sigma = A\dot{\epsilon}_{\epsilon,T}^m \quad (3-3)$$

where A is a material constant, and m refers to the strain rate sensitivity exponent.

At specified strain and temperature, m controls the power partition between G and J , which can be converted from the change rate of G and J :

$$\frac{\partial J}{\partial G} = \frac{\dot{\epsilon} \partial \sigma}{\sigma \partial \dot{\epsilon}} = \left| \frac{\partial(\ln \sigma)}{\partial(\ln \dot{\epsilon})} \right|_{\epsilon,T} = m \quad (3-4)$$

J reveals the microstructure evolution like DRV, DRX, wedge creaking, etc. Substituting **Eq. (3-3)** into the equation express of J , the specific equation for J can be obtained:

$$J = \int_0^{\sigma} \dot{\epsilon} d\sigma = \left(\frac{m}{m+1} \right) \sigma \dot{\epsilon} = \left(\frac{m}{m+1} \right) P \quad (3-5)$$

To maximize work efficiency and performance, consuming maximum input energy to activate microevolution is required. Furthermore, the J is hard to measure. With the severe nonlinear characteristic, the dissipation covariance with the maximum value (J_{max}) was adopted to introduce the power dissipation factor (η), which indicates the consumed energy proportion by microstructure evolution. When the material is under ideal linear dissipation conditions, the m will be 1, and the J reaches J_{max} . Now,

$$\eta = \frac{J}{J_{max}} = \frac{\left(\frac{m}{m+1} \right) P}{\left(\frac{1}{1+1} \right) P} = \frac{2m}{m+1} \quad (3-6)$$

Eight strain points were selected from 0.1 to 0.8 at an interval of 0.1. At each strain, the logarithms of stress and strain rate were adopted to plot curves. To describe the nonlinear relationship between flow stress and strain rate, a cubic polynomial function was used:

$$\ln \sigma = a + b \ln \dot{\epsilon} + c(\ln \dot{\epsilon})^2 + d(\ln \dot{\epsilon})^3 \quad (3-7)$$

The relationship plots and fitting results are displayed in **Fig. 3. 1**. It is worth noting here that the R values of all fitting results are 1, which means the fitting accuracy is perfect, and the third polynomial function can be fully trusted.

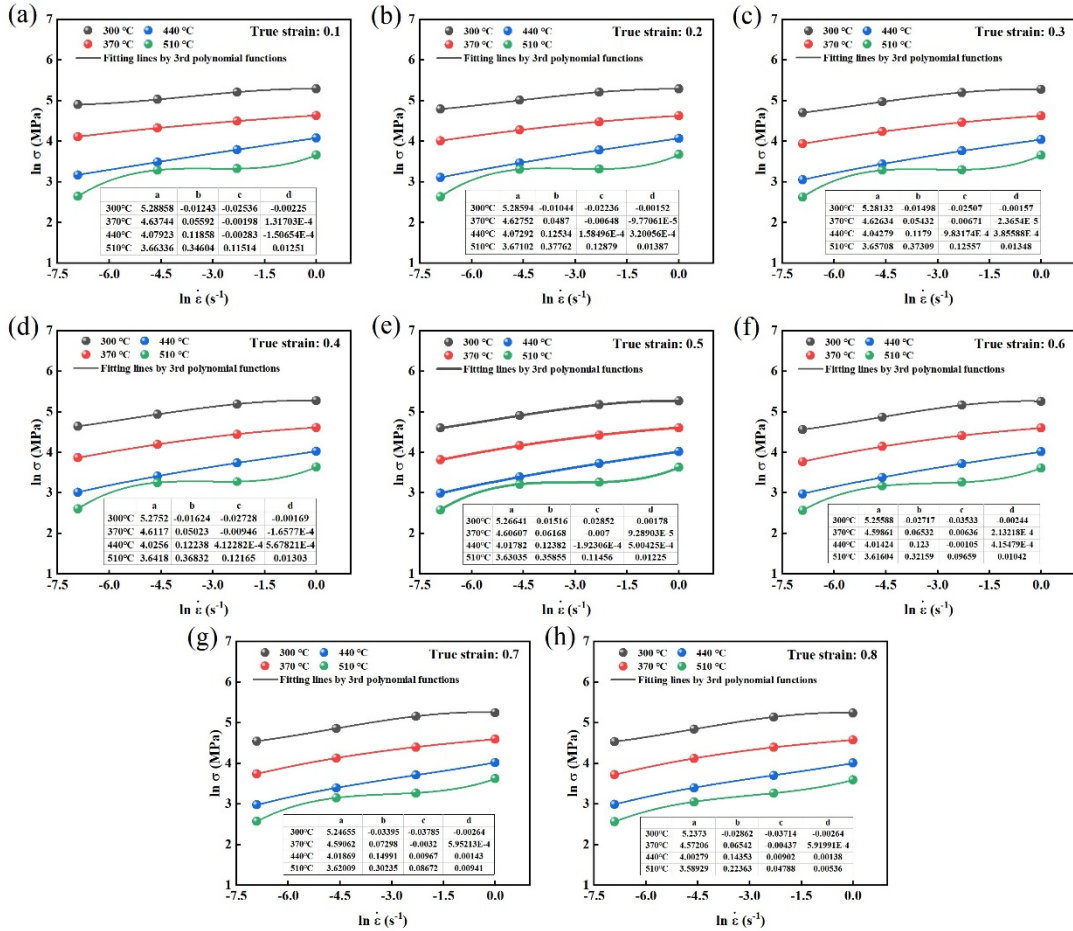


Fig.3. 1. $\ln \sigma - \ln \dot{\epsilon}$ plots and cubic polynomial function fitting results under each strain: (a) 0.1; (b) 0.2; (c) 0.3; (d) 0.4; (e) 0.5; (f) 0.6; (g) 0.7; (h) 0.8.

According to **Eq. (3-3)**, m can be obtained by taking the derivative of $\ln \dot{\epsilon}$ on both sides of the **Eq. (3-7)**:

$$m = b + 2c \ln \dot{\epsilon} + 3d(\ln \dot{\epsilon})^2 \quad (3-8)$$

Different η values under different deformation conditions can be calculated by Eq. (3-6),

then the received η values were scattered on the $\dot{\epsilon} - T$ two-dimensional planar. The specified

m and η refer to Table 3-1.

Table 3-1 m values and η values under different conditions.

ϵ	T (°C)	$\dot{\epsilon}$ (s ⁻¹)							
		m				η			
		0.001	0.01	0.1	1	0.001	0.01	0.1	1
0.1	300	0.015841	0.077993	0.068569	0.01243	0.031188	0.1447	0.128339	-0.02517
	370	0.102128	0.082536	0.067133	0.05592	0.185329	0.152486	0.125819	0.105917
	440	0.136112	0.13506	0.129216	0.11858	0.23961	0.237979	0.22886	0.212019
	510	0.546144	0.081482	0.014781	0.34604	0.70646	0.150686	0.029131	0.51416
0.2	300	0.080885	0.098797	0.068355	0.01044	0.149664	0.179827	0.127963	-0.0211
	370	0.124238	0.102167	0.076987	0.0487	0.221017	0.185392	0.142968	0.092877
	440	0.168967	0.144243	0.129701	0.12534	0.289087	0.25212	0.22962	0.222759
	510	0.583828	0.073868	0.005132	0.37762	0.737237	0.137574	0.010212	0.548221
0.3	300	0.106627	0.116035	0.0755	0.01498	0.192707	0.207942	0.140399	-0.03042
	370	0.143636	0.114616	0.084844	0.05432	0.251192	0.205661	0.156418	0.103043
	440	0.18668	0.151488	0.128561	0.1179	0.314626	0.263116	0.227831	0.210931
	510	0.567955	0.074183	0.009228	0.37309	0.724453	0.138119	0.018286	0.543431
0.4	300	0.118722	0.127496	0.082508	0.01624	0.212245	0.226157	0.152439	-0.03302
	370	0.157195	0.126813	0.091158	0.05023	0.271682	0.225083	0.167085	0.095655
	440	0.197968	0.154709	0.129513	0.12238	0.330507	0.267962	0.229325	0.218072
	510	0.552924	0.076887	0.015352	0.36832	0.712107	0.142795	0.03024	0.538354
0.5	300	0.124049	0.13427	0.087867	0.01516	0.220718	0.236752	0.161541	-0.03079
	370	0.171686	0.132062	0.095394	0.06168	0.293058	0.233313	0.174172	0.116193
	440	0.198113	0.15743	0.132665	0.12382	0.330709	0.272033	0.234253	0.220356
	510	0.529448	0.082792	0.025826	0.35855	0.692339	0.152924	0.050352	0.527842
0.6	300	0.111643	0.142992	0.096721	0.02717	0.200861	0.250206	0.176382	-0.05586
	370	0.183709	0.137463	0.098	0.06532	0.310396	0.241701	0.178507	0.12263
	440	0.196983	0.159105	0.134444	0.123	0.329132	0.27453	0.237022	0.219056
	510	0.478786	0.094913	0.042514	0.32159	0.647539	0.17337	0.08156	0.486671
0.7	300	0.111048	0.146697	0.098365	0.03395	0.199897	0.25586	0.179111	-0.07029
	370	0.202395	0.140322	0.097184	0.07298	0.336653	0.24611	0.177151	0.136032
	440	0.22102	0.151827	0.128123	0.14991	0.362026	0.263627	0.227144	0.260733
	510	0.451322	0.10232	0.052662	0.30235	0.621946	0.185644	0.100055	0.464315
0.8	300	0.106569	0.145488	0.100425	0.02862	0.192611	0.254019	0.18252	-0.05893
	370	0.210538	0.143333	0.094961	0.06542	0.347842	0.250729	0.17345	0.122806
	440	0.216463	0.148252	0.123941	0.14353	0.355889	0.258222	0.220548	0.25103
	510	0.329434	0.123657	0.088389	0.22363	0.4956	0.220097	0.162422	0.365519

η values were connected with each other after dense interpolation, and the contour maps

were plotted. This contour map is the power dissipation map, as follows:

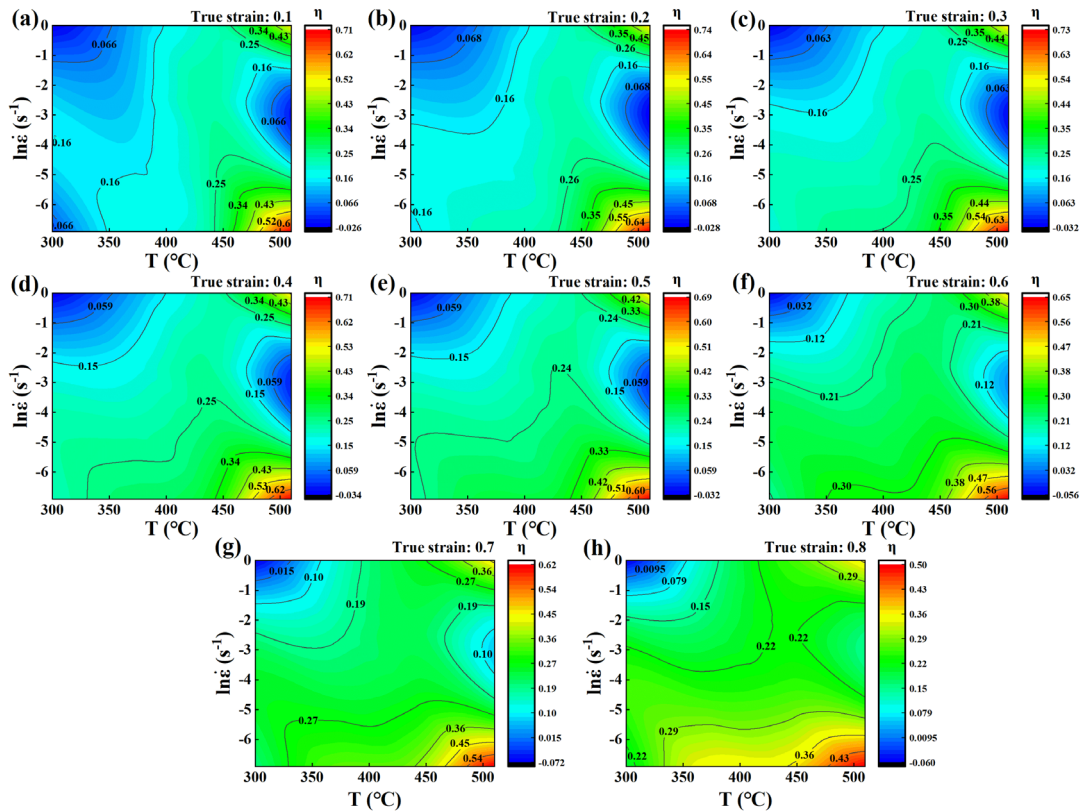


Fig.3. 2. Dissipation maps under experimental conditions divided by true strain: (a) 0.1; (b) 0.2; (c) 0.3; (d) 0.4; (e) 0.5; (f) 0.6; (g) 0.7; (h) 0.8.

As can be checked in From **Fig.3. 2**, at strain perspective, as the degree of deformation deepens, η generally shows an increasing trend and is applicable to all temperature and strain rate states.

The power dissipation rate is relatively low at deformation conditions like low temperature and high strain rate, high temperature, and moderate strain rate. In these situations, the majority of deformation system power converts to the deformation kinetic energy and deformation heat. The deformation process always keeps high energy dissipation at high temperatures and low strain rates (more than 0.5 in most deformation cases). This means more than 50% of system energy is consumed by microstructure evolution in these cases. The power dissipation will increase with the deformation process going because the

temperature and strain rate do not change. As plastic deformation of materials progresses, more and more external energy is input into the system as deformation potential energy.

Moreover, the increase in system temperature caused by material deformation increases the driving force of molecular motion, which provides a prerequisite for the triggering and propulsion of DRV and CDRX. Since temperature and strain rate are important factors affecting the formability of materials, and an increase in temperature and a decrease in strain rate can greatly exacerbate material softening, it is crucial to study the effects of strain rate and temperature on processing performance. To realize a comprehensive observation of the η variation along temperature, strain rate, and strain, respectively, 3D power dissipation maps were drawn.

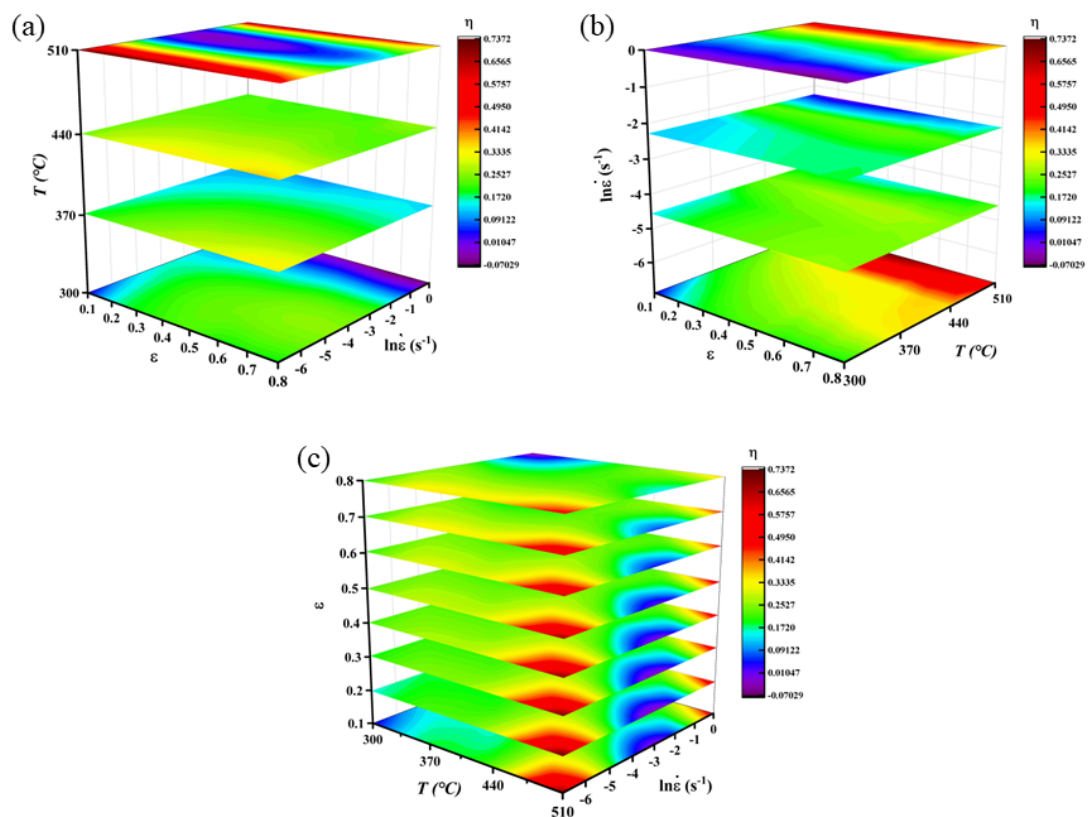


Fig.3. 3. 3D power dissipation maps of AA6061-T6 alloy: (a) along temperature; (b) along

strain rate; (3) along strain.

From the η distribution maps, we can see that η varies significantly with the strain variation, temperature, and strain rate and presents a significant nonlinear characteristic. From the perspective of temperature (**Fig.3. 3(a)**), η exhibits different trends with temperature changes at different strain rates: 1. At low and high strain rates, η increases with increasing deformation temperature, which reveals that at low and high strain rates, the increase in temperature has a positive effect on the evolution of the microstructure of AA6061-T6 alloy; 2. At the moderate strain rate, with the increase in temperature, the η shows a trend of first increasing and then decreasing. In general, the power dissipation rate increases with the temperature rise, especially from 440 °C to 510 °C.

In terms of strain rate (**Fig.3. 3(b)**), the η variation depends on temperature significantly: 1. At low temperature, the η decreases with the increase of strain rate. This can be attributed to the short time of microstructure changes caused by the increase in deformation rate and the insufficient implementation of DRV and CDRX; 2. At moderate temperatures, the η decreases slightly but is not apparent; 3. At high temperatures, the η displays a decrease and increase combined change condition along the strain rate change. At high temperatures and low strain rates, the internal molecular activity of the alloy system is high, and the plastic deformation rate is low. The microstructure can evolve sufficiently in this context, resulting in a high energy dissipation rate. Under the premise of high molecular activity provided by high temperature, in order to maintain a high strain rate external environment, a vast amount of kinetic energy is applied to the alloy in a short time, which greatly promotes the evolution of the internal microstructure of the material, especially CDRX.

3.2.2 Establishment of instability map

If the entropy increase caused by external loading in the deformation system cannot be fully released, it will lead to plastic instability. This phenomenon is mainly manifested as internal cracking, void nucleation, local plastic flow, and adiabatic shear bands in the material [77]–[79]. Ziegler [80] holds the opinion that when the deformation process meets the following equation, the deformation condition will be stable:

$$\frac{dD}{d\dot{\varepsilon}} > \frac{dD}{\dot{\varepsilon}} \quad (3-9)$$

where D is the dissipation function, a material dissipative energy characteristic indicator. Meanwhile, D can be regarded as J when the input energy is divided into G and J . On the above basis, combining **Eq. (3-9)** with **Eq. (3-8)**, a dimensionless parameter equation for determining flow stability can be obtained:

$$\xi(\dot{\varepsilon}) = \frac{\partial \ln \frac{m}{m+1}}{\partial \ln \dot{\varepsilon}} + m = \frac{m'}{m+1} + m = \frac{2c + 6d \ln \dot{\varepsilon}}{m(m+1)} + m < 0 \quad (3-10)$$

When the $\xi(\dot{\varepsilon})$ is lower than 0, the plastic deformation condition is unstable. The received $\xi(\dot{\varepsilon})$ values calculated from Eq. (3-10) are listed in **Table 3-2**.

Table 3-2 $\xi(\dot{\varepsilon})$ values under different conditions.

ε	T (°C)	$\dot{\varepsilon}$ (s ⁻¹)			
		$\xi(\dot{\varepsilon})$			
		0.001	0.01	0.1	1
0.1	300	2.743471	0.236248	-0.23742	4.017301
	370	0.000486	-0.01713	-0.02474	-0.01886
	440	0.140987	0.122477	0.097943	0.065189
	510	-0.2698	-1.44996	3.958853	1.241792
0.2	300	0.325148	0.068536	-0.30238	4.228365
	370	0.043607	-0.00852	-0.08543	-0.23038
	440	0.079387	0.076605	0.093948	0.128186
	510	-0.27691	-1.75296	12.92328	1.317314

0.3	300	0.261589	0.051024	-0.32977	3.28201
	370	0.044591	-0.00953	-0.08257	-0.20615
	440	0.072592	0.055556	0.064536	0.099256
	510	-0.28112	-1.68265	7.108178	1.297365
0.4	300	0.264634	0.057955	-0.32699	3.288806
	370	0.068492	-0.0006	-0.1079	-0.34536
	440	0.060545	0.043761	0.068288	0.129942
	510	-0.28051	-1.55808	4.200763	1.272187
0.5	300	0.275689	0.067899	-0.31387	3.690333
	370	0.049867	-0.00995	-0.0801	-0.1793
	440	0.070354	0.052943	0.070355	0.120329
	510	-0.27536	-1.34746	2.404259	1.226688
0.6	300	0.415032	0.117091	-0.32226	2.502832
	370	0.044808	-0.01654	-0.07752	-0.14213
	440	0.079582	0.060171	0.068289	0.103827
	510	-0.25844	-0.99795	1.24953	1.115473
0.7	300	0.448409	0.125169	-0.33965	2.1201
	370	0.017816	-0.04534	-0.06791	-0.02112
	440	0.000436	-0.00121	0.124459	0.298261
	510	-0.24511	-0.8303	0.920903	1.04943
0.8	300	0.471437	0.134984	-0.31385	2.492488
	370	0.01921	-0.05686	-0.10012	-0.07692
	440	-0.00358	-0.00736	0.114643	0.287258
	510	-0.18063	-0.35197	0.355704	0.747597

Through Spline dense interpolation method, a 100×100 matrix was built. Adopting T as the abscissa and $\ln \dot{\epsilon}$ as the ordinate, the instability maps of AA6061-T6 alloy under experimental conditions were plotted, see **Fig.3. 4**. The regions with $\xi(\dot{\epsilon})$ values lower than 0 are regarded as the instability regions. Based on these figures, the hot working maps will be established in the next section.

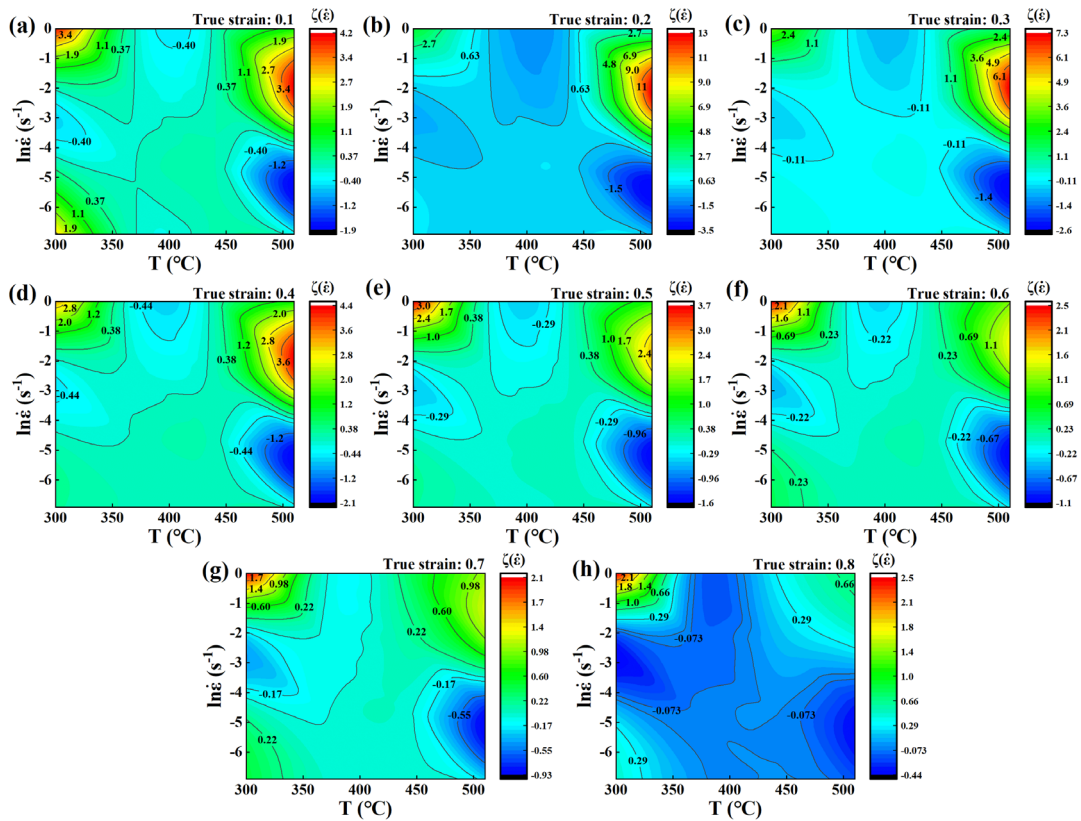


Fig.3. 4. Instability maps under experimental conditions divided by true strain: (a) 0.1; (b) 0.2; (c) 0.3; (d) 0.4; (e) 0.5; (f) 0.6; (g) 0.7; (h) 0.8.

3.2.3 Establishment of thermal processing map

The thermal processing map is obtained by superposing the instability map onto the power dissipation map and focusing on the instability area, namely, the area $\xi(\dot{\epsilon}) < 0$. The processing maps of AA6061-T6 alloy under different strains are shown in **Fig.3. 5**.

The instability areas and the deformation of instability areas are highly nonlinear. As the processing of the deformation, the instability area becomes larger and larger. Furthermore, as can be seen in the maps, the deformation conditions (low temperature and high strain rate, low temperature and low strain rate, and high temperature and high strain rate) keep stable

throughout the whole forming process. The instability region is mainly reflected in the local flow, shear bands, and even the formation of cracks caused by uneven energy dissipation inside the alloy. The deformation conditions (low temperature and moderate strain rate, moderate temperature and high strain rate, and high temperature and moderate strain rate, high temperature and low strain rate) keep unstable from the start to the end, which should be avoided during thermal processing.

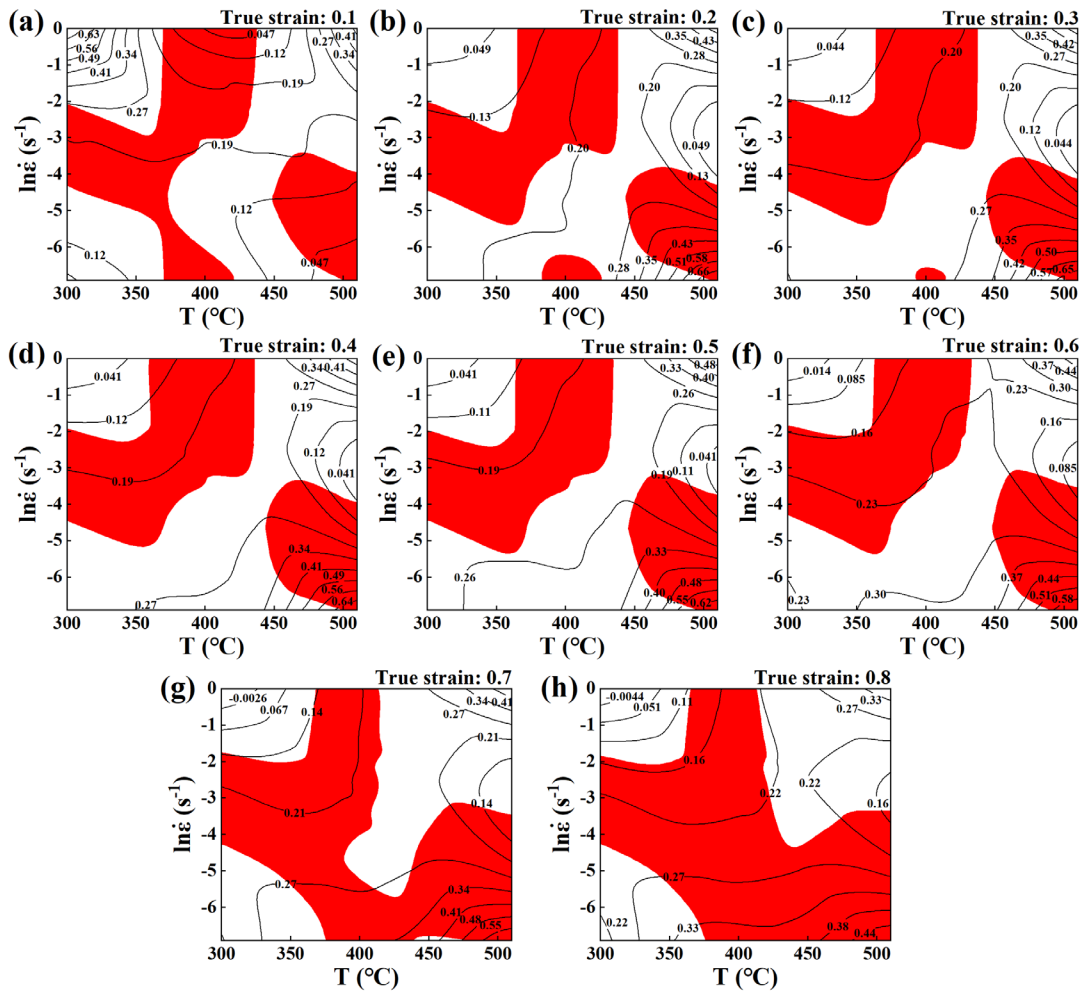


Fig.3. 5. Thermal processing maps under experimental conditions divided by true strain: (a) 0.1; (b) 0.2; (c) 0.3; (d) 0.4; (e) 0.5; (f) 0.6; (g) 0.7; (h) 0.8.

Similarly, the 3D instability maps were established for a deeper study of instability change along temperature and strain rate, see Fig.3. 6.

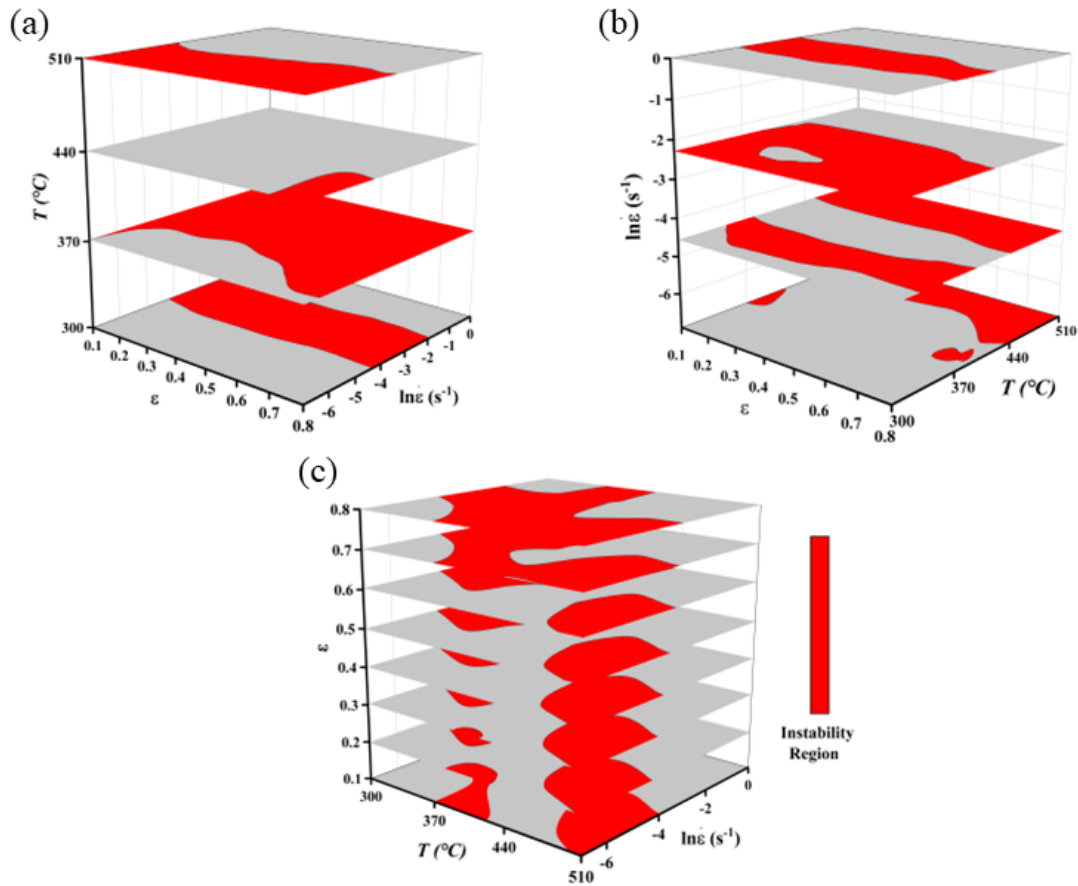


Fig.3. 6. 3D thermal processing maps of AA6061-T6 alloy: (a) along temperature; (b) along strain rate; (3) along strain.

From Fig.3. 6(a), we can see that the instability region varies significantly with the change in processing temperature. When the deformation temperature is 300 °C, the unstable phenomenon happens at moderate strain rates. When the temperature rises to 370 °C, the instability region increases significantly. Only at low strain rates the deformation is able to keep stable. Consequently, a temperature range of 370 °C should be avoided during the processing. At 440 °C, the instability area shrinks to a tiny scale. All deformation conditions are stable except for the low strain rate and high deformation degree. 440 °C seems like the most desirable process temperature. At 510 °C, the deformation of AA6061-T6 alloy cannot maintain stability under medium-low strain rates.

Fig.3. 6(b) displays the instability region distribution at different strain rates. At the strain rate 0.001 s^{-1} , there are discrete instability region distributions, while most instability area concentrates on high temperature and medium-high strain deformation conditions. At strain rate 0.01 s^{-1} , two main unstable regions exist, one at low and the other at high temperatures. At the strain rate of 0.1 s^{-1} , the material undergoes unstable deformation at low and moderate temperatures, while a small region at low temperature and strain 0.3 keeps stable. When the strain rate is at 1 s^{-1} , the instability area occurs at moderate deformation temperature. To sum up, the desirable stable deformation region can be determined as $300 \text{ }^{\circ}\text{C}$ and 0.001 s^{-1} , $300 \text{ }^{\circ}\text{C}$ and 1 s^{-1} , $440\sim 510 \text{ }^{\circ}\text{C}$ and $0.1\sim 1 \text{ s}^{-1}$.

3.3 Summary of the chapter

Based on DMM theory, the power dissipation maps, instability maps, and thermal processing maps were plotted for experimental deformation conditions of AA6061-T6 alloy. The detailed establishment process is introduced in this chapter.

The power dissipation and instability region variation along temperature, strain rate, and temperature of established maps were explained. Finally, the ideal deformation regions under experimental deformation ranges were determined as $300 \text{ }^{\circ}\text{C}$ and 0.001 s^{-1} , $300 \text{ }^{\circ}\text{C}$ and 1 s^{-1} , $440\sim 510 \text{ }^{\circ}\text{C}$ and $0.1\sim 1 \text{ s}^{-1}$. The work conducted in this chapter can provide instruction for AA6061-T6 alloy industrial hot processing.

4 LINEAR REGRESSION MODEL

Linear is a statistical analysis method that uses linear regression analysis in mathematical statistics to determine the quantitative interdependence relationship among two or more variables [81]. This kind of regression equation is a linear combination of one or more regression eco-efficiencies. The regression equation with only one argument is called simple linear regression, while the equation with two or more arguments is called multi-linear regression. Linear regression is the first type of regression analysis rigorously studied and widely used in practical applications. This is because models that rely linearly on their unknown parameters are more accessible to fit than models that rely nonlinearly on their unknown parameters. The statistical characteristics of the resulting estimates are also easier to determine. The least squares approximation is always used to fit linear regression models. Besides, other methods like the gradient descent method are adopted widely as well.

4.1 MLL regression

As the most straightforward regression algorithm in ML, MLL regression refers to a regression question with multiple features. In terms of a sample with n features, the regression function can be written as:

$$\hat{y} = w_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n \quad (4-1)$$

where \hat{y} is the target variable (label), w_0 is the intercept, w_i is the regression coefficient, x_i is the feature in the sample. The equation can be expressed in the form of matrixes:

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_m \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1n} \\ 1 & x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \times \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix} \quad (4-2)$$

the above equation can be simplified as:

$$\hat{y} = \mathbf{X}w \quad (4-3)$$

where \mathbf{X} is a matrix with the structure of $(m, n+1)$, w is a column matrix with the structure $(n+1,1)$.

To explain the operation mechanism of the MLL, the loss function should be introduced.

Here, the maximum likelihood function is adopted to explain the loss function. The error between the real label and the predicted label is:

$$\epsilon_i = y_i - x_i w \quad (4-4)$$

where ϵ_i is the error between y_i and $x_i w$. y_i is the real label. ϵ_i distribute independently and obey $(0, \sigma^2)$ Gaussian distribution:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\epsilon_i - \mu)^2}{2\sigma^2}\right] \quad (4-5)$$

As the errors obey the average value $\mu=0$, the variance as σ^2 Gaussian distribution. Hence, the above error probability Gaussian distribution is:

$$p(\epsilon_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\epsilon_i)^2}{2\sigma^2}\right] \quad (4-6)$$

Substituting **Eq. (4-4)** to **Eq. (4-6)**, the equation yields to:

$$p(y_i | \mathbf{X}_i; w) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_i - \mathbf{X}_i w)^2}{2\sigma^2}\right] \quad (4-7)$$

The smaller the ϵ_i , the greater the $p(y_i|\mathbf{X}_i; w)$, indicating that the predicted value is closer to the real value.

Since the MLL refers to using a hyperplane to fit multiple points, all ϵ_i need to be the minimum values. Namely, the product of all probabilities should be maximum, and this corresponds to the likelihood function:

$$L(w) = \prod_{i=1}^m p(y_i|x_i; w) = \prod_{i=1}^m \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_i - \mathbf{X}_i w)^2}{2\sigma^2}\right] \quad (4-8)$$

Taking the logarithm of both sides, the equation yields to:

$$\begin{aligned} \log L(w) &= \log \prod_{i=1}^m \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_i - \mathbf{X}_i w)^2}{2\sigma^2}\right] \\ &= m \log \frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{\sigma^2} \cdot \frac{1}{2} \sum_{i=1}^m (y_i - \mathbf{X}_i w)^2 \end{aligned} \quad (4-9)$$

Because $m \log \frac{1}{\sqrt{2\pi}\sigma}$ can be regarded as a constant, the target transforms to solving the minimum value of $\frac{1}{\sigma^2} \cdot \frac{1}{2} \sum_{i=1}^m (y_i - \mathbf{X}_i w)^2$. Therefore, the loss function was received:

$$J(w) = \frac{1}{2m} \sum_{i=1}^m (y_i - \mathbf{X}_i w)^2 \quad (4-10)$$

The loss function is divided by m to average the loss. The loss function measures the difference between the real label and the predicted result. In order to obtain better regression results, the difference is expected to be as small as possible. As a consequence, solving the target function transforms to solve the L2 normal form of $y - \mathbf{X}w$:

$$\min_w \|y - \mathbf{X}w\|_2^2 \quad (4-11)$$

Take the derivative of w from the above equation:

$$\begin{aligned}
\frac{\partial \|y - Xw\|_2^2}{\partial w} &= \frac{\partial (y - Xw)^T (y - Xw)}{\partial w} \\
&= \frac{\partial (y^T - w^T X^T)^T (y - Xw)}{\partial w} \\
&= \frac{\partial (y^T y - w^T X^T y - y^T X w + w^T X^T X w)}{\partial w} \quad (4-12) \\
&= 0 - X^T y - X^T y + 2X^T X w \\
&= X^T X w - X^T y
\end{aligned}$$

Make the first derivative of the above equation 0 after taking the derivative, values of w will be obtained:

$$\begin{aligned}
X^T X w - X^T y &= 0 \\
X^T X w &= X^T y \\
w &= (X^T X)^{-1} X^T y \quad (4-13)
\end{aligned}$$

The operation mechanism flow chart of the MLL regression is displayed in **Fig.4. 1**.

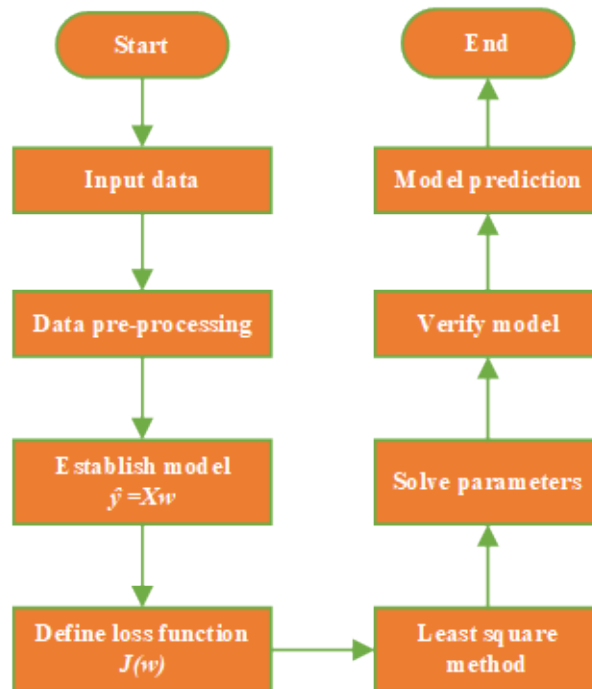


Fig.4. 1. Operation mechanism flow chart of the MLL regression algorithm.

4.2 Regression verification

The flow stress data were imported to a linear regression package with the order of strain rate, temperature, strain, and stress for regression work. 4098 data sets were imported and randomly disrupted. It is worth to noting that the input data is divided into three groups based on the proportion of 8: 1: 1= training: validating: testing. The following chapters still follow this classification standard. Finally, the regression model was built, and the parameters were obtained:

$$w = [0 \quad 0.2487 \quad -0.8474 \quad -0.0512]^T \quad (4-14)$$

Strain points from 0.1 to 0.8 at an interval of 0.1 at all experimental deformation conditions was selected for regression verification. The predicted flow stress values from the MLL regression model were scattered on the flow stress curves plots to make an intuitive comparison. The verification results are plotted in **Fig.4. 2**.

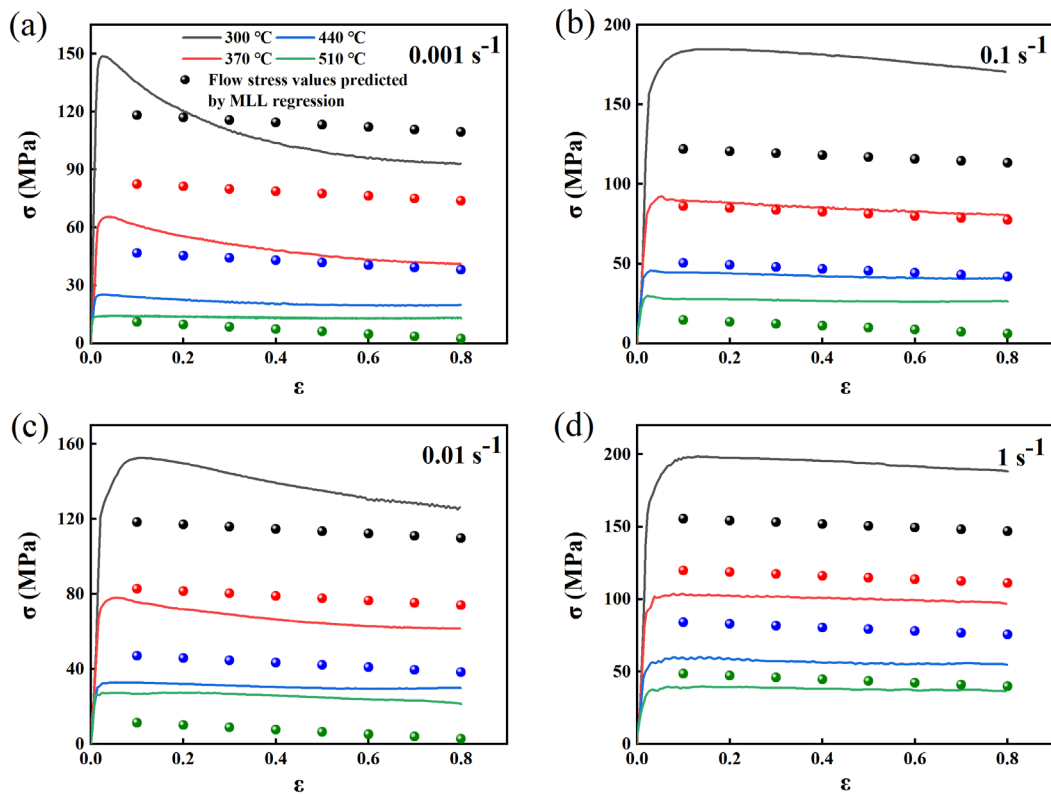


Fig.4. 2. Regression verification of the MLL model at different experimental deformation conditions, divided by strain: (a) 0.001 s^{-1} , (b) 0.01 s^{-1} , (c) 0.1 s^{-1} , and (d) 1 s^{-1} .

This model has the typical linear characteristic. However, it is also evident that the MLL model does not meet the regression requirement of the AA6061-T6 alloy. Nearly all predicted points locate far away from their corresponding flow stress curves, except for a few deformation conditions (strain rate 0.1 s^{-1} and temperature $370\text{ }^{\circ}\text{C}$, strain rate 0.1 s^{-1} and temperature $440\text{ }^{\circ}\text{C}$).

It can be concluded that the flow behavior of the Al-Mg-Si AA6061-T6 alloy is highly nonlinear. The MLL model based on the pure linear regression theory fails to regress the complex flow characteristic of the studied material. Therefore, new regression models need to be exploited.

5 GAM REGRESSION MODEL

Although the linear regression model is simple and easy to understand, the function variation in research works is nonlinear. Linear regression is likely to fail to meet the actual needs, and this problem occurred in the last chapter. The GAM is a free and flexible statistical model that extends from the linear regression model, which can be used to detect the impact of nonlinear regression. The additive model (AM) was first proposed by Stone [82]. He used single smooth functions to estimate every additive term. The dependent variables change with the independent variables can be explained in each additive term. The additive concept presented by Stone solves the problem of model variance increasing with the number of independent variables effectively. To expand the application field of the additive model, Hastie and Tibshirani [83] proposed GAM, the model widely adopted nowadays.

5.1 GAM

5.1.1 Form of the GAM

AM changes the linear regression coefficient w to smooth function, then the MLL is expanded to:

$$E(y|x) = \alpha + f_1(x_1) + f_2(x_2) + \cdots + f_p(x_p) = \alpha + \sum_{j=1}^p f_j(x_j) \quad (5-1)$$

where α is the intercept, $f_j(x_j)$ is the smooth function.

In order to make Eq. (5-1) estimable, the soothing function needs to meet the standardization conditions $E [f_j(x_j)] = 0$. In GAM, the nonparametric smooth functions like local regression smooth function, kernel function, and smooth spline function substitute smooth function give the model significant flexibility and powerful nonlinear fitting ability:

$$\begin{aligned} g(\mu) &= s_0 + s_1(x_1) + s_2(x_2) + \dots + s_p(x_p) \\ n &= s_0 + \sum_{j=1}^p s_j(x_j) \end{aligned} \quad (5-2)$$

where g is the link function, s_j is the nonparametric smooth function, n is the linear prediction value. The distribution of labels y belongs to the natural exponential family like Gamma distribution, Poisson distribution, binomial distribution, etc. In addition, there is no need to make any assumption of y on x . Instead, the link function g takes the work of connecting additive component n and random component y .

5.1.2 Calculation principle of the AM: back-fitting algorithm

The back-fitting algorithm was first adopted to calculate s_i for AM. In the initial back-fitting algorithm, the residual is selected as the indicator for evaluation. The following equation explains the detailed process:

$$\begin{aligned} \text{Initialization} \quad & \hat{f}_j(x) = 0 \quad \forall x_j \text{ and } \forall j, \hat{\alpha} = \bar{y}. \\ \text{Cycle} \quad & j = 1, 2, \dots, p, 1, 2, \dots, p, 1, 2, \dots, \\ & r_{ij} = y_i - \hat{\alpha} - \sum_{\substack{k=1 \\ k \neq i}}^p \hat{f}_k(X_{ki}), i = 1, \dots, n, \\ & \hat{f}_j(x_{ji}) = S(r_j | x_{ji}), i = 1, \dots, n, \\ \text{until} \quad & \text{the functions } \hat{f}_j \text{ converge.} \end{aligned} \quad (5-3)$$

Two modified algorithms were designed based on the back-fitting algorithm: The unweighted back-fitting algorithm and the weighted back-fitting algorithm. Instead of residual, the unweighted back-fitting algorithm chooses root sum square (*RSS*) as the evaluation indicator:

$$\begin{aligned}
 &\text{Initialization } \hat{f}_i(x_i) = 0 \quad \forall x_j \text{ and } \forall j, \hat{\alpha} = \bar{y}. \\
 &\text{Cycle } j = 1, 2, \dots, p, 1, 2, \dots, p, 1, 2, \dots, \\
 &RSS_{ij} = \frac{1}{n} \left\| y - \alpha - \sum_{\substack{k=1 \\ k \neq i}}^p \hat{f}_k(X_{ki}) \right\|^2, \quad i = 1, \dots, n, \\
 &\hat{f}_j(x_{ji}) = E(RSS_j | x_{ji}), \quad i = 1, \dots, n, \\
 &\text{until the functions } \hat{f}_j \text{ converge.}
 \end{aligned} \tag{5-4}$$

RSS never increases in any step for smooth functions such as linear regression, univariate and bivariate splines, or their combinations, which means that the above algorithms will always converge, and the following condition is always regarded as the convergence threshold:

$$\frac{\sum_{i=1}^n \sum_{j=1}^p [f_i^{m-1} X_{ij} - f_j^m X_{ij}]^2}{1 + \sum_{i=1}^n \sum_{j=1}^p [f_i^{m-1} X_{ij}]^2} \leq \varepsilon \tag{5-5}$$

Where *m* means *m*-th iteration. ε is the threshold value. At specified regression problems, thresholds can be customized to be different magnitudes.

Numerical instability with weights may cause convergence problems for distributions other than the normal distribution. Even when the algorithm converges, individual functions do not need to be completely different from each other, and even for the same fitting surface, the dependency between covariates can lead to more than one expression appearing.

The form of the weighted algorithm is the same as that of the unweighted algorithm, except that the smooth function is weighted. The weighted back-fitting algorithm allows for using weights in the local scoring process if the data is not a normal distribution.

5.1.3 Calculation principle of the GAM: local scoring algorithm

The link function is one of the essential parts of the local scoring algorithm. Once the distribution is specified, the corresponding link and adjusted dependent weights are also defined. **Table 5-1** lists the most typical link functions.

Table 5-1 Typical link functions in GAM.

Distribution	Link	Adjusted dependent (z)	Weights (w)
Normal	μ	y	1
Binomial	$\log\left(\frac{\mu}{n-\mu}\right)$	$\eta + \frac{y-\mu}{n\mu(1-\mu)}$	$n\mu(1-\mu)$
Gamma	$-\frac{1}{\mu}$	$\eta + \frac{y-\mu}{\mu^2}$	μ^2
Poisson	$\log \mu$	$\eta + \frac{y-\mu}{\mu}$	μ
Inverse Gaussian	$\frac{1}{\mu^2}$	$\eta + \frac{2(y-\mu)}{\mu^3}$	$\frac{\mu^3}{4}$

Here goes the calculation process:

$$\begin{aligned}
 &\text{Initialization } \hat{s}_j^0(x_j) = 0 \forall x_j \text{ and } \forall j, \hat{\alpha} = g[E(y)]. \\
 &\text{Loop over outer iteration counter } m. \\
 &\hat{\eta}^{m-1}(\mathbf{X}_i) = \hat{\alpha} + \sum_{j=1}^p \hat{s}_j^{m-1}(x_{ji}), \\
 &\hat{\mu}_i = g^{-1}(\hat{\eta}^{m-1}(\mathbf{X}_i)), \\
 &z_i = \hat{\eta}^{m-1}(\mathbf{X}_i) + [y_i - \hat{\mu}_i] \cdot \left(\frac{\partial \mu}{\partial \eta}\right)_i^{m-1}, \\
 &w_i = [V_i^{m-1}]^{-1} \cdot \left[\left(\frac{\partial \mu}{\partial \eta}\right)_i^{m-1}\right]^2, i = 1, \dots, n.
 \end{aligned} \tag{5-6}$$

where V_i^{m-1} is the variance of y at $\hat{\mu}_i$. By using a back-fitting algorithm with weight w , an additive model is fitted to z to obtain the estimated functions $s_j^m, j = 1, \dots, p$.

The iteration will stop until the convergence threshold is reached or the dispersion expanded from RSS no longer decreases. Here, the convergence condition is:

$$\frac{\sum_{i=1}^n w_i \sum_{j=1}^p [s_j^{m-1}(x_{ij}) - s_j^m(x_{ij})]^2}{\sum_{i=1}^n w_i \{1 + \sum_{j=1}^p [s_j^{m-1}(x_{ij})]^2\}} \leq \varepsilon^3 \quad (5-7)$$

In the present work, the ε is defined as 10^{-6} .

Two circles are included in the GAM: the local scoring algorithm as the outer circle and the back-fitting algorithm as the interior circle. After receiving the weighted parameters, the interior circle keeps running till meeting the convergence condition, or RSS no longer decreases. After that, the outer circle calculates a new set of weights based on the estimated values from the interior circle. The system keeps operating until it meets the stop conditions.

The flow chart of GAM is shown in **Fig. 5. 1**.

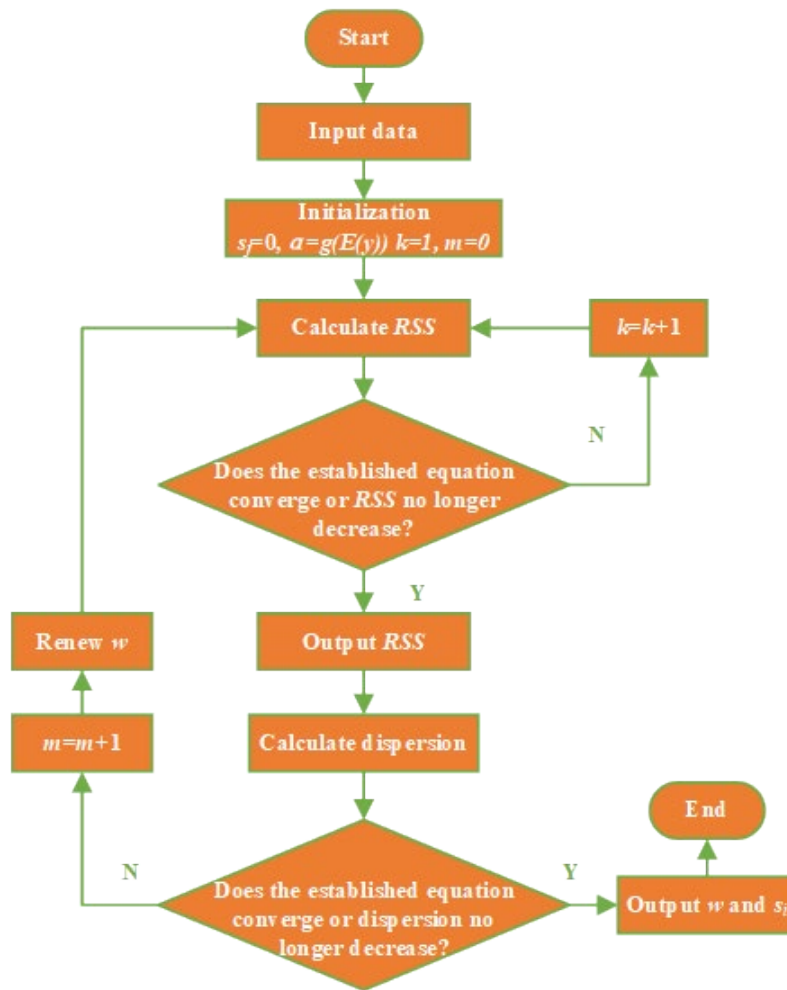


Fig.5. 1. Operation mechanism flow chart of the GAM regression algorithm.

5.2 Regression verification

By inputting the flow stress data into the GAM model for iteration, the GAM regression model for AA6061-T6 alloy was finally obtained. The α is determined as 2.8744×10^{-15} . The received GAM model was used to predict the flow stress under designed experiment deformation conditions, namely, $0.001 \text{ s}^{-1} \sim 1 \text{ s}^{-1}$ and $300 \text{ }^\circ\text{C} \sim 510 \text{ }^\circ\text{C}$. Fig.5. 2 compares the predicted stress with the stress data received from thermal compression tests.

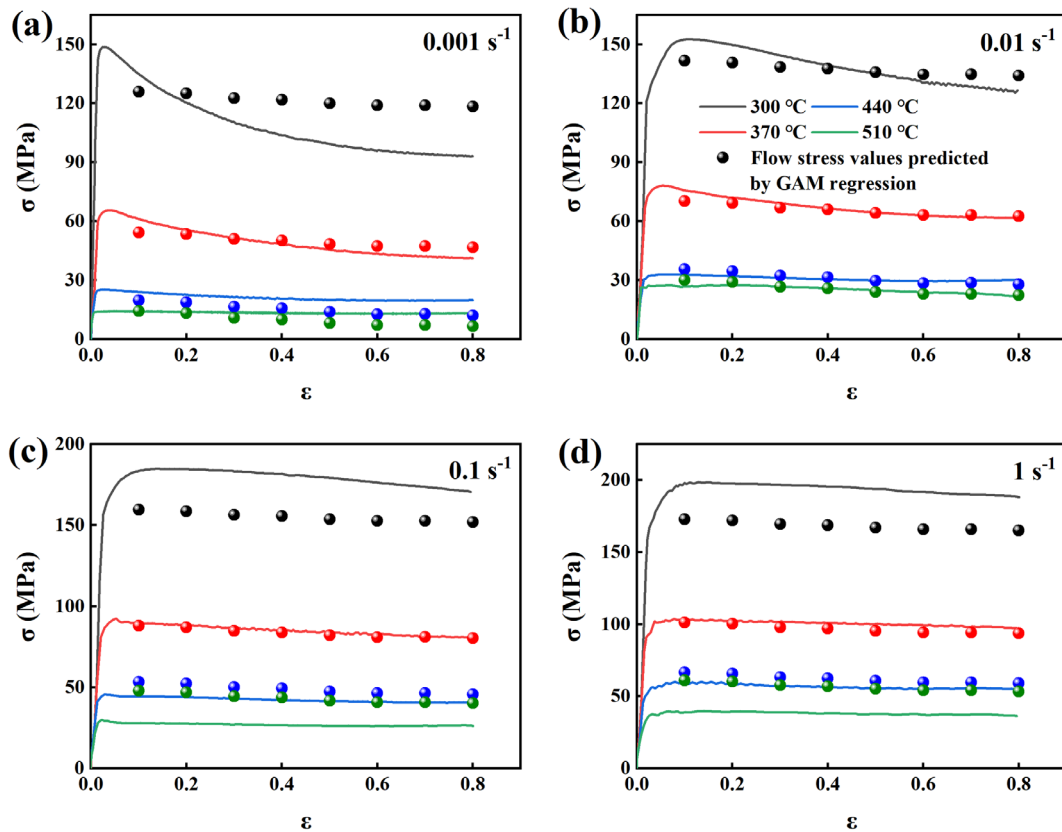


Fig.5. 2. Regression verification of the GAM model at different experimental deformation conditions, divided by strain: (a) 0.001 s^{-1} , (b) 0.01 s^{-1} , (c) 0.1 s^{-1} , and (d) 1 s^{-1} .

The fitting result of the GAM model is much better than the MLL model. At various deformation conditions like $370 \text{ }^{\circ}\text{C}$ and 0.01 s^{-1} , $440 \text{ }^{\circ}\text{C}$ and 0.01 s^{-1} , $510 \text{ }^{\circ}\text{C}$ and 0.01 s^{-1} , $370 \text{ }^{\circ}\text{C}$ and 0.1 s^{-1} , $370 \text{ }^{\circ}\text{C}$ and 1 s^{-1} . However, the GAM model loses the most accuracy at low-temperature deformation conditions. In deformation conditions $300 \text{ }^{\circ}\text{C}$ and 0.001 s^{-1} , $300 \text{ }^{\circ}\text{C}$ and 0.01 s^{-1} , the established model is cannot predict the flow trends. The predicted flow curves will overlap at high temperatures of $440 \text{ }^{\circ}\text{C}$ and $510 \text{ }^{\circ}\text{C}$. Therefore, this model may not be able to distinguish high-temperature flow characteristics.

6 SVR MODEL

SVM is a generalized linear classifier that processes binary data classification based on supervised learning. The decision boundary is the maximum-margin hyperplane for solving learning samples. The hinge loss is adopted to calculate the empirical risk, and the structure risk is optimized by adding a regularizer in the solving system, which endows its sparsity and robustness. The SVM can use the kernel method to process nonlinear classification, a commonly used kernel learning method [84].

6.1 SVM

6.1.1 Linear separability

SVM is based on linear separability. In a classification problem, the input data and learning objectives are given:

$$\begin{aligned} X &= \{x_1, \dots, x_N\}, & x_i &= [x_{i1}, \dots, x_{in}] \in \chi \\ y &= \{y_1, \dots, y_N\}, & y &\in \{-1, 1\} \end{aligned} \quad (6-1)$$

where X is the input data, y is the learning objective, X_i is the feature space, and y is the binary variable: negative and positive classes. When the classification problem meets the following condition, we regard the problem as linear separability: if the hyperplane regards as a decision boundary that separates the learning objects as a negative class and a positive

class exists in the input data feature space and the distances between arbitrary sample points and hyperplane greater than or equal to 1:

$$\begin{aligned} \text{Decision boundary: } & w^T X + b = 0 \\ \text{Point to hyperplane distance: } & y_i(w^T X_i + b) \geq 1 \end{aligned} \quad (5-2)$$

where the w is the normal vector of the hyperplane, and the b is the intercept of the hyperplane. Therefore, two parallel hyperplanes are constructed as interval boundaries to judge the classification of samples:

$$\begin{aligned} w^T x_i + b \geq +1, & \rightarrow y_i = +1 \\ w^T x_i + b \leq -1, & \rightarrow y_i = -1 \end{aligned} \quad (6-3)$$

All samples below the lower interval boundary belong to the negative class. Relatively, all samples above the upper interval boundary belong to the positive class. The distance between two interval hyperplanes is defined as margin: $d = \frac{2}{\|w\|}$. The samples at the hyperplanes are defined as support vectors. Then, the optimization problem of the SVM is obtained:

$$\begin{aligned} \max_{w, b} & \frac{2}{\|w\|} \\ \text{s. t. } & y_i(w^T x_i + b) \geq 1, \quad i = 1, 2, \dots, m \end{aligned} \quad (6-4)$$

In order to maximize the margin, only $\|w\|^{-1}$ needs to maximize, equal to minimize $\|w\|^2$.

Consequently, the equation yields:

$$\begin{aligned} \min_{w, b} & \frac{1}{2} \|w\|^2 \\ \text{s. t. } & y_i(w^T x_i + b) \geq 1, \quad i = 1, 2, \dots, m \end{aligned} \quad (6-5)$$

By adding Lagrange multipliers to each constrained condition, the dual problem will be obtained:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 + \sum_{i=1}^m \alpha_i (1 - y_i(w^T x_i + b)) \quad (6-6)$$

Take the partial derivative of w and b for the above equation, and set the partial derivative to 0, then:

$$\begin{aligned} w &= \sum_{i=1}^m \alpha_i y_i x_i \\ 0 &= \sum_{i=1}^m \alpha_i y_i \end{aligned} \quad (6-7)$$

Substitute **Eq. (6-7)** to **Eq. (6-6)**, then:

$$\begin{aligned} L(w, b, \alpha) &= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j x_i^T x_j + \sum_{i=1}^m \alpha_i \\ &\quad - \sum_{i=1}^m \alpha_i y_i \sum_{j=1}^m \alpha_j y_j x_j^T x_i - \sum_{i=1}^m \alpha_i y_i b \\ &= \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j x_i^T x_j - \sum_{i=1}^m \alpha_i y_i b \end{aligned} \quad (6-8)$$

The optimal term in the last equation is 0, the dual problem obtained finally:

$$\begin{aligned} \max_{\alpha} & \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j x_i^T x_j \\ \text{s. t.} & \sum_{i=1}^m \alpha_i y_i = 0, \\ & \alpha_i \geq 0, i = 1, 2, \dots, m \end{aligned} \quad (6-9)$$

Taking the derivative of α from the above equation, the final model will be obtained by solving w and b :

$$f(x) = w^T x + b = \sum_{i=1}^m \alpha_i y_i x_i^T x + b \quad (6-10)$$

Because **Eq. (6-5)** is an optimization problem with inequality constraint, the optimization needs to satisfy the Karush-Kuhn-Tucker (KKT) condition:

$$\begin{cases} \alpha_i \geq 0 \\ y_i(w^T x_i + b) - 1 \geq 0, & i = 1, 2, \dots, m \\ \alpha_i(y_i(w^T x_i + b) - 1) = 0, & i = 1, 2, \dots, m \end{cases} \quad (6-11)$$

Sequential minimal optimization (SMO) is used to solve **Eq. (6-9)**, a quadratic programming problem. The α can be obtained after iteration. For arbitrary support (x_s, y_s) , there is

$y_s(\sum_{i \in S} \alpha_i y_i x_i^T x_s + b) = 1$, $S = \{i | \alpha_i > 0, i = 1, 2, \dots, m\}$. b can be received by calculating the average value of all support vectors:

$$b = \frac{1}{|S|} \sum_{s \in S} \left(\frac{1}{y_s} - \sum_{i \in S} \alpha_i y_i x_i^T x_s \right) \quad (6-12)$$

Meanwhile, kernel functions are usually adopted in SVM, and then the corresponding model can be written as:

$$f(x) = w^T \phi(x) + b \quad (6-13)$$

The original objective function will be:

$$\begin{aligned} \min_{w, b} & \frac{1}{2} \|w\|^2 \\ \text{s. t. } & y_i (w^T \phi(x_i) + b) \geq 1, \quad i = 1, 2, \dots, m \end{aligned} \quad (6-14)$$

The dual function:

$$\begin{aligned} \max_{\alpha} & \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j \phi(x_i)^T \phi(x_j) \\ \text{s. t. } & \sum_{i=1}^m \alpha_i y_i = 0, \\ & \alpha_i \geq 0, i = 1, 2, \dots, m \end{aligned} \quad (6-15)$$

6.1.2 Loss function

The form of the SVM introduced above is on the basis that all samples are divided correctly, with the equation form $y_i(w^T x_i + b) \geq 1$, called hard margin. However, the problems in most cases are not linearly separable. The use of hyperplanes as the boundary will bring classification loss. That is, some support vectors fall on the wrong side of the decision boundary. At this time, the soft margin is introduced. When maximizing the margin, the soft margin needs to guarantee that the sample number should be as small as possible. Now, the optimization function is:

$$\min_{w, b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m l_{0/1}(y_i(w^T x_i + b) - 1) \quad (6-16)$$

where $C > 0$ is a constant, $l_{0/1}$ is the 0/1 loss function:

$$l_{0/1}(z) = \begin{cases} 1, & z < 0 \\ 0, & z \geq 0 \end{cases} \quad (6-17)$$

When C is infinite, **Eq. (6-16)** degenerates to hard margin SVM. When C is a finite value, **Eq. (6-16)** allows some samples do not meet the requirement.

Where $l_{0/1}$ is nonconvex and noncontinuous, which is hard to solve, loss functions such as exponential loss, logistic loss, and hinge loss are developed. In SVM, the hinge loss function is utilized:

$$l_{hinge}(z) = \max(0, 1 - z) \quad (6-18)$$

Substituting **Eq. (6-18)** to **Eq. (6-16)**:

$$\min_{w, b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \max(0, 1 - y_i(w^T x_i + b)) \quad (6-19)$$

Introducing slack variables $\xi_i = \max(0, 1 - y_i(w^T x_i + b)) \geq 0$, **Eq. (6-19)** yields to:

$$\begin{aligned} \min_{w, b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i \\ \text{s. t. } y_i(w^T x_i) \geq 1 - \xi_i, \quad \xi_i \geq 0, i = 1, 2, \dots, m \end{aligned} \quad (6-20)$$

Applying Lagrange multiplier method:

$$L(w, b, \alpha, \xi, \mu) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i + \sum_{i=1}^m \alpha_i (1 - \xi_i - y_i(w^T x_i + b)) - \sum_{i=1}^m \mu_i \xi_i \quad (6-21)$$

where $\alpha_i \geq 0, \mu_i \geq 0$ are Lagrange multipliers.

Applying partial derivative of w, b, ξ_i to 0:

$$\begin{aligned}
w &= \sum_{i=1}^m \alpha_i y_i x_i, \\
0 &= \sum_{i=1}^m \alpha_i y_i, \\
C &= \alpha_i + \mu_i
\end{aligned} \tag{6-22}$$

Substituting **Eq. (6-22)** to **Eq. (6-21)**, the dual problem of the SVR will be obtained:

$$\begin{aligned}
\max_{\alpha} \quad & \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j x_i^T x_j \\
s. t. \quad & \sum_{i=1}^m \alpha_i y_i = 0, \\
& 0 \leq \alpha_i \leq C, i = 1, 2, \dots, m
\end{aligned} \tag{6-23}$$

For soft margin SVM, there is KKT condition requirement:

$$\begin{cases}
\alpha_i \geq 0, \mu_i \geq 0, \\
y_i(w^T x_i + b) - 1 + \xi_i \geq 0, \\
\alpha_i(y_i(w^T x_i + b) - 1 + \xi_i) = 0, \\
\xi_i \geq 0, \mu_i \xi_i = 0.
\end{cases} \tag{6-24}$$

6.2 SVR

SVR was proposed as a branch of SVM. As explained in **Fig.6. 1**, SVM aims to find the largest margin from samples to the hyperplane. On the contrary, the SVR tries to find the smallest margin between the hyperplane and the samples. In other words, the two algorithms differ in defining objective and loss function.

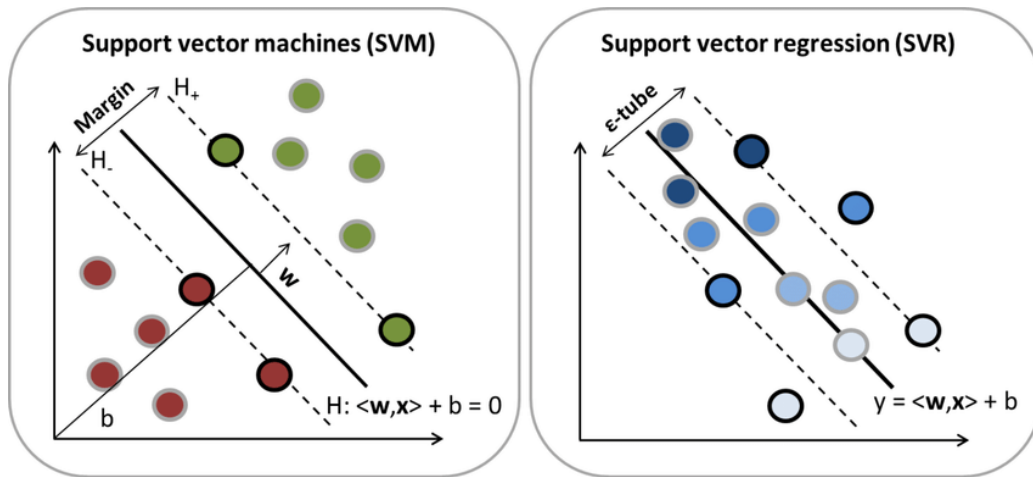


Fig.6. 1. The difference between SVM and SVR.

Unlike the traditional regression model that calculates the loss by outputting the difference between the model output $f(x)$ and the real value y , the SVR allows the error value ϵ between $f(x)$ and y . When the error between the $f(x)$ and y is over ϵ , the loss will be countered.

Consequently, the SVR problem can be written as follows:

$$\min_{w, b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m l_{\epsilon}(f(x_i) - y_i) \quad (6-25)$$

C is the regularization constant, l_{ϵ} is:

$$l_{\epsilon}(z) = \begin{cases} 0, & \text{if } |z| < \epsilon, \\ |z| - \epsilon, & \text{otherwise} \end{cases} \quad (6-26)$$

Introducing slack variable ξ_i and $\hat{\xi}_i$, **Eq. (6-25)** can be written as:

$$\begin{aligned} \min_{w, b, \xi_i, \hat{\xi}_i} &= \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \hat{\xi}_i) \\ \text{s. t. } & f(x_i) - y_i \leq \epsilon + \xi_i, \\ & y_i - f(x_i) \leq \epsilon + \hat{\xi}_i, \\ & \xi_i \geq 0, \hat{\xi}_i \geq 0, i = 1, 2, \dots, m \end{aligned} \quad (6-27)$$

Introducing Lagrange multipliers $\mu_i \geq 0, \hat{\mu}_i \geq 0, \alpha \geq 0, \hat{\alpha} \geq 0$, the Lagrange function will be obtained:

$$\begin{aligned}
L(w, b, \alpha, \hat{\alpha}, \xi, \hat{\xi}, \mu, \hat{\mu}) = & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \hat{\xi}_i) - \sum_{i=1}^m \mu_i \xi_i - \\
& \sum_{i=1}^m \hat{\mu}_i \hat{\xi}_i + \sum_{i=1}^m \alpha_i (f(x_i) - y_i - \epsilon - \xi_i) + \\
& \sum_{i=1}^m \hat{\alpha}_i (y_i - f(x_i) - \epsilon - \hat{\xi}_i)
\end{aligned} \quad (6-28)$$

Taking partial derivatives of the above equation:

$$\begin{aligned}
w &= \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) x_i, \\
0 &= \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i), \\
C &= \alpha_i + \mu_i, \\
C &= \hat{\alpha}_i + \hat{\mu}_i.
\end{aligned} \quad (6-29)$$

Substituting **Eq. (6-29)** into **Eq. (6-28)**, the dual problem of SVR is received:

$$\begin{aligned}
& \max_{\alpha, \hat{\alpha}} \sum_{i=1}^m y_i (\hat{\alpha}_i - \alpha_i) - \epsilon (\hat{\alpha}_i + \alpha_i) \\
& - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m (\hat{\alpha}_i - \alpha_i) (\hat{\alpha}_j - \alpha_j) x_i^T x_j \\
& s. t. \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) = 0, \\
& 0 \leq \hat{\alpha}_i, \alpha_i \leq C.
\end{aligned} \quad (6-30)$$

The solution of SVR is:

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) x_i^T x + b \quad (6-31)$$

The samples that meet $\hat{\alpha}_i - \alpha_i \neq 0$ are support vectors of SVR, which locate outside ϵ band.

In this case, support vectors are merely a part of the training samples. This keeps the sparsity

of the solution. From KKT, each sample has $(C - \alpha_i) \xi_i = 0$ and $\alpha_i (f(x_i) - y_i - \epsilon - \xi_i) =$

0. Therefore, when α_i is obtained, if $0 < \alpha_i < C$, there must be $\xi_i = 0$, then

$$b = y_i + \epsilon - \sum_{j=1}^m (\hat{\alpha}_j - \alpha_j) x_j^T x_i \quad (6-32)$$

Like the solution process of SVM, an approach with more robustness is adopted: choosing to multiply or all samples meet the condition $0 < \alpha_i < C$ to calculate b , then taking their average value.

Similarly, the model can be transformed into the following equation after introducing the kernel trick:

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_j - \alpha_j) \phi(x_i)^T \phi(x_i) + b \quad (6-33)$$

The SVR flow operation mechanism chart is illustrated in **Fig.6. 2**.

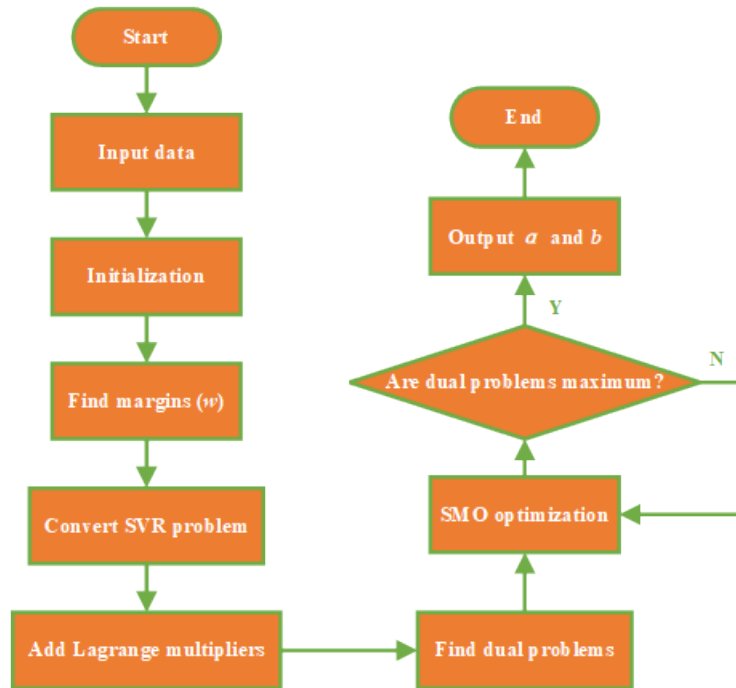


Fig.6. 2. Operation mechanism flow chart of the SVR algorithm.

6.3 Regression verification

The experimental flow data were imported into the SVR regression model for training. The Gaussian function was selected as the kernel function and the maximum iteration was set as 1000000. In order to eliminate the amount of dimensional quantity between different data and facilitate data comparison and joint processing before the training, the input data were standardized. The standardization follows the following equation:

$$X = \frac{X_{input} - \text{mean}(X)}{\text{std}(X)} \quad (6-34)$$

where X is the processed data for training, X_{input} is the import experimental data.

Ultimately, the SVM model was received after 77 iterations. The α values are listed in **Table 6-1**. The α value is -0.831394 and the bias b is 0.9011.

Table 6-1 α values in SVM model.

	α_{1i}	α_{2i}	α_{3i}	α_{4i}	α_{5i}	α_{6i}	α_{7i}
1	0.0725382	-0.088496	-0.085379	-0.834394	0.8343938	0.6773392	-0.834394
2	-0.509712	-0.30096	-0.054197	0.8343938	0.3036914	0.5838901	-0.074411
3	-0.699148	-0.094739	-0.182132	0.8343938	0.6768462	0.4148462	-0.296469
4	-0.52481	-0.380653	-0.834394	0.1926026	0.8343938	0.6612647	-0.069821
5	-0.556462	-0.284181	-0.428086	0.0373544	0.8343938	0.3004168	-0.69187
6	-0.449626	-0.132842	-0.038492	0.0884715	-0.093255	0.4339955	-0.257135
7	-0.834394	-0.834394	-0.812447	0.8343938	0.1037369	0.0133007	-0.834394
8	-0.611876	-0.174221	-0.250048	0.8343938	0.8343938	0.8343938	
9	-0.155883	-0.081201	-0.508321	0.0442538	0.8343938	0.8343938	
10	-0.733562	-0.078088	-0.332931	0.8188501	0.5054829	-0.069398	

Moreover, 201 support vectors were obtained during the iteration, see **Table 6-2**. The support vectors finally obtained are: $w_1 = 3.3928756$, $w_2 = -0.755554$, and $w_3 = 1.2354467$.

Table 6-2 Support vectors of the SVM model.

w_1			w_2			w_3		
-0.353574	-0.353574	-0.012988	-1.5954	0.9241379	0.9241379	-1.724346	-1.719544	-1.448208
-0.353574	-0.353574	3.3928756	-1.5954	0.9241379	-1.5954	-1.14009	-1.712118	-1.726772
-0.353574	-0.353574	3.3928756	-1.5954	0.9241379	-1.5954	-0.704448	-0.724498	-1.702069
-0.353574	-0.353574	3.3928756	-1.5954	0.9241379	-1.5954	-0.080787	-0.435093	-1.677267

-0.353574	-0.353574	3.3928756	-1.5954	0.9241379	-1.5954	0.3550039	0.505845	-0.549251
-0.353574	-0.353574	3.3928756	-1.5954	0.9241379	-1.5954	0.6321812	0.5824289	-0.433855
-0.353574	-0.353574	3.3928756	-1.5954	0.9241379	-1.5954	1.5827723	1.7411382	-0.238608
-0.353574	-0.353574	3.3928756	-1.5954	0.9241379	-1.5954	1.7215838	1.7459896	0.5508449
-0.353574	-0.012988	3.3928756	-0.755554	-1.5954	-1.5954	-1.724247	-1.735633	0.6510921
-0.353574	-0.012988	3.3928756	-0.755554	-1.5954	-1.5954	-1.694643	-1.710782	0.7011909
-0.353574	-0.012988	3.3928756	-0.755554	-1.5954	-1.5954	-0.684597	-0.867963	1.6728711
-0.353574	-0.012988	3.3928756	-0.755554	-1.5954	-1.5954	-0.565835	-0.818656	1.6979205
-0.353574	-0.012988	3.3928756	-0.755554	-1.5954	-1.5954	-0.239103	-0.669745	1.7230689
-0.353574	-0.012988	3.3928756	-0.755554	-1.5954	-0.755554	0.3450534	-0.62034	-1.672762
-0.353574	-0.012988	3.3928756	-0.755554	-1.5954	-0.755554	0.6224783	-0.545884	-1.650435
-0.353574	-0.012988	3.3928756	-0.755554	-1.5954	-0.755554	1.7313857	-0.496528	-1.533604
-0.353574	-0.012988	3.3928756	-0.755554	-1.5954	-0.755554	1.7412867	0.3460435	-0.662023
-0.353574	-0.012988	3.3928756	0.084292	-1.5954	-0.755554	-1.724247	0.4648055	-0.353806
-0.353574	-0.012988	3.3928756	0.084292	-1.5954	-0.755554	-0.625241	0.5840131	-0.048213
-0.353574	-0.012988	3.3928756	0.084292	-1.5954	-0.755554	-0.308954	1.6090594	1.1844072
-0.353574	-0.012988	3.3928756	0.084292	-1.5954	-0.755554	0.4441125	1.6847522	1.2354467
-0.353574	-0.012988		0.084292	-1.5954		0.572825	1.7335639	
-0.353574	-0.012988		0.084292	-0.755554		1.7414847	-0.496479	

The constructed SVM model was adopted to predict the flow stress under experimental conditions, and the predicted values were compared with the experiment flow behavior, as seen in **Fig.6. 3**.

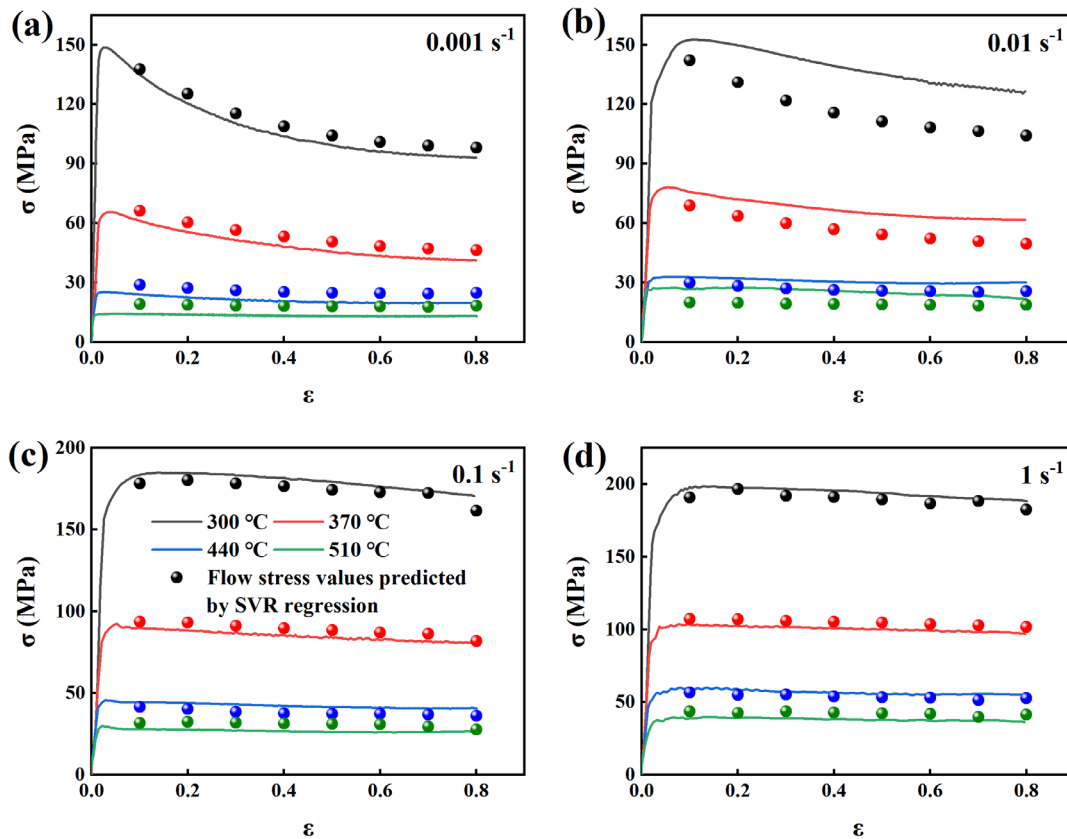


Fig.6. 3. Regression verification of the SVR model at different experimental deformation conditions, divided by strain: (a) 0.001 s^{-1} , (b) 0.01 s^{-1} , (c) 0.1 s^{-1} , and (d) 1 s^{-1} .

Based on the observation on the plots, the SVM model can predict the deformation behavior of aluminum alloys under most experimental conditions. The model is able to describe the flow behavior change along the strain. However, some large differences occur in the conditions of $300 \text{ }^{\circ}\text{C}$ and 0.01 s^{-1} , and $370 \text{ }^{\circ}\text{C}$ and 0.01 s^{-1} . In general, the SVM model is sufficient to fit and predict the flow behavior of the AA6061-T6 alloy under deformation conditions $300\sim 510 \text{ }^{\circ}\text{C}$ and strain rate $0.001\sim 1 \text{ s}^{-1}$.

7 RT MODEL

Before explaining the operation algorithm of RT, it is better to introduce DT first, which is because RT is derived from DT. DT is a typical supervised machine learning classification algorithm, and meanwhile, it is a non-parametric model. The DT is a tree structure where each internal node represents a judgment on an attribute, each branch represents the output of a judgment result, and finally, each leaf node represents a classification result. Decisions can be divided into ID3, C4.5, and CART according to different generation algorithms (classification error, entropy, and Gini index) [85]. Besides, the regression work only can be processed on the CART algorithm. **Fig. 7.1** illustrates the clear classification.

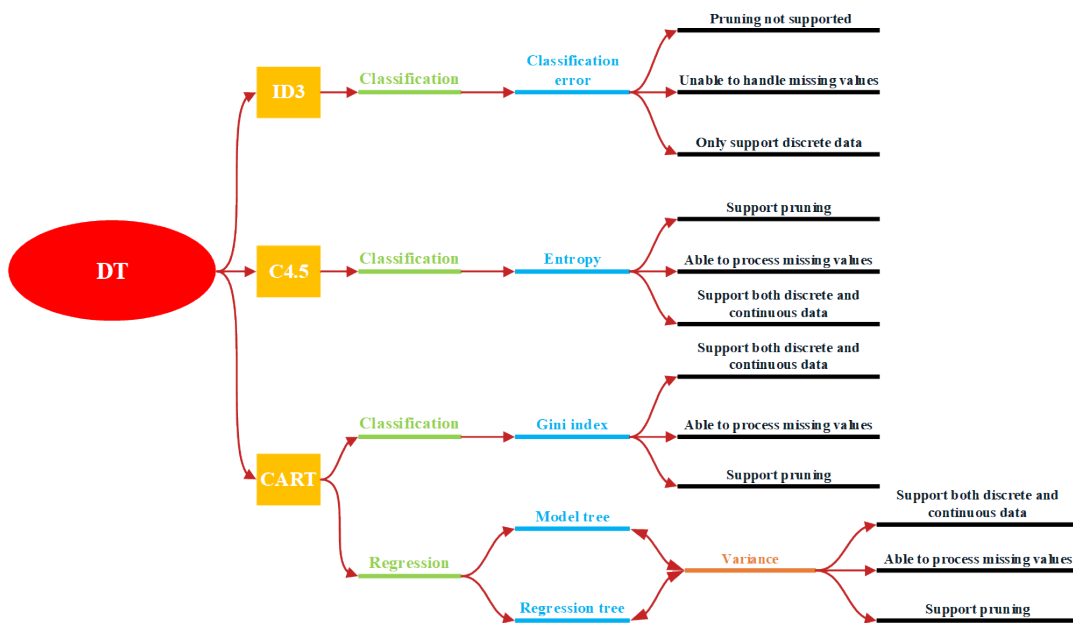


Fig.7. 1. Types of DT.

In the case of DT classification, no domain knowledge or parameter setting is required, so it is suitable for exploratory knowledge discovery. In addition, most other classification methods need data normalization processing. While for the DT, data normalization is not necessary. In addition, the DT can be visualized to see the final tree form. However, overfitting may occur for the DT results. Through tree pruning methods or setting reasonable tree depth and the number of leaf nodes can avoid overfitting.

7.1 DT

The decision function of classification model of DT can be formulated to:

$$y(x_n) = \sum_{m=1}^M c_m \cdot \Pi \{x_n \in R_m\} \quad (7-1)$$

where y is the mapping function, x is the sample point, c is the class, and R is the sample space.

We assume that the DT is a binary tree. According to different classification standards, different sample spaces can be divided into two different categories, see **Fig.7. 2**.

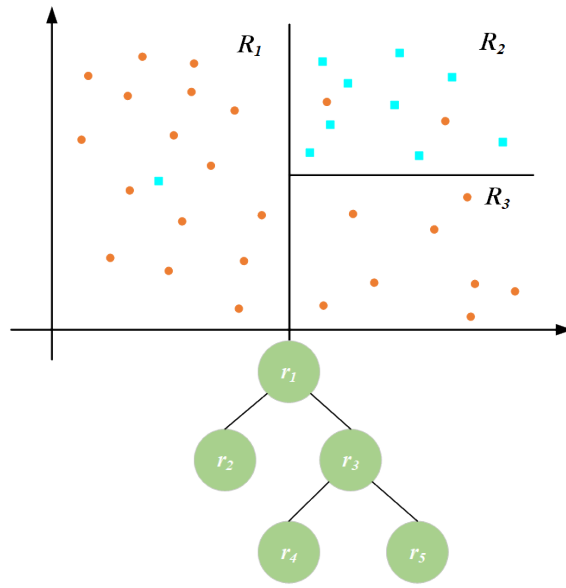


Fig.7. 2. Classification of a binary tree.

From Fig.7. 2, when a particular class accounts for the majority of the space, we divide the space into the class. In a mathematic form:

$$c_m = \max_k \sum_{x_n \in R_m} \Pi \{t_n = k\} \quad (7-2)$$

where k is the class after classification, t is the sample point. For binary classification, $k=1, 2$.

At the same time, the posterior probability of the class can be output as well:

$$p_k^m = \frac{1}{|R_m|} \sum_{x_n \in R_m} \Pi \{t_n = k\} \quad (7-3)$$

$$p(C_k | \mathbf{x}) = \sum_{m=1}^M p_k^m \cdot \Pi \{\mathbf{x} \in R_m\}$$

Where p is the class posterior probability. As a top-down greedy heuristic search algorithm, decision trees need to find the best classification feature for each classification. This is to say, the classification needs to meet the following condition:

$$\begin{aligned} R^-(j, s) &= \{x_i \in R | x_i(j) \leq s\} \\ R^+(j, s) &= \{x_i \in R | x_i(j) > s\} \end{aligned} \quad (7-4)$$

where j is the classification feature, s is the threshold. To determine the best classification feature, the effects (loss) of before and after classification need to be evaluated first.

Different DT types have different evaluation functions:

$$\left\{ \begin{array}{ll} \text{Classification error:} & 1 - \max_k p_k^m \quad \text{ID3} \\ \text{Entropy:} & - \sum_k p_k^m \ln p_k^m \quad \text{C4.5} \\ \text{Gini index:} & 1 - \sum_k (p_k^m)^2 \quad \text{CART} \end{array} \right. \quad (7-5)$$

Then, the loss can be calculated and the minimum loss can be found:

$$\begin{aligned} L(j, s) &= (p_{R_m^-} \cdot E(R_m^-) + p_{R_m^+} \cdot E(R_m^+)) - E(R_m) \\ (j^*, s^*) &= \arg \min_{(j, s)} L(j, s) \end{aligned} \quad (7-6)$$

where L is the loss function, $p_{R_m^-}$ and $p_{R_m^+}$ are the posterior probabilities after binary classification, $E(R_m^-)$ and $E(R_m^+)$ are sample spaces after binary classification, $E(R_m)$ is the space before the classification.

7.2 RT

The mathematical expression of RT also can be written in the **Eq. (7-1)** form. **Fig.7. 3** shows the graphical RT three times regression growing.

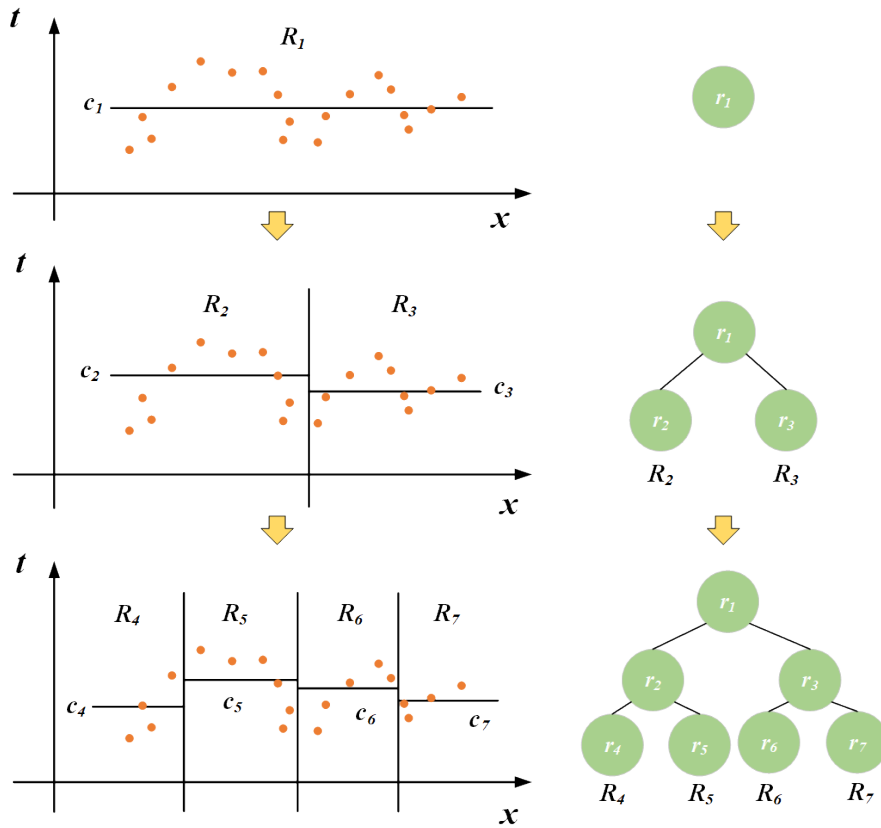


Fig.7. 3. Regression process of RT (three times growing).

The difference between DT and RT is the definition of the loss function. In RT, the variance is used as the evaluation indicator:

$$\begin{aligned}
 L(c_m) &= \arg \min_{c_m} \sum_{x_n \in R_m} (t_n - c_m)^2 \\
 &= \text{ave}(t_n | x_n \in R_m)
 \end{aligned}
 \tag{7-7}$$

Then, for each decision space:

$$E(R_m) = \sum_{x_n \in R_m} (t_n - c_m)^2 = \sum_{x_n \in R_m} (t_n - \text{ave}(t_n | x_n \in R_m))^2
 \tag{7-8}$$

The **Eq. (7-6)** also applies here. The operation flow chart of DT (RT) is displayed in **Fig. 7.4**.

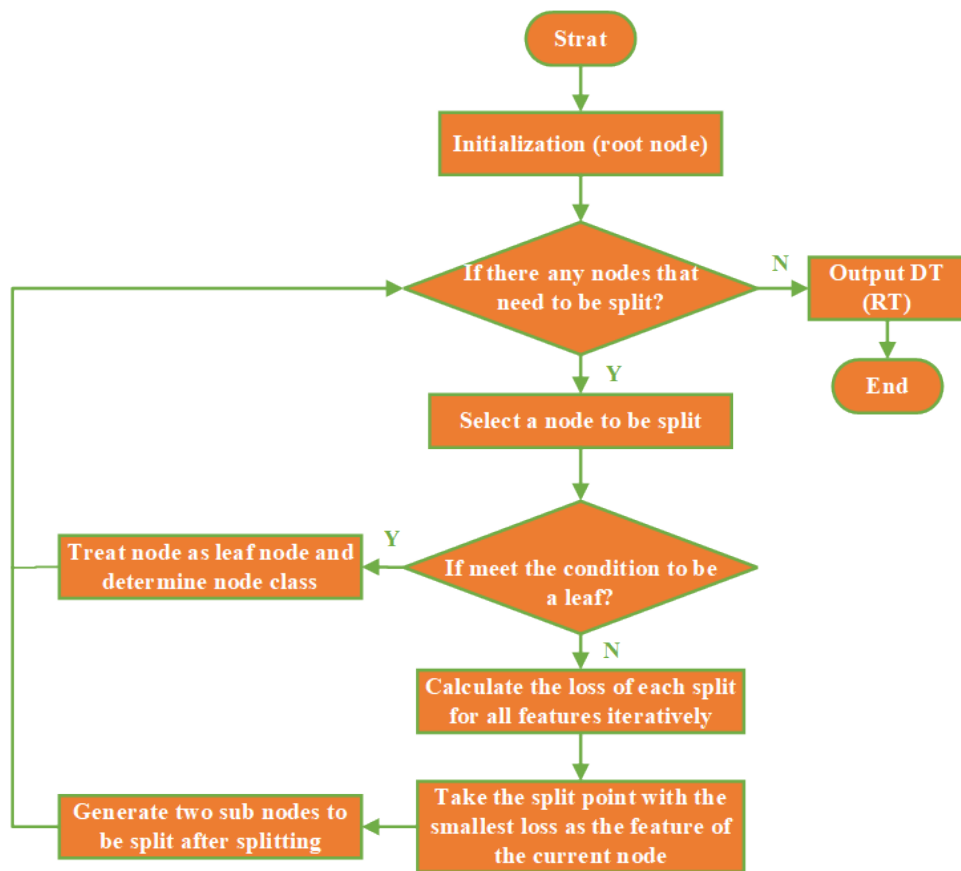


Fig.7. 4. Operation mechanism flow chart of DT (RT).

7.3 Regression verification

The received experimental flow stress data were imported to the RT. The split time was controlled to determine the best tree depth to check the error change. The split number was set from 1 to 100, and the *MSE* was considered as the error evaluation standard, see Fig.7. 5.

The equation of *MSE* is [86]:

$$MSE = \frac{1}{n} \sum_{i=1}^n (p - y)^2 \quad (9-22)$$

where p is the final predicted values. y is the expected output data. where n is the number of the test data.

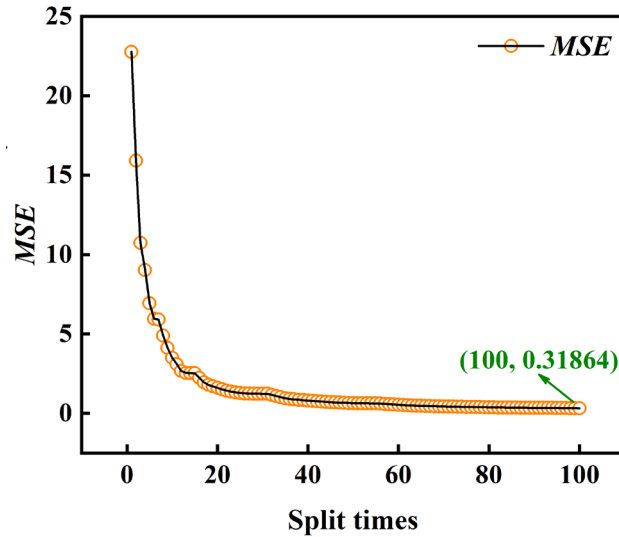


Fig.7. 5. The *MSE* variation with the split times.

The *MSE* shows a down trend along the split times. The decrease rate is highest in the initial stage and then gradually slows down as the number of splits increases. However, the *MSE* has always shown a decreasing trend. This is attributed to the fact that as the depth of the RT deepens, the model's classification of features becomes more refined, and the model's fitting ability also becomes more robust. In order to avoid overfitting caused by the boundless split of the RT, the split is terminated 100 times. At this point, the *MSE* of RT is 0.31864, which is a relatively low value to announce the regression reliability of the established RT model.

Eventually, an RT with 100 branch nodes and 101 leaf nodes was built. Fig.7. 6 displays the structure of the RT model (Triangles represent branch nodes, and points represent leaf nodes).

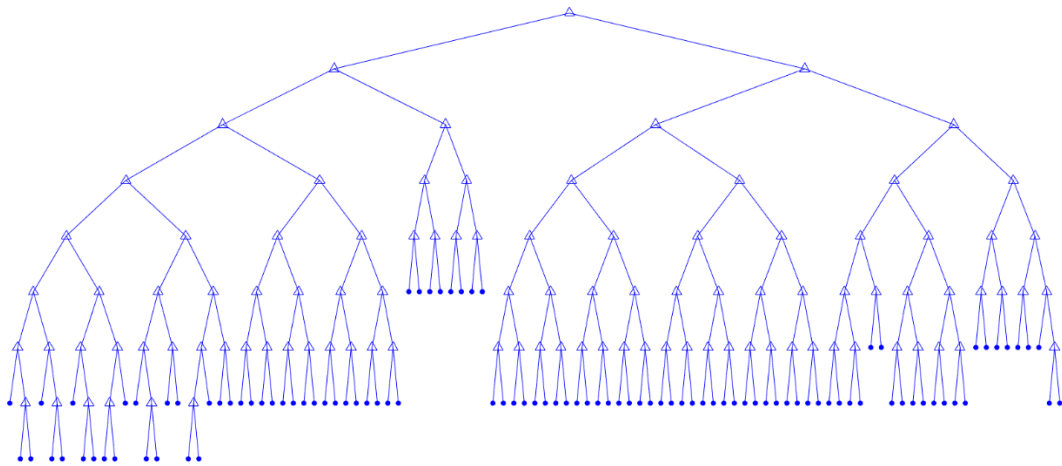


Fig.7. 6. The structure of the RT model.

The parameters of the RT model are listed in **Table 7-1**. The values marked by red color are parameters of leaf nodes. Additionally, in each branch node, the left branch represents categories that are less than this value, and the right branch represents categories that are greater than or equal to this value. Each leaf node corresponds to a specific numerical value.

Table 7-1 The parameters of the RT model.

1	2	3	4	5	6	7	8	9
-1.17548	1.68994	-0.183281	-0.709349	-1.27375	-1.52128	-1.64009	1.88647	1.82372
	-0.335631	0.364187	-0.236281	0.181935	-1.01623	-1.41239	-1.57081	1.78028
		-0.183281	-0.48969	-0.905289	-0.303261	-1.15492	1.72785	1.68225
		-0.183281	-0.4419	0.628642	0.795745	-0.857839	-1.36289	1.64605
			1.68994	-0.837294	-1.25227	-0.531058	1.59303	1.54563
			0.504215	0.0144604	-0.607765	-0.115292	-1.07581	1.51247
			0.504215	0.713617	0.209757	0.478939	-0.937022	1.48155
				1.46178	1.27512	1.14223	1.40059	1.44582
				-1.17472	2.97651	-1.42578	1.34343	1.2971
				0.439212	2.949	-1.054	-0.39237	1.26914
				-0.434449	2.91232	-0.781379	1.23295	1.14697
				-0.376949	2.87323	-0.409647	1.1838	1.11918
				-0.64982	2.83144	-0.013139	0.300549	
				-0.559523	2.78689	0.474558	1.08619	
				-0.549251	2.74293	0.927948	1.05338	
				0.149881	2.69903	1.49804	1.02154	
					-1.46187	-1.62031	2.28615	
					-0.847938	-1.33313	2.25362	

-0.095416	-1.00638	2.21842
0.983863	-0.639993	2.1792
-1.12836	-0.263707	2.14643
0.383246	0.201613	2.10877
-1.00328	0.69676	2.06614
0.366068	1.34037	2.0263
-1.23412	-1.45044	1.9861
-0.075638	-0.781304	1.94446
-1.03747	-0.087544	1.90341
-0.118831	0.853888	1.86455
-1.14395	-1.31613	1.82252
-0.004104	-0.709473	1.78834
-0.593632	-0.106529	1.75735
0.928071	0.916215	1.72848
	-1.48175	0.306524
	-0.996403	0.268798
	-0.551155	0.234577
	-0.572207	0.203464
	-1.3617	0.174527
	-0.841676	0.145197
	-0.277247	0.112707
	0.173099	0.0837158
	-0.300887	0.0575
	-0.321291	0.0329769
	-0.341561	0.0060609
	-0.356947	-0.024385
	-0.414199	-0.047418
	-0.441829	-0.068992
	-0.474372	-0.087772
	1.4447	-0.104658
		0.614859
		0.562012
		0.522889
		0.478836
		0.439219
		0.4007
		0.367269
		0.340353
		0.925581
		0.90221
		0.877045
		0.857293
		0.840807
		0.818002

0.790675
0.766203
-0.489277
-0.504683
-0.518935
-0.533291
-0.695178
-0.699547
-0.702984
-0.706319
-0.709143
-0.712129
-0.714438
-0.716827
-0.496005
-0.519763

The flow stress outputs from the constructed model are compared with the original experiment flow stress values, and the comparison result is shown in **Fig.7. 7**. The magnitude of flow stress and the trend of stress variation are highly consistent with experimental data. All in all, the RT model can fully predict the flow stress variation of the studied alloy under experimental plastic deformation conditions. However, as we can see in **Fig.7. 6** and **Table7-1**, in order to achieve high accuracy, the RT model needs to be cumbersome. This will lead to the consumption of computer computing power during the simulation process.

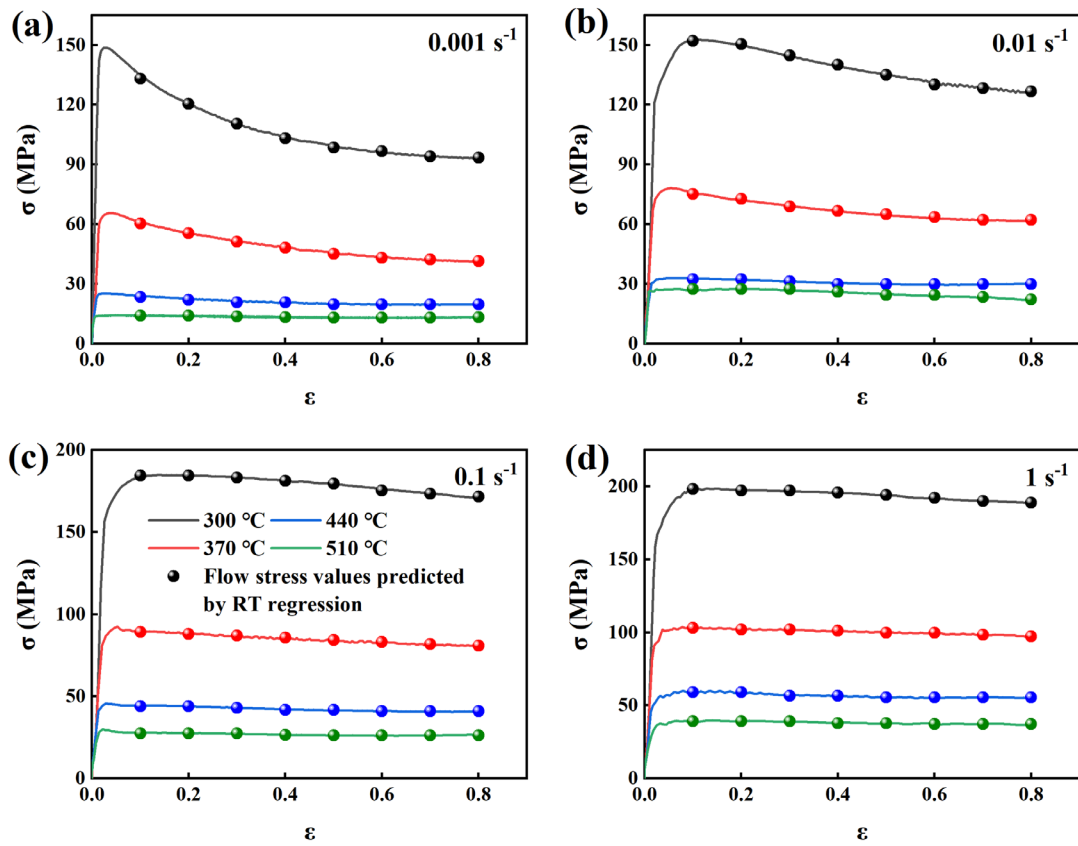


Fig. 7. 7. Regression verification of the RT model at different experimental deformation conditions, divided by strain: (a) 0.001 s^{-1} , (b) 0.01 s^{-1} , (c) 0.1 s^{-1} , and (d) 1 s^{-1} .

8 RF MODEL

RF is an ML algorithm based on DT, which adopts many DTs to make decisions. Moreover, RF is a typical bagging learning process combining multiply single models (weak learning) outputs [87]. In the bagging algorithm, the substitution methods are always adopted to select random samples in the training set, which means single data points can be selected many times. After generating multiple samples, the models will be trained separately. Based on task types (regression or classification), the estimated values with more accuracy will be generated because of the average predicted values or multiple values. The RF is composed of randomly created DTs. Each node in the DT is a random subset of features used to calculate the output. The working mechanism of RF integrates the outputs of single DTs to generate the final output result. In brief, the RF algorithm uses multiple (randomly generated) DTs to generate the final output. In addition to bagging learning, another assembled learning model named boosting learning exists. Different from the random selection thought of bagging learning, boosting trees enhances the classification or regression ability of models by concatenating multiple CART tree patterns. Boosting trees adjust weights based on the error from the last CART tree to minimize the error step by step [88]. As a consequence, boosting trees cannot solve the overfitting problem of DT to some extent. Besides, the DT receives an excellent regression result in the last chapter, and there is no need to introduce boosting trees to reduce the regression error a step more.

As mentioned in Chapter 7, overfitting is a tricky problem that constantly occurring in DT. However, when multiple DTs consist of an RF algorithm, their prediction results will be more accurate, especially when the DTs are unrelated. Furthermore, RF is able to deal with classification and regression problems with high precision. Since feature bagging can stay accurate when some data is missing, the RF classifier is always used as an effective tool for predicting missing values. However, because data in every DT in RF is required to be calculated, the processing process can be pretty time-consuming.

8.1 RF

The construction of RF consists of two parts: random selection of data and random selection of selected features, see Fig.8. 1.

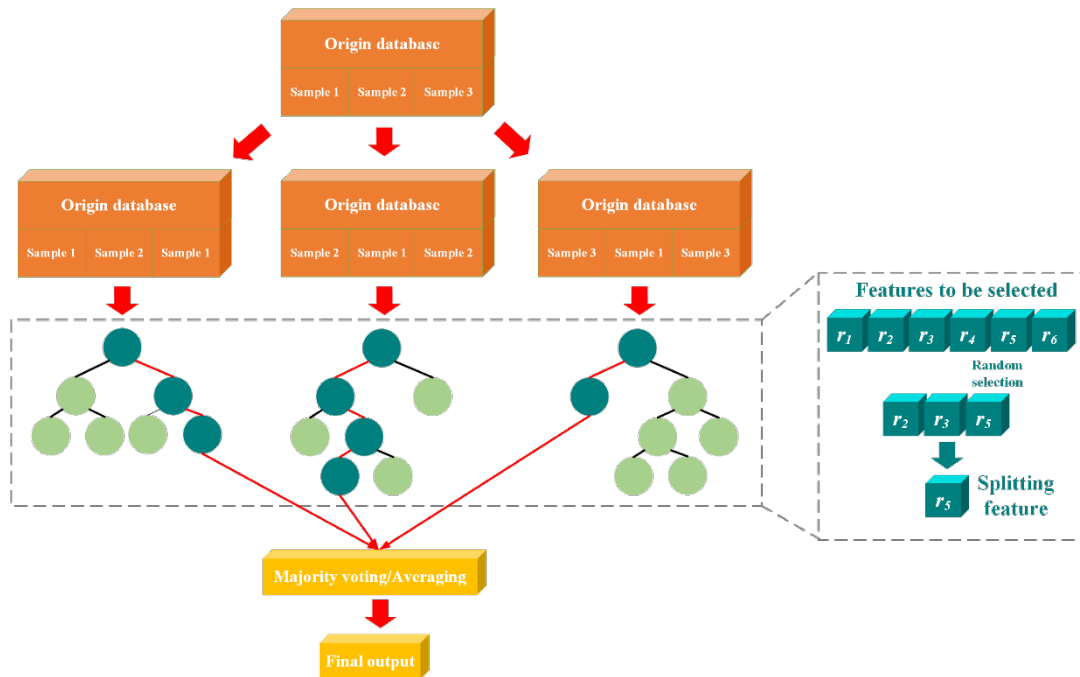


Fig.8. 1. Data random selection and feature random selection in RF.

The sub-datasets are constructed by sampling with returns from the origin database. The data volume of the sub dataset keeps the same as the original database. Samples in different sub-datasets can be duplicated, and the samples in the same sub-dataset can be duplicated as well. After that, the sub-datasets are utilized to construct sub-RTs, and each sub-RT outputs a result based on the data in the sub-dataset. Finally, the constructed sub-RTs form an RF. When a new sample datum is input to ask a classification result through RF, the RF will output it by voting on the judgment results of sub-decision trees. In the regression task, the average value of sub-RT will be the output.

Similar to the random selection of the dataset, each splitting process of the subtree in the RF does not use all the features to be selected but randomly selects a certain feature from all the features to be selected, and then selects the best feature from the randomly selected features. This process will guarantee that the RTs in the RF are different from each other, improve the diversity of the system, and thus improve the regression performance. The flow chart of RF operation is illustrated in **Fig.8. 2**.

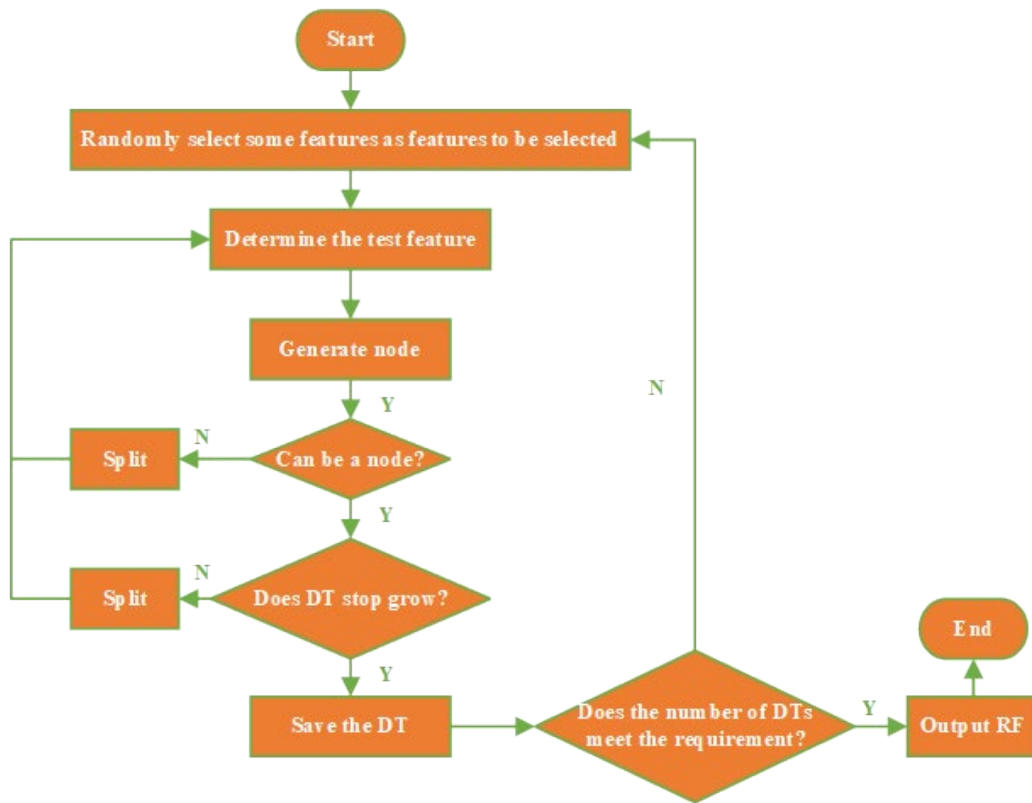


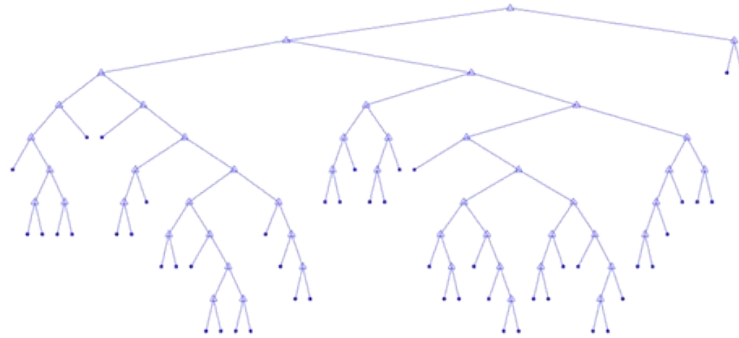
Fig.8. 2. Operation mechanism flow chart of RF.

8.2 Regression verification

As a bagging RT algorithm, the bagging quantity of RTs must be determined in advance. In the present work, the bagging quantity was set as 100. A RF containing 100 regression trees is constructed after a random selection combination with returns. Taking the first and second RTs as examples, Fig.8. 3 shows the structures. From the graph, it can be seen that the selected features influence the trend and depth of RT splitting. This randomness ensures the advantage that RF is challenging to fall into overfitting. Since the structure of the RF is a

nonparametric model, and its structure is highly complex, no specific model expression will be given here.

(a)



(b)

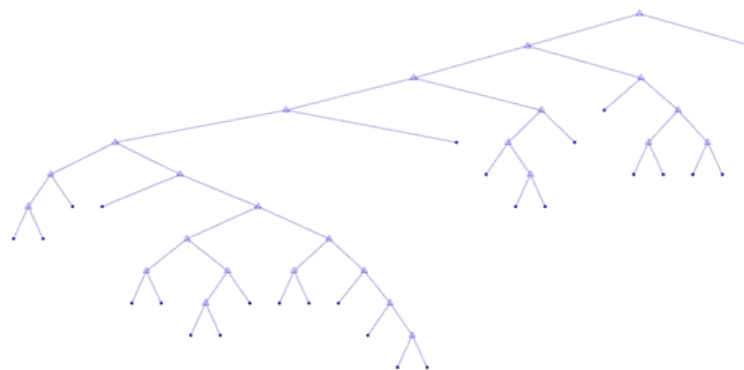


Fig.8. 3. The structure of the first RT and the structure of the second RT

Eventually, the flow data were extracted from the RT model by inputting the corresponding strains, temperatures, and strain rates. The data points were plotted on the experimental flow behavior curves. **Fig.8. 4** displays the regression effect. The fitting result is not as good as RT model. Except for the four deformation conditions (440 °C and 0.001 s⁻¹, 440 °C and 0.1 s⁻¹, 440 °C and 1 s⁻¹, and 370 °C and 0.01 s⁻¹), the fitting performance under all other deformation conditions is not ideal. This model has the worst accuracy in predicting stress at low temperatures. At the deformation conditions 440 °C and 0.01 s⁻¹, 510 °C and 0.01 s⁻¹, the

model gives the different stress change trends. Furthermore, at strain rates 0.001 s^{-1} , and 0.01 s^{-1} , the established RF model may fail to distinguish the flow stress values.

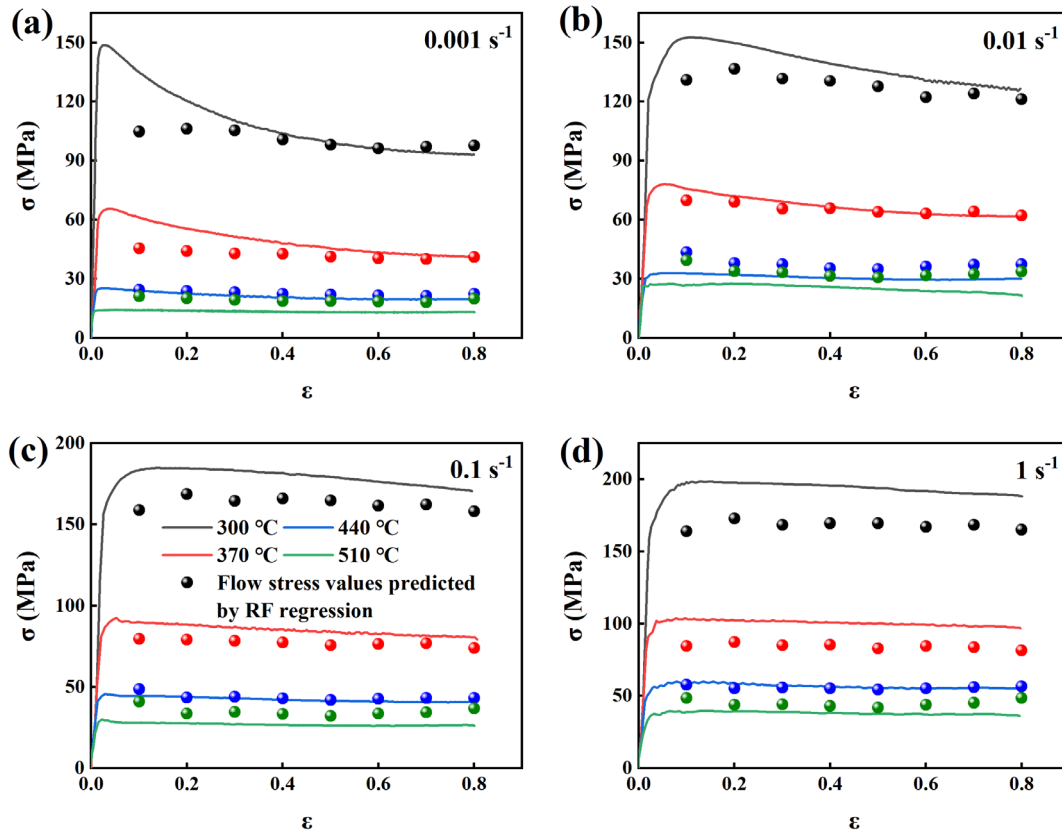


Fig.8. 4. Regression verification of the RT model at different experimental deformation conditions, divided by strain: (a) 0.001 s^{-1} , (b) 0.01 s^{-1} , (c) 0.1 s^{-1} , and (d) 1 s^{-1} .

However, Due to the non-parametric nature and complexity of the RT and RF models, there should be difficulties in incorporating the established models into simulation software. There is no report about substituting tree models to the simulation environments. Consequently, encoding tree models into numerical simulation software is a research problem that needs to be overcome.

9 MLP MODEL

Perceptron is a binary linear classifier and is regarded as the simplest feedforward neural network [89]. The structure of perceptron comes from biological neurocytes. A single neurocyte can be regarded as a machine with two outputs according to its activation condition: yes or no. The activation condition of neurocytes depends on the received semaphore and the effect of synapses. When the received semaphore exceeds a threshold, the neurocyte will be activated, and the signal pulse will be transferred to other neurons. With the bias, weight, and activation function corresponding to the threshold, synapse, and cell body of the neurocyte, the concept of perceptron was proposed.

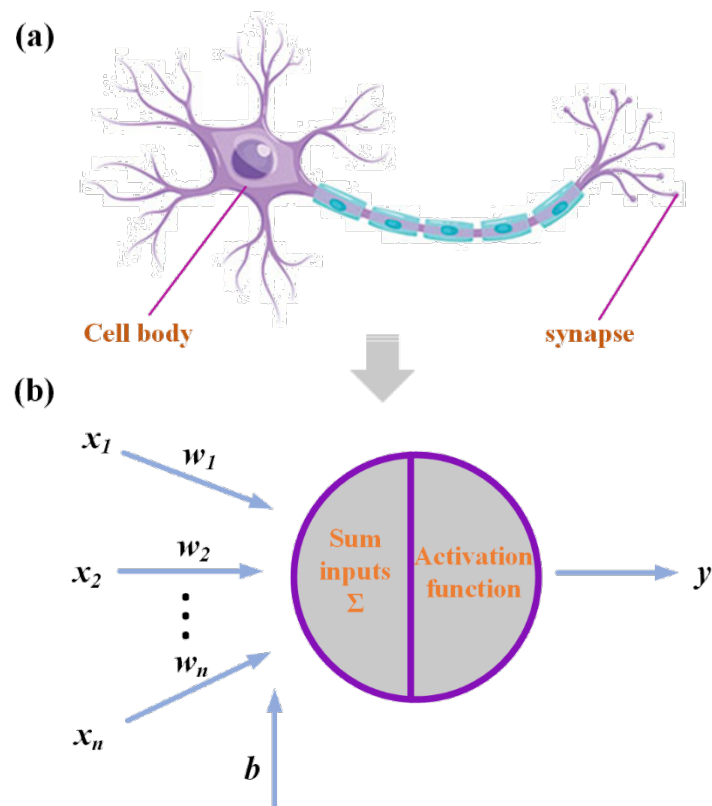


Fig.9. 1. Perceptron (b), inspired by neurocyte (a).

Fig.9 .1(b) shows the structure of a perceptron, which can be written in a mathematical form:

$$\begin{aligned}
 y &= f\left(\sum_{i=1}^m w_i x_i + b\right) = f(W^T X) \\
 W &= [w_1 \ w_2 \ \dots \ w_n \ b]^T \\
 X &= [x_1 \ x_2 \ \dots \ x_n \ 1]^T
 \end{aligned}
 \tag{9-1}$$

where y is the scalar output, f is the activation function, w is the weight, x is the input, b is the bias. As an anti-symmetric sign function, the definition of $f(n)$ is:

$$f(n) = \begin{cases} +1 & \text{if } n \geq 0 \\ -1 & \text{if } n < 0 \end{cases}
 \tag{9-2}$$

In order to deal with the nonlinear problems, the activation functions introduce the nonlinearity to the outputs of neurons. There are three activation functions mainly used:

$$\left\{ \begin{array}{l} \text{Sigmoid: } \frac{1}{1 + e^{-x}} \\ \quad \text{tanh: } \frac{e^x - e^{-x}}{e^x + e^{-x}} \\ \quad \text{ReLU: } \max(0, x) \end{array} \right.
 \tag{8-3}$$

The output of the perceptron is a scalar result through an activation function after obtaining the inner product of the input vector and weight vector.

9.1 MLP

Two types of perceptions can be obtained through the combination of perceptions: SLP and MLP. SLP has no hidden layer inside it, and is the simplest feedforward neural network. Different from SLP, there is one hidden layer at least in MLP [35].

The feedforward neural network is the first and simplest ANN. It contains multiple neurons (nodes) arranged in multiple layers. The nodes of adjacent layers have connections or edges.

All connections are equipped with weights. There are three types of neurons in a feedforward neural network: input, hidden, and output. The neurons receive information from outside called input nodes. The input neurons do not do any calculations, which deliver information to hidden neurons. Hidden neurons have no connection with the outside world, which take the job of calculating and delivering the information from input neurons to output neurons. The output neurons undertake the task of computing and transmitting information to the outside world. The input, hidden, and output neurons consist of input, hidden, and output layers, respectively. A feedforward network has one input layer, one output layer, and zero or more hidden layers. In a feedforward network, the information is transmitted sequentially along the input, hidden, and output layers, and there is no loop within the network.

Different from the models mentioned before, the operation of the MLP model is a black box. People cannot explain how the parameters are obtained, and parameters inside the MLP model have no meaning.

9.1.1 Training of MLP

The training of MLP is based on the gradient descent algorithm. It is assumed that p is the function of x , and the relationship between the independent and dependent variables follows the following equation:

$$\Delta p \approx \frac{\partial p}{\partial x_1} \Delta x_1 + \frac{\partial p}{\partial x_2} \Delta x_2 + \cdots + \frac{\partial p}{\partial x_m} \Delta x_m \quad (9-4)$$

In a vector form, the equation can be rewritten as:

$$\begin{cases} \Delta p = \nabla p \Delta x \\ \nabla p = \left[\frac{\partial p}{\partial x_1} \quad \frac{\partial p}{\partial x_2} \quad \dots \quad \frac{\partial p}{\partial x_m} \right] \\ \Delta x = [\Delta x_1 \quad \Delta x_2 \quad \dots \quad \Delta x_m]^T \end{cases} \quad (9-5)$$

In each iteration, the value of Δp should be negative to guarantee the decrease of the loss function. As ∇p is the function of Δx , we can assume that $\nabla p = -\eta \Delta x$, then:

$$\Delta p = -\eta \Delta x \cdot \Delta x \quad (9-6)$$

where η is the learning rate. In each iteration, the gradient Δp of p can be calculated, and the variable will be updated to take a step toward the fastest descent. The backpropagation algorithm is adopted to update weights in each iteration. Taking two hidden layers MLP as an example, the Fig. 9.2 illustrates the structure of the MLP.

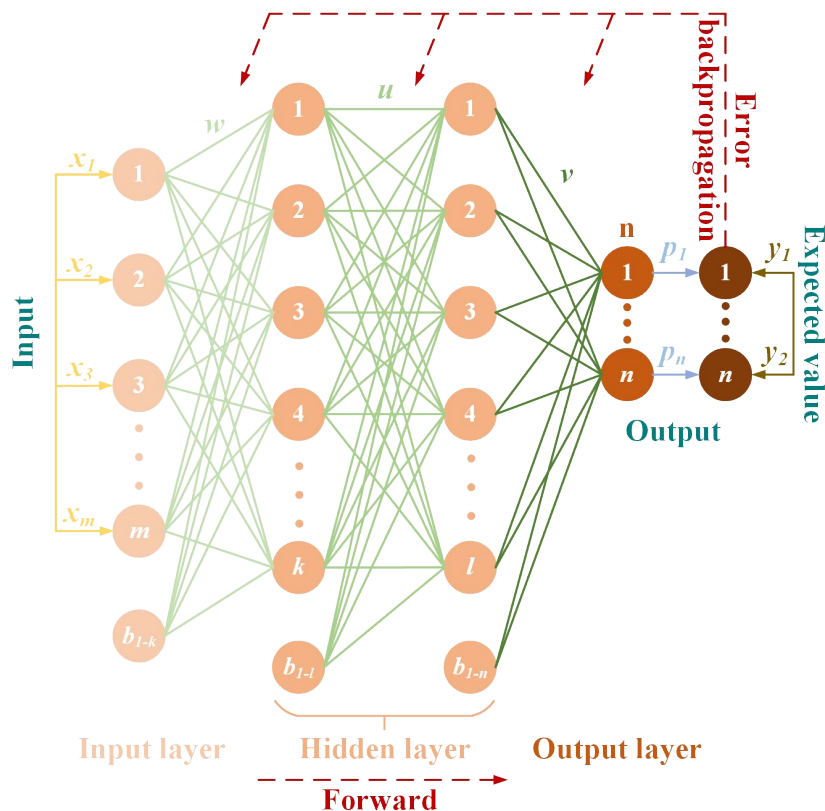


Fig.9. 2. Topological structure of a two hidden layer MLP.

In the feedforward propagation process, the output of the first hidden layer is:

$$[p_{11} \dots p_{1k}] = f \left([x_1 \dots x_m] \begin{bmatrix} w_{11} & \dots & w_{1k} \\ \vdots & \ddots & \vdots \\ w_{m1} & \dots & w_{mk} \end{bmatrix} \right) + [b_{11} \dots b_{1k}] \quad (9-7)$$

where k is the hidden neuron number in the first hidden layer, m is the input neuron number, x is the input data, w is the weight, and b is the bias. Similarly, the output of the second hidden layer is:

$$[p_{21} \dots p_{2l}] = f \left([x_{11} \dots x_{1k}] \begin{bmatrix} u_{11} & \dots & u_{1l} \\ \vdots & \ddots & \vdots \\ u_{k1} & \dots & u_{kl} \end{bmatrix} \right) + [b_{21} \dots b_{2l}] \quad (9-8)$$

where l is the hidden neuron number in the second hidden layer. The output of the MLP is:

$$[p_1 \dots p_n] = f \left([x_{21} \dots x_{2l}] \begin{bmatrix} v_{11} & \dots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{l1} & \dots & v_{ln} \end{bmatrix} \right) + [b_{31} \dots b_{3n}] \quad (9-9)$$

At the end of the feedforward propagation process, the loss is calculated by output and the expected value, **Eq. (9-10)** shows the loss function:

$$L(w, b) = \frac{1}{2s} \sum_{i=1}^s \|p_{w,b}(x^{(i)}) - y^{(i)}\|^2 \quad (9-10)$$

where s is the number of the input data, y is the expected output data, namely, the experimental flow stress values. For a single input datum, there is:

$$L(w, b) = \frac{1}{2} \|y - p\|^2 \quad (9-11)$$

In the backpropagation process, the loss error information propagates sequentially along the output, hidden, and input layers. In order to renew weights, the weights' effect on the loss function should be evaluated first. Take the weight and bias between the first neuron in the second hidden layer and the first neuron in the output neuron as an example:

$$\begin{cases} \frac{\partial L}{\partial v_{11}} = \frac{\partial L}{\partial p} \times \frac{\partial p}{\partial c} \times \frac{\partial c}{\partial v_{11}} = -(y - p)f'(p)p_{21} \\ \frac{\partial L}{\partial b_{31}} = \frac{\partial L}{\partial p} \times \frac{\partial p}{\partial c} \times \frac{\partial c}{\partial b_{31}} = -(y - p)f'(p) \end{cases} \quad (9-12)$$

where c is the add value before the activation function operation. Then the weight and bias can be updated:

$$\begin{aligned} v'_{11} &= v_{11} - \eta(y - p)f'(p)p_{21} \\ b'_{31} &= v_{11} - \eta(y - p)f'(p) \end{aligned} \quad (9-13)$$

The updated weights of the output neurons are:

$$\begin{aligned} \begin{bmatrix} v'_{11} & \dots & v'_{1n} \\ \vdots & \ddots & \vdots \\ v'_{l1} & \dots & v'_{ln} \end{bmatrix} &= \begin{bmatrix} v_{11} & \dots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{l1} & \dots & v_{ln} \end{bmatrix} - \eta(y - p)f'(p) \begin{bmatrix} p_{21} & \dots & p_{21} \\ \vdots & \ddots & \vdots \\ p_{2l} & \dots & p_{2l} \end{bmatrix} \\ [b'_{31} & \dots & b'_{3n}] &= [b_{31} & \dots & b_{3n}] - \eta(y - p)f'(p) \end{aligned} \quad (9-14)$$

Take the weight between the first neuron in the first hidden layer and the first neuron in the second hidden layer an example to calculate the renewed weight:

$$\frac{\partial L}{\partial u_{11}} = \frac{\partial L}{\partial p_{21}} \times \frac{\partial p_{21}}{\partial c_{21}} \times \frac{\partial c_{21}}{\partial u_{11}} \quad (9-15)$$

the first part $\frac{\partial L}{\partial p_{21}}$ can be calculated:

$$\begin{cases} \frac{\partial L}{\partial p_{21}} = \sum_{i=1}^n \frac{\partial L_i}{\partial p_{21}} = \sum_{i=1}^n \frac{\partial L_i}{\partial c_i} \times \frac{\partial c_i}{\partial p_{21}} \\ \frac{\partial L_i}{\partial c_i} = \frac{\partial L_i}{\partial p_i} \times \frac{\partial p_i}{\partial c_i} \\ \frac{\partial c_i}{\partial p_{21}} = v_{1i} \end{cases} \quad (9-16)$$

the second part $\frac{\partial p_{21}}{\partial c_{21}}$:

$$\frac{\partial p_{21}}{\partial c_{21}} = f'(p_{21}) \quad (9-17)$$

the third part $\frac{\partial c_{21}}{\partial u_{11}}$:

$$\frac{\partial c_{21}}{\partial u_{11}} = p_{11} \quad (9-18)$$

Combining Eq. (9-15), Eq. (9-16), and Eq. (9-17):

$$\frac{\partial L}{\partial u_{11}} = \left(\sum_{i=1}^n \frac{\partial L_i}{\partial p_i} \times \frac{\partial p_i}{\partial c_i} \times v_{1i} \right) \times f'(p_{21}) \times p_{11} \quad (9-19)$$

Consequently, the weight and bias can be updated:

$$\begin{aligned} u'_{11} &= u_{11} - \eta \left[\left(\sum_{i=1}^n \frac{\partial L_i}{\partial p_i} \times \frac{\partial p_i}{\partial c_i} \times v_{1i} \right) \times f'(p_{21}) \times p_{11} \right] \\ b'_{21} &= b_{21} - \eta \left[\left(\sum_{i=1}^n \frac{\partial L_i}{\partial p_i} \times \frac{\partial p_i}{\partial c_i} \times v_{1i} \right) \times f'(p_{21}) \right] \end{aligned} \quad (9-20)$$

For the same reason, the weight w_{11} can be renewed as well:

$$\begin{aligned} w'_{11} &= w_{11} - \eta \left[\left(\sum_{i=1}^n \sum_{j=1}^l \frac{\partial L_{ij}}{\partial p_{ij}} \times \frac{\partial p_{ij}}{\partial c_{ij}} \times \frac{\partial c_{ij}}{\partial p_{1j}} \times \frac{\partial p_{1j}}{\partial c_{1j}} u_{1j} \right) \times f'(p_{11}) \times x_1 \right] \\ b'_1 &= w_{11} - \eta \left[\left(\sum_{i=1}^n \sum_{j=1}^l \frac{\partial L_{ij}}{\partial p_{ij}} \times \frac{\partial p_{ij}}{\partial c_{ij}} \times \frac{\partial c_{ij}}{\partial p_{1j}} \times \frac{\partial p_{1j}}{\partial c_{1j}} u_{1j} \right) \times f'(p_{11}) \right] \end{aligned} \quad (9-21)$$

After finishing an update of weights and biases, the backpropagation MLP finishes one iteration. The process will continue till the loss meet the set value or the iteration meets the maximum number. The operation process flow chart of the backpropagation MLP is illustrated in **Fig.9. 2**.

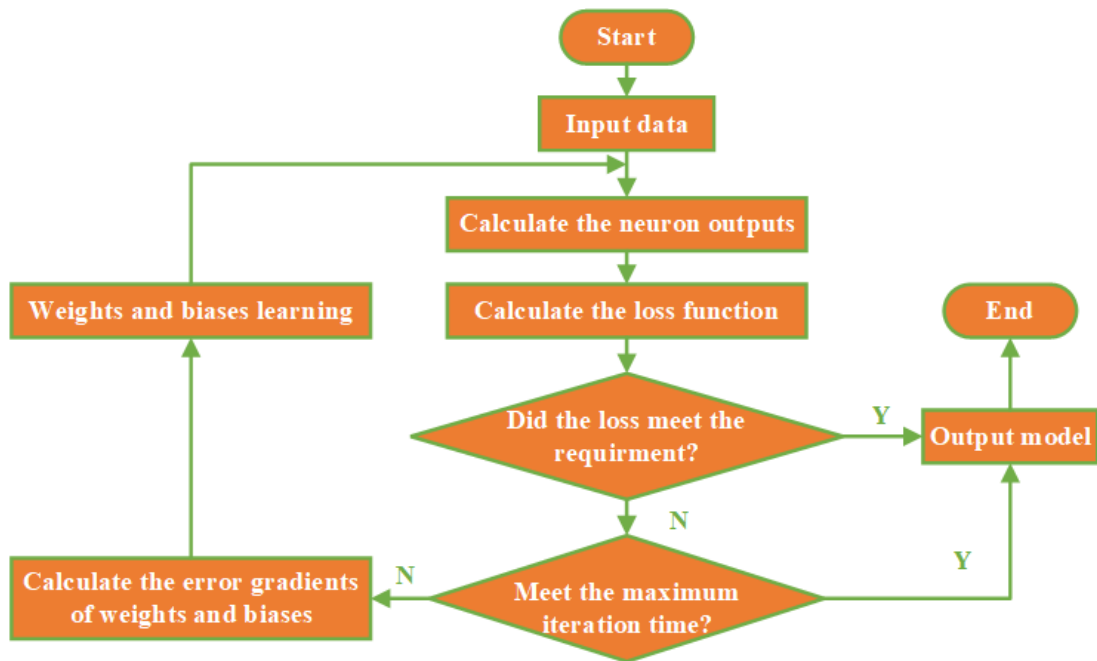


Fig.9. 3. Operation mechanism flow chart of back propagation MLP.

9.2 Regression verification

Due to flow behavior modeling being a four-dimensional nonlinear problem, a one-hidden layer perceptron is able to handle it. In this work, Bayesian regularization was adopted in the training process. In order to determine the hidden layer number, the trial-and-error method was adopted, and the MSE and R value were hired as indicators. The variation of model accuracy with the number of hidden layer neurons is shown in the following figure.

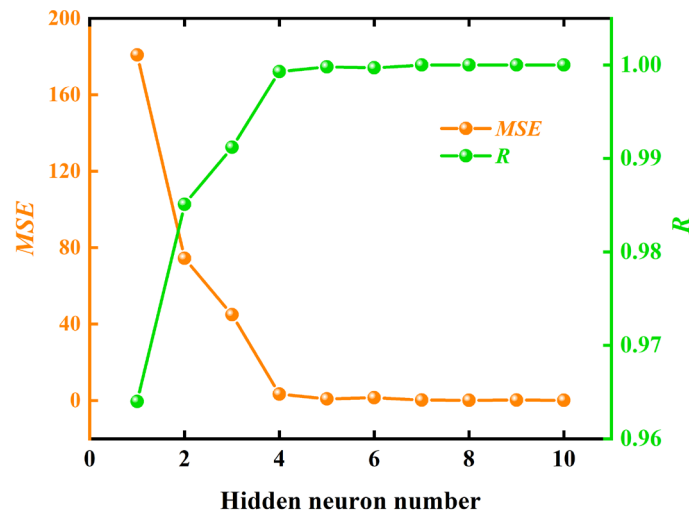


Fig.9. 4. The MSE and R variation along with the hidden layer neuron number.

It is obvious in **Fig.9. 4** that the MSE and R keep stable after the neuron number reaches 5. Therefore, the hidden neuron number is determined as 5. The maximum iteration number was set as 1000. The error curve drops sharply in the initial stage, then becomes slight after 100 iterations. After 400 iterations, the error change becomes stable until the algorithm found the best training result after 798 iterations, with the MSE 0.82284 and R 0.9998, see **Fig.9. 5** and **Fig.9. 6**.

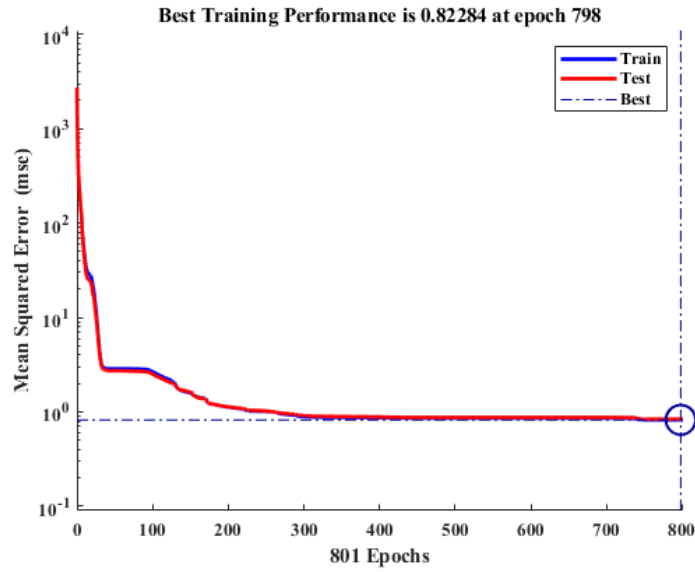


Fig.9. 5. MSE variation during the optimization process.

The Fig.9. 5 shows the fitting performances of the training and test data groups.

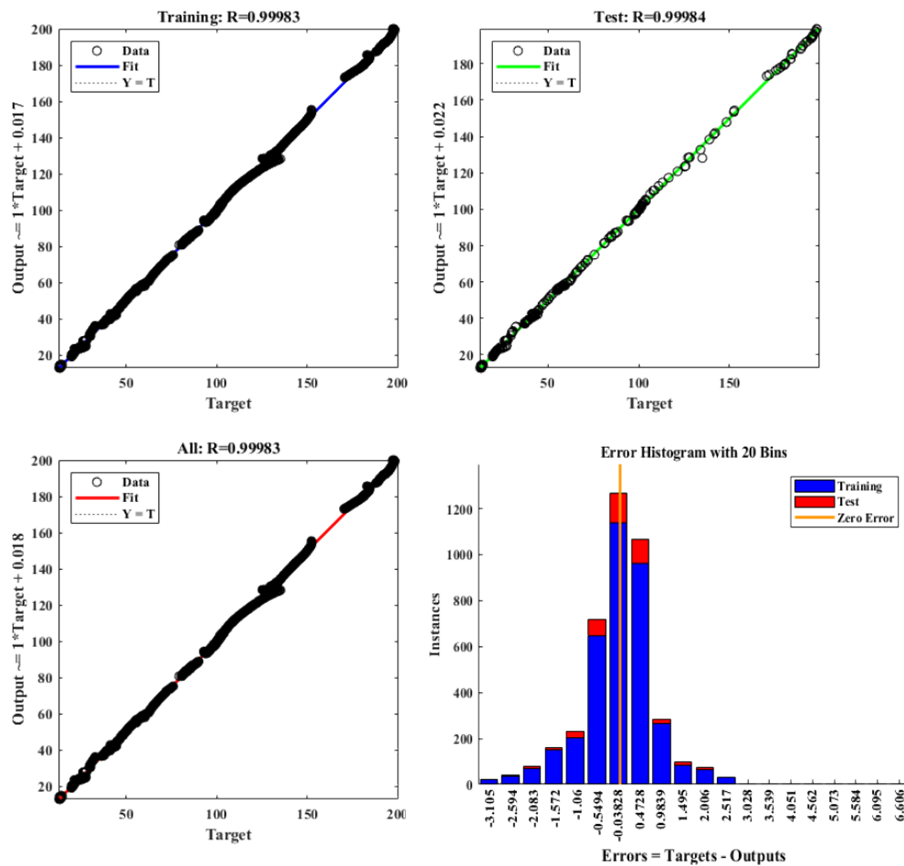


Fig.9. 6. The fitting performances of the training and test data groups.

As can be checked in **Fig.9. 6**, all R values are above 99.98%, which indicates the predicted data keep excellent correlation relationships with the expected data. In addition to this, the vast majority of errors are concentrated around 0, which reveals the high accuracy of the established back propagation MLP model. In sum, the established model is fully able to describe the flow behavior of experimental Al-Mg-Si alloy.

Finally, the weights and biases were obtained. For the hidden layer:

$$\begin{cases} w = \begin{bmatrix} -9.886 & -0.011 & 0.026 & -8.362 & -17.965 \\ -0.269 & -1.130 & -0.358 & -0.249 & -0.987 \\ 0.171 & -0.037 & 0.011 & 0.165 & 0.082 \end{bmatrix} \\ b = [-10.852 \quad 0.443 \quad -0.424 \quad -10.001 \quad -21.141] \end{cases} \quad (9-23)$$

For the output layer:

$$\begin{cases} w = [-17.784 \quad -1.052 \quad 5.469 \quad 60.066 \quad -26.371] \\ b = 18.045 \end{cases} \quad (9-24)$$

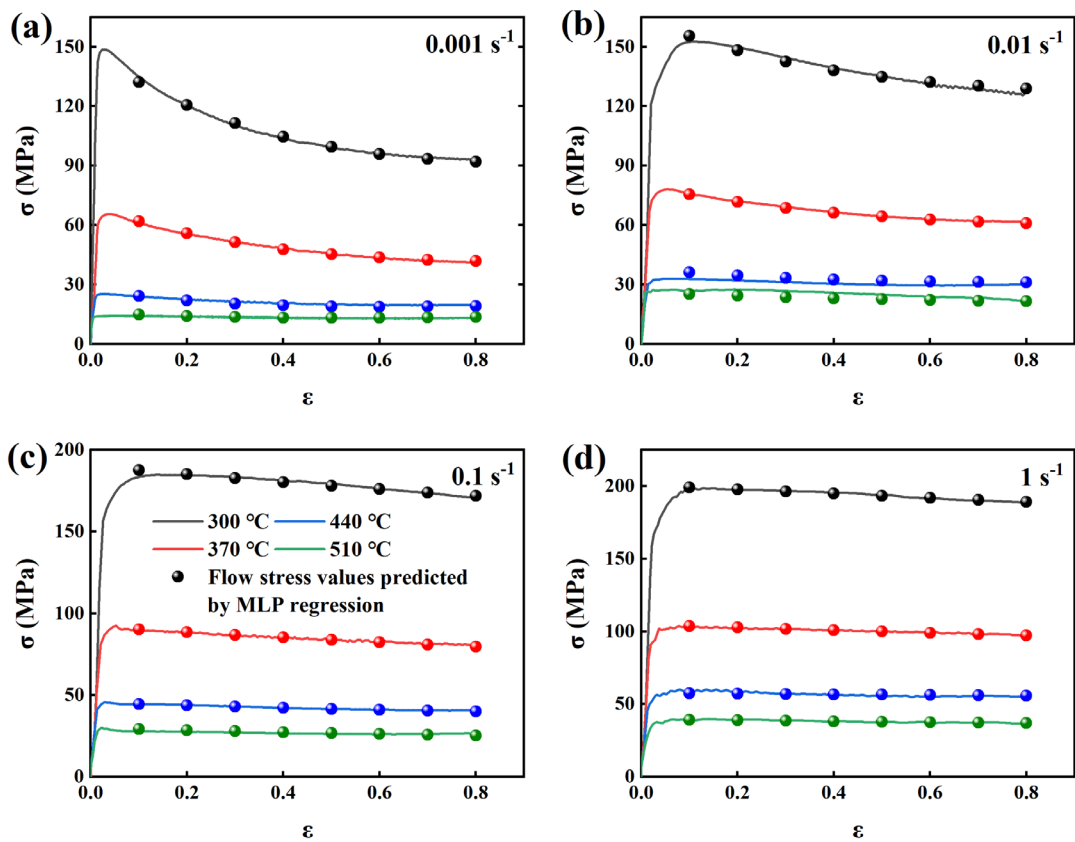


Fig.9. 7. Regression verification of the MLP model at different experimental deformation

conditions, divided by strain: (a) 0.001 s^{-1} , (b) 0.01 s^{-1} , (c) 0.1 s^{-1} , and (d) 1 s^{-1} .

In order to make a comparison, the predicted flow stress is scattered on the experimental flow stress curves. **Fig.9. 6** illustrates that all predicted points locate well at the curves. The trend of flow stress also maintains a high degree of consistency under all designed deformation conditions. In order to guarantee the computing time of simulation works, we controlled the structure of the MLP model as simply as possible. The accuracy of the MLP model will improve as the model becomes more complex. Compared with the RT and RF models, the structure of the MLP model is relatively concise.

10 DISCUSSION

10.1 Performances evaluation of models by R

R was proposed by Karl Pearson and was designed to describe the linear correlation closeness between the studied variables. The Pearson correlation coefficient is the most commonly used one adopted in this work. The R is calculated using the product difference method, which is also based on the deviation between two variables and their respective mean values. The degree of correlation between the two variables is reflected by multiplying the two deviations:

$$R = \frac{Cov(Y, P)}{\sqrt{Var|Y| Var|P|}} \quad (10-1)$$

where $Cov(Y, P)$ is the covariance of Y and P , Y is the set of the experimental flow stress data, P is the set of the predicted flow stress data, $Var|Y|$ and $Var|P|$ are variances of Y and P .

Correlation plots are generally used to reflect the correlation direction and relationship between two variables. The correlation plots about predicted stress values and experimental values are presented in **Fig.10. 1**. The flow stress points were selected from strain 0.1 to 0.8 at an interval of 0.1. To make a clear display, the confidence bands, confidence ellipses, and prediction bands are added, with a degree of confidence of 95%. In addition to this, the perfect fit lines (namely, all predicted stress values are consistent with the experimental stress values) are drawn.

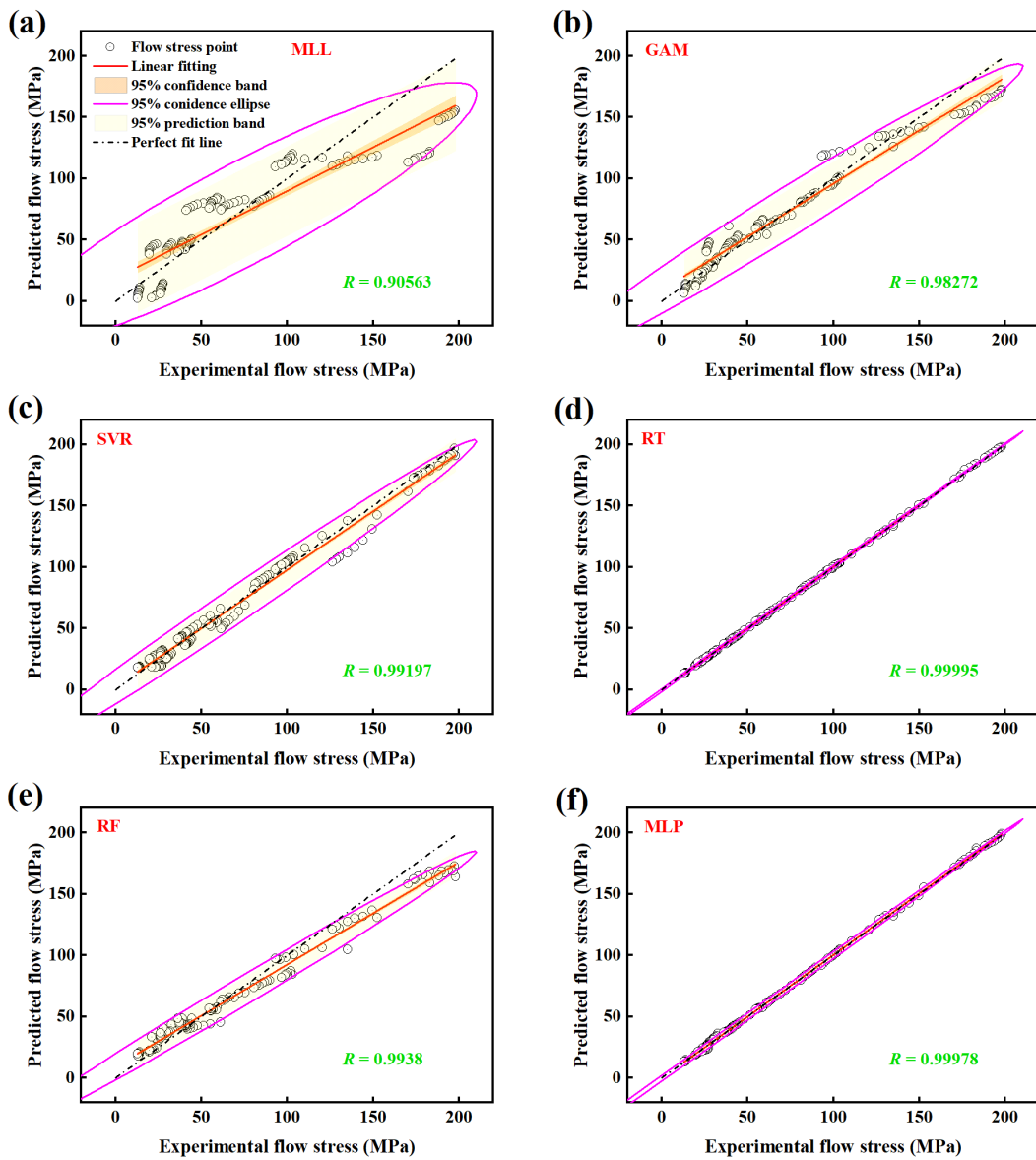


Fig.10. 1. Correlation plots of established models: (a) MLL, (b) GAM, (c) SVR, (d) RT, (e) RF, (f) MLP.

As can be seen in **Fig.10. 1(a)**, the MLL model presents the worst correlation. To be specified, the confidence band and confidence ellipse are the most prominent among all models, revealing the most significant error. Moreover, the largest angle exists between the linear fitting and the perfect match lines. The areas of confidence bands, confidence ellipses, and prediction bands gradually decrease in the order of MLL, GAM, SVR, RF, MLP, and RT,

and so do the angles between linear fitting lines and perfect fit lines. On the contrary, the R gradually increases and approaches the critical correlation value of 1, indicating the related performance enhancement. It is clear that the RT model and the MLP model see the most excellent correlation, with R values of 0.99998 and 0.99978.

However, the determination of error by R heavily depends on the number of evaluation samples. For example, when the number of sample groups is 1, the absolute value of the correlation coefficient is always 1. Consequently, more evaluation methods are needed.

10.2 Performances evaluation of models by RE

RE refers to the difference between the actual observation value and the estimated value in mathematical statistics. If the regression model is correct, the RE can be regarded as the observation value of the error. Residual analysis is often used to examine the reliability of data and the rationality of model assumptions. The mathematic form of RE is quite simple in the present work:

$$RE = p_i - y_i \quad (10-2)$$

i is the sample number of the flow stress.

The residual plots can help determine whether the fitted linear model meets the relevant requirements, see **Fig.10. 2**. The closer the residual value approaches the horizontal axis of 0, the smaller the difference. The RE range of the MLL model is -30~30 MPa, accounting for the most extensive error range, and followed by the GAM model, with the RE range of -

20~30 MPa. At moderate stress, the RE s are the largest, indicating the poor performance of the model in this stress range.

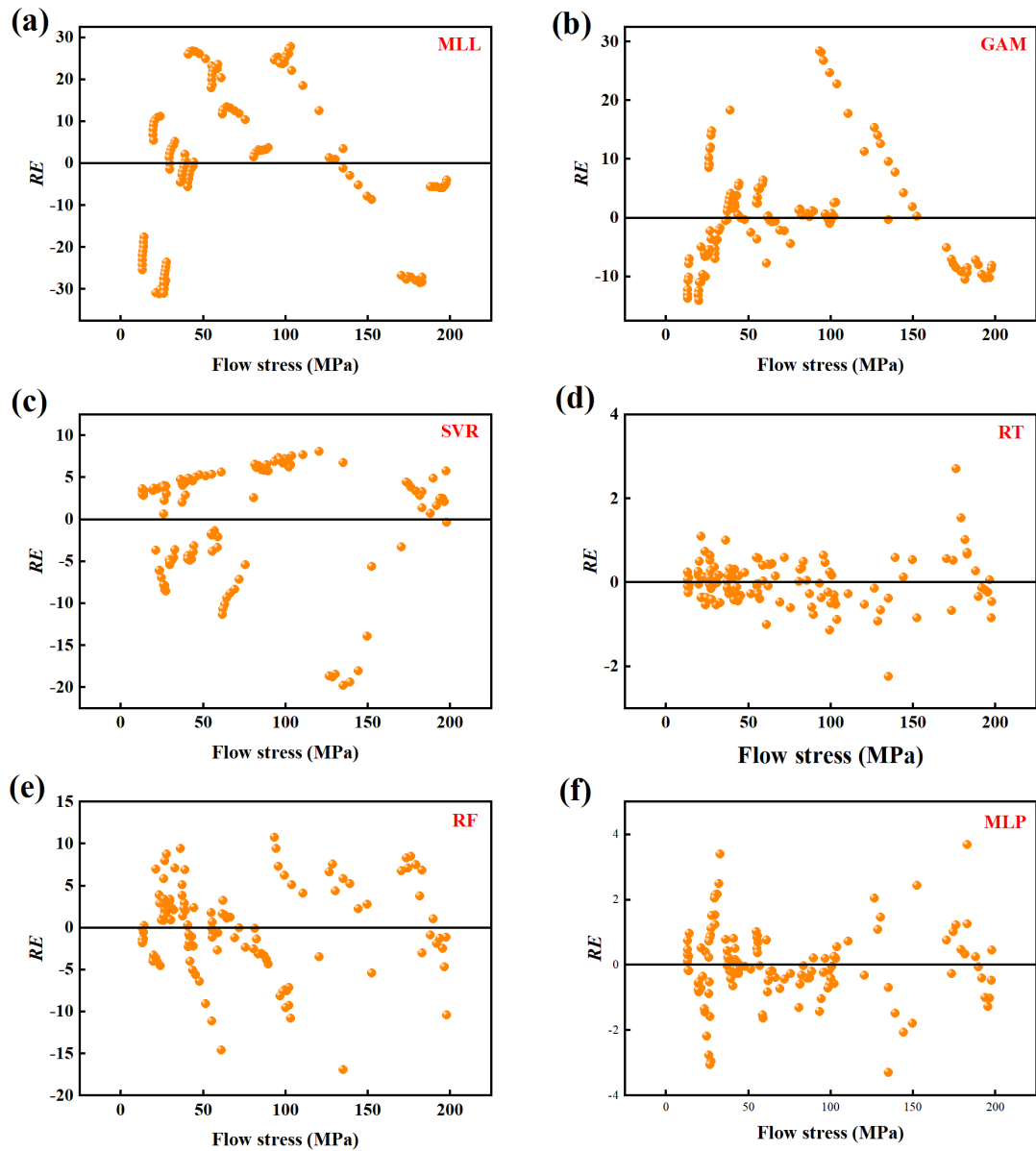


Fig.10. 2. Residual error plots of established models: (a) MLL, (b) GAM, (c) SVR, (d) RT, (e) RF, (f) MLP.

The RE ranges of the SVR model and the RF model are similar (-20~10 MPa), while the SVR model shows large errors in the stress range of 100~150 MPa. As for the RF model, the error size is evenly distributed among various stress values.

The *RE* values of the RT model and MLP model have shrunk to the single-digit range, demonstrating extreme error accuracy. Compared with the MLP model, the error of the RT model is more compact, which means that the model has the highest accuracy.

10.3 Performances evaluation of models by δ distribution

The relative error refers to the ratio percentage value of measured absolute error and the true value. Generally speaking, the δ is able to reflect the credibility of measurements. It can be presented by the ratio of the absolute error and the real value:

$$\delta = \frac{y_i - p_i}{y_i} \times 100\% \quad (10-3)$$

The relative error not only indicates the size of the error but also points out the directions of positive or negative. By reflecting the predicted deviation values with the same unit dimension, the relative error represents the actual size of the true value deviation accurately. At the same time, the Gaussian distribution function is utilized to verify the uniformity of error distribution:

$$\begin{cases} y = y_0 + Ae^{-\frac{(\delta-\mu)^2}{2\omega^2}} \\ \omega = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\delta - \mu)^2} \\ \mu = \frac{1}{N} \sum_{i=1}^N \delta_i \end{cases} \quad (10-4)$$

where both y_0 and A are constants, ω is the standard deviation of δ , μ is the average value of δ , y is the probability of the δ . In general, the closer ω and μ to 0, the higher precise the model is. The δ distribution plots are presented in **Fig.10. 3** for further evaluation.

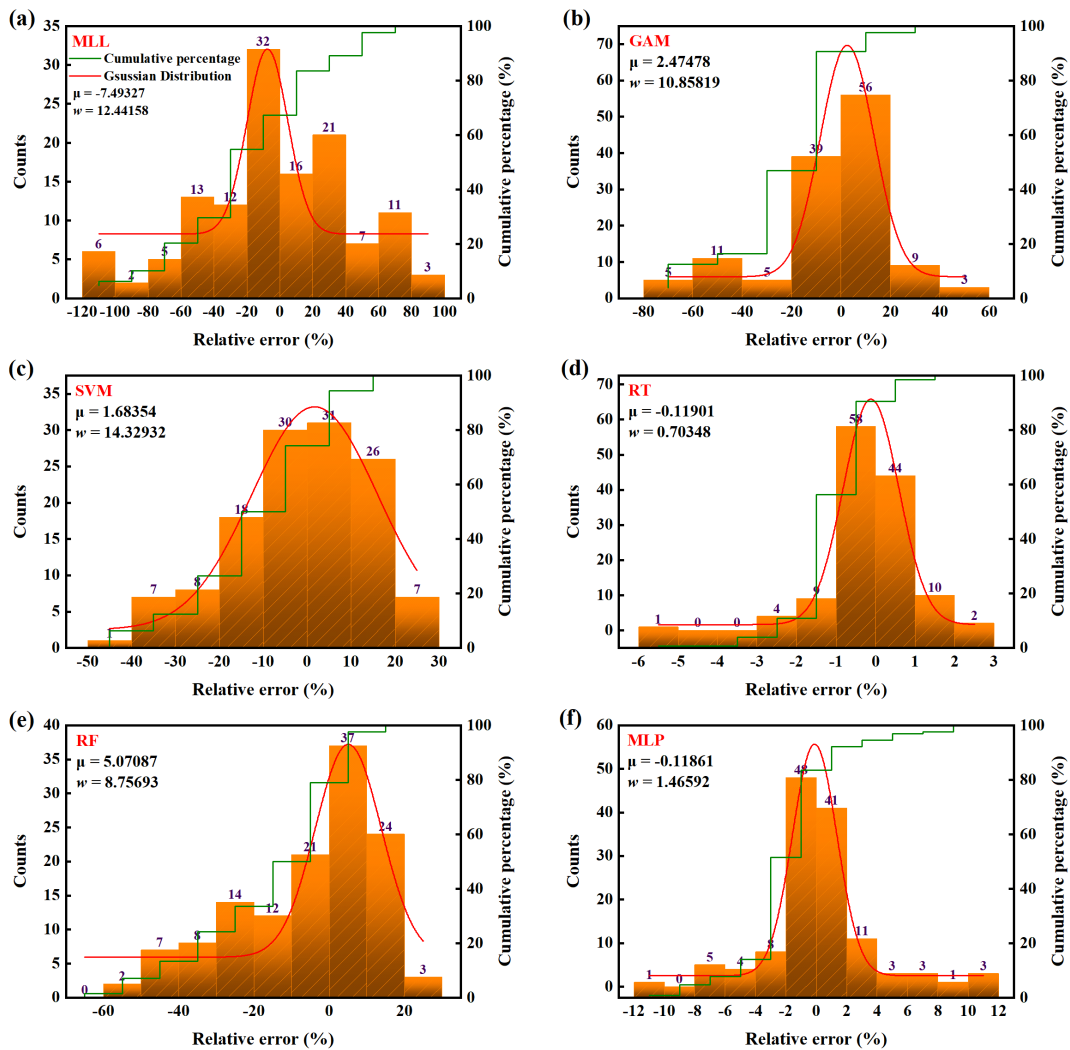


Fig.10. 3. δ distribution plots of established models: (a) MLL, (b) GAM, (c) SVR, (d) RT, (e) RF, (f) MLP.

In **Fig.10. 3(a)**, the δ distribution has good normal distribution characteristics, whereas the δ s are relatively high (some errors even out of 100%). Additionally, the MLL model shows the biggest ω (12.44158) and μ (-7.49327), which explains the insufficient accuracy of the model. Compared with the MLL model, more relative errors of the GAM, SVM, and RF models concentrate at the low error range. Although the SVM model provides the lowest μ (14.32932) among the mentioned three models, it does not mean this model owns the highest

precision. This is because the low μ credits to the offset of positive and negative errors. However, the low μ implies the uniform error distribution symmetry. Due to the introduction of square error, the ω does not subject to the influence of positive and negative values. The RF model shows the least ω among three models. With respect to the RT model and MLP model, the RT model has the narrowest δ distribution range (from -6 % to 3 %), but the MLP model offers better δ distribution. Therefore, the μ value of the RT model is higher than that of the MLP model. On the opposite, the ω value of the RT model is lower than that of the MLP model.

10.4 Performances evaluation of models by *AARE*

AARE refers to the average absolute values of deviations of all single observation values and arithmetic mean. Different from the relative error, the problem of mutual cancellation of errors can be avoided by *AARE*. Consequently, the *AARE* can better reflect the actual situation of the prediction error. The following is the mathematic form of *AARE*:

$$AARE = \frac{\sum_{i=1}^N \left| \frac{y_i - p_i}{y_i} \right|}{N} \quad (10-5)$$

All predicted and experimental flow stress values were imported into the above equation, and the *AAREs* for each model are listed in **Table 10-1**.

Table 10-1 *AAREs* of different established models.

	MLL	GAM	SVR	RT	RF	MLP
AARE	0.361758	0.165236	0.12158	0.00708	0.149465	0.019797

It can be seen that the RT model exhibits the smallest *AARE* (0.00708), namely, the highest model accuracy. Next is the MLP model, with a value of 0.19797. The errors of GAM, SVR, and RF models are generally close, with corresponding *AARE* values of 0.165236, 0.12158, and 0.149465, respectively. Besides, the MLL model performs worst, with a value of 0.361758. In order to make a detailed study, the *AARE*s in deformation temperatures and different process strain rates were summarized, and the *AARE* histograms are displayed in Fig.10. 4.

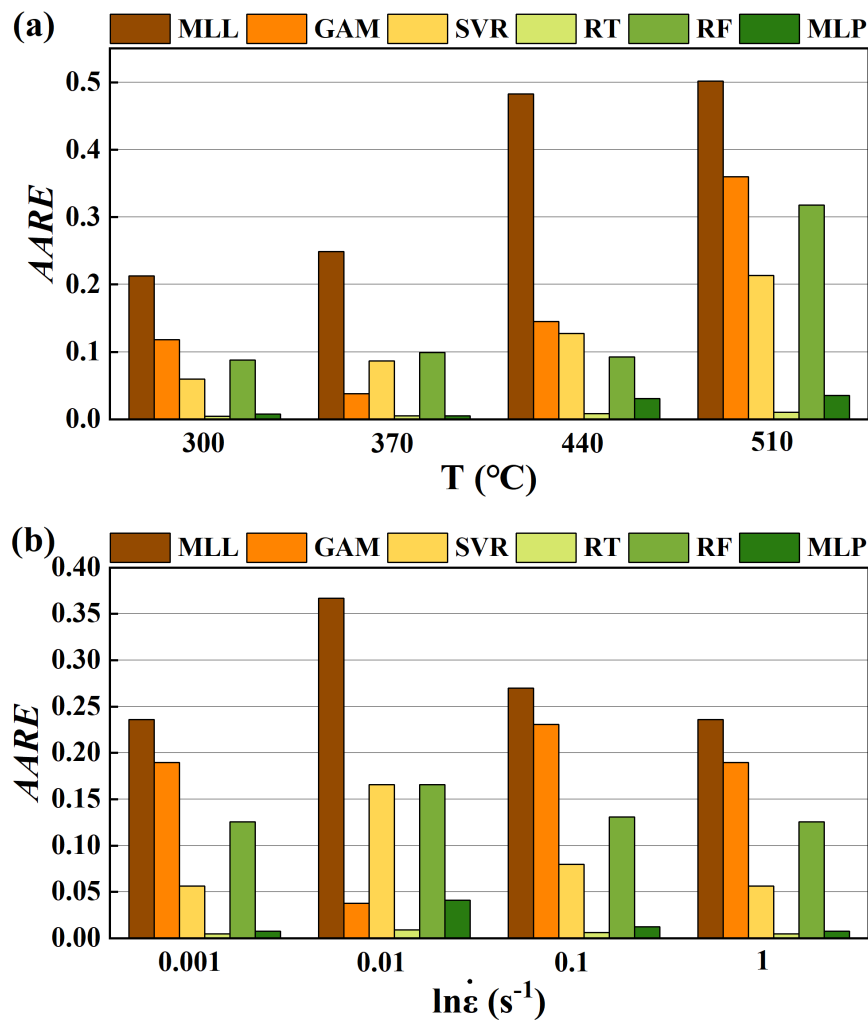


Fig.10. 4. *AARE*s of established models (a) in different temperatures, and (b) in different strain rates.

From **Fig.10. 4(a)**, we can see that the MLL model shows the highest *AAREs* in all temperatures, and the errors increase with the increase of temperature (from 0.21 to 0.5). The SVR, RT, and MLP models show similar increase trends (0.05 to 0.21, 0.005 to 0.01, and 0.01 to 0.05, respectively), but the *AAREs* are much lower than the MLL model. From 300 °C to 370 °C, the GAM sees a downtrend from 0.12 to 0.04 in *AARE*. After that, the *AARE* rises to 0.38 till the temperature 510 °C. The *AARE* variation of the RF model performs a difference, and it keeps a nearly constant (nearly 0.1) from 300 to 440 °C, while it soars to 0.32 at 510 °C.

As for **Fig.10. 4(b)**, the MLL model still contributes the largest *AAREs*, but the difference is less obvious than **Fig.10. 4(a)**. The *AARE* (0.37) of the MLL model occurs at the strain rate of 0.01 s⁻¹. The GAM keeps high *AAREs* under all strain rates except 0.01 s⁻¹, with an *AARE* of 0.035. The *AAREs* of the SVR, RF, RT, and MLP models show a convex distribution with increasing strain rate, with the maximum values (0.17, 0.17, 0.01, and 0.04, respectively) also appearing at 0.01 s⁻¹ strain rate. The above phenomenon indicates that most models lose accuracy at strain rate 0.01 s⁻¹. It is worth noting that the RT and MLP models still keep low errors under these conditions.

10.5 Summary of the chapter

In this chapter, the accuracy performances of the established models are examined by statistical strategies such as correlation coefficient, residual error, relative error distribution,

and average absolute relative error. The comparison results show that the RT model gives the strongest correlation, shows the narrowest residual and relative error ranges, and provides the lowest *AARE*. All evaluation results point out the fact that the RT model owns the highest accuracy. The next is the MLP model, which presents excellent regression ability as well. Additionally, the MLP model has a more uniform error distribution than the RT model. Although the GAM model, the SVR model, and the RF model are able to take the regression work, errors still exist when compared with the MLP model and the RT model. Regardless of which statistical analysis method is used as the background, the MLL model performs the worst. In other words, the MLL model fails to describe the flow behavior of the AA6061-T6 alloy at the experimental conditions. Eventually, the MLP and RT models are recommended as the flow stress indicator of the studied alloy.

11 CONCLUSION

Three parts were divided in this work to study the deformation mechanism and the performance of the AL-Mg-Si AA6061-T6 alloy. In the first part, the Gleeble-3800 thermal simulator was adopted to carry out the thermal axial compression tests. Sixteen flow behavior curves were obtained. The deformation mechanism dominated by DRV and CDRX is discussed by combining the flow stress characteristics. Based on the DMM theory, the thermal processing maps are plotted in the second part. In the third part, six different ML regression models are utilized to regress the flow behavior of the AA6061-T6 alloy at designed deformation conditions. Statistical analysis methods evaluate the description performances. Through the analysis and comparisons of the experiment data, the following conclusions are received:

- (1) The flow behavior of the AA6061-T6 alloy is the result of the combined effect of WH and DS. WH is a manifestation of dislocation multiplication and accumulation. DS is caused by DRV and DRX. In this work, CDRX is a typical example of softening;
- (2) CDRX can be divided into two situations: one is a continuous decrease after the stress reaches its peak, during which CDRX occurs continuously; Another type is the gradual storage of potential energy in materials at low strain rates, with the intermittent occurrence of CDRX, mainly manifested by a wavy flow curve;
- (3) On the basis of DMM and Ziegler's theory, the power dissipation is studied, and the thermal processing maps along different deformation conditions are established. The

desirable thermal processing conditions are determined as 300 °C and 0.001 s⁻¹, 300 °C and 1 s⁻¹, 440~510 °C and 0.1~1 s⁻¹;

- (4) The operation mechanisms of ML algorithms (MLL, GAM, SVR, RT, RF, and MLP) are introduced and induced in this work. Aims at describing the flow behavior of the studied alloy, six regression models were built;
- (5) Statistical methods (*R*, *RE*, relative error distribution, and *AARE*) were adopted to evaluate the regression performances of the ML models. RT possesses the best performance (*R* 0.99995, *AARE* 0.00708), and the MLP model also shows excellent performance (*R* 0.99978, *AARE* 0.019797). The precision of the MLL model is the worst (*R* 0.90563, *AARE* 0.361758);
- (6) In practical applications, models can be flexibly selected according to different needs. If high accuracy is required, RT models can be selected; If a high-precision and a structurally simple model is needed, the MLP model can be considered. In addition, the MLP model has specific equation expressions, but the operation cannot be explained.

REFERENCES

- [1] X. Zhu, H. Yang, X. Dong, and S. Ji, "The effects of varying Mg and Si levels on the microstructural inhomogeneity and eutectic Mg₂Si morphology in die-cast Al–Mg–Si alloys," *J. Mater. Sci.*, vol. 54, no. 7, pp. 5773–5787, Apr. 2019, doi: 10.1007/s10853-018-03198-6.
- [2] S. Zhu, H.-C. Shih, X. Cui, C.-Y. Yu, and S. P. Ringer, "Design of solute clustering during thermomechanical processing of AA6016 Al–Mg–Si alloy," *Acta Mater.*, vol. 203, p. 116455, Jan. 2021, doi: 10.1016/j.actamat.2020.10.074.
- [3] J. R. Kissell and R. L. Ferry, *Aluminum Structures: A Guide to Their Specifications and Design*. John Wiley & Sons, 2002.
- [4] G. Lu, B. Sun, J. Wang, Y. Liu, and C. Liu, "Enhanced high-temperature age-hardening behavior and mechanical properties of Al–Mg–Si alloys via microalloying with Cd," *Mater. Sci. Eng. A*, vol. 825, p. 141910, Sep. 2021, doi: 10.1016/j.msea.2021.141910.
- [5] M. Ramamurthy, P. Balasubramanian, N. Senthilkumar, and G. Anbuhezhiyan, "Influence of process parameters on the microstructure and mechanical properties of friction stir welds of AA2014 and AA6063 aluminium alloys using response surface methodology," *Mater. Res. Express*, vol. 9, no. 2, p. 026528, Feb. 2022, doi: 10.1088/2053-1591/ac5777.
- [6] B. Valdez, S. Kiyota, M. Stoytcheva, R. Zlatev, and J. M. Bastidas, "Cerium-based conversion coatings to improve the corrosion resistance of aluminium alloy 6061-T6," *Corros. Sci.*, vol. 87, pp. 141–149, Oct. 2014, doi: 10.1016/j.corsci.2014.06.023.
- [7] J. Zhang *et al.*, "Reinforcement study of anodizing treatment with various temperatures on aluminum substrates for stronger adhesive bonding with carbon fiber composites," *Surf. Coat. Technol.*, vol. 462, p. 129473, Jun. 2023, doi: 10.1016/j.surfcoat.2023.129473.
- [8] N. R. J. Hynes *et al.*, "Effect of stacking sequence of fibre metal laminates with carbon fibre reinforced composites on mechanical attributes: Numerical simulations and experimental validation," *Compos. Sci. Technol.*, vol. 221, p. 109303, Apr. 2022, doi: 10.1016/j.compscitech.2022.109303.
- [9] W. Chen, S. Li, S. B. Krishna, N. Kosimov, and D. Jung, "Comparison of Shell Thickness Integration Rules and Element Types on Single Point Incremental Forming (SPIF) Simulation," *J. Adv. Technol. Eng. Res.*, vol. 7, no. 1, pp. 17–24, 2021, doi: <https://doi.org/10.20474/jater-7.1.3>.
- [10] S. Li, W. Chen, N. Kosimov, S. B. Krishna, and D. Jung, "Reduction of Explicit Incremental Forming Simulation Time by Mass Scaling Parameters Adjusting and Symmetrical Setting," *Int. J. Technol. Eng. Stud.*, vol. 7, no. 1, pp. 27–34, doi: <https://dx.doi.org/10.20469/ijtes.7.10004-1>.
- [11] X. Chen *et al.*, "Identification of the Constitutive Model Parameters by Inverse Optimization Method and Characterization of Hot Deformation Behavior for Ultra-Supercritical Rotor Steel," *Materials*, vol. 14, no. 8, Art. no. 8, Jan. 2021, doi: 10.3390/ma14081958.
- [12] X. Chen *et al.*, "An enhanced lemaître damage model and verification method for X12 alloy steel in hot forming process," *Eng. Fract. Mech.*, vol. 273, p. 108711, Oct. 2022, doi:

- 10.1016/j.engfracmech.2022.108711.
- [13] A. Rudnytskyj, P. Simon, M. Jech, and C. Gachot, “Constitutive modelling of the 6061 aluminium alloy under hot rolling conditions and large strain ranges,” *Mater. Des.*, vol. 190, p. 108568, May 2020, doi: 10.1016/j.matdes.2020.108568.
- [14] G. Geng, D. Ding, L. Duan, and H. Jiang, “A modified Johnson-Cook model of 6061-T6 Aluminium profile,” *Aust. J. Mech. Eng.*, vol. 20, no. 2, pp. 516–526, Mar. 2022, doi: 10.1080/14484846.2020.1721966.
- [15] X. H. Fan, M. Li, D. Y. Li, Y. C. Shao, S. R. Zhang, and Y. H. Peng, “Dynamic recrystallisation and dynamic precipitation in AA6061 aluminium alloy during hot deformation,” *Mater. Sci. Technol.*, vol. 30, no. 11, pp. 1263–1272, Sep. 2014, doi: 10.1179/1743284714Y.0000000538.
- [16] W. Chen, Y. Guan, and Z. Wang, “Modeling of Flow Stress of High Titanium Content 6061 Aluminum Alloy Under Hot Compression,” *J. Mater. Eng. Perform.*, vol. 25, no. 9, pp. 4081–4088, Sep. 2016, doi: 10.1007/s11665-016-2224-z.
- [17] K. Wang, D. Wen, J. Li, Z. Zheng, and Y. Xiong, “Hot deformation behaviors of low-alloyed ultrahigh strength steel 30CrMnSiNi2A: Microstructure evolution and constitutive modeling,” *Mater. Today Commun.*, vol. 26, p. 102009, Mar. 2021, doi: 10.1016/j.mtcomm.2021.102009.
- [18] Y. C. Lin, H. Yang, D.-G. He, and J. Chen, “A physically-based model considering dislocation–solute atom dynamic interactions for a nickel-based superalloy at intermediate temperatures,” *Mater. Des.*, vol. 183, p. 108122, Dec. 2019, doi: 10.1016/j.matdes.2019.108122.
- [19] D. Wen *et al.*, “Hot tensile behavior of a low-alloyed ultrahigh strength steel: fracture mechanism and physically-based constitutive model,” *J. Mater. Res. Technol.*, vol. 13, pp. 1684–1697, Jul. 2021, doi: 10.1016/j.jmrt.2021.05.100.
- [20] J.-Y. Zhang, P. Jiang, Z. Zhu, Q. Chen, J. Zhou, and Y. Meng, “Tensile properties and strain hardening mechanism of Cr-Mn-Si-Ni alloyed ultra-strength steel at different temperatures and strain rates,” *J. Alloys Compd.*, vol. 842, p. 155856, Nov. 2020, doi: 10.1016/j.jallcom.2020.155856.
- [21] J. Lin and J. Yang, “GA-based multiple objective optimisation for determining viscoplastic constitutive equations for superplastic alloys,” *Int. J. Plast.*, vol. 15, no. 11, pp. 1181–1196, Jan. 1999, doi: 10.1016/S0749-6419(99)00031-5.
- [22] H. R. R. Ashtiani and A. A. Shayanpoor, “New constitutive equation utilizing grain size for modeling of hot deformation behavior of AA1070 aluminum,” *Trans. Nonferrous Met. Soc. China*, vol. 31, no. 2, pp. 345–357, Feb. 2021, doi: 10.1016/S1003-6326(21)65500-0.
- [23] W. Chen, S. Li, K. S. Bhandari, S. Aziz, X. Chen, and D. W. Jung, “Genetic optimized Al–Mg alloy constitutive modeling and activation energy analysis,” *Int. J. Mech. Sci.*, vol. 244, p. 108077, Apr. 2023, doi: 10.1016/j.ijmecsci.2022.108077.
- [24] L. He, H. Su, J. Xu, and L. Zhang, “Inverse identification of constitutive parameters of Ti 2 AlNb intermetallic alloys based on cooperative particle swarm optimization,” *Chin. J. Aeronaut.*, vol. 31, no. 8, pp. 1774–1785, Aug. 2018, doi: 10.1016/j.cja.2018.01.002.
- [25] F. Jiang, L. Fei, H. Jiang, Y. Zhang, Z. Feng, and S. Zhao, “Constitutive model research on the hot deformation behavior of Ti6Al4V alloy under wide temperatures,” *J. Mater. Res.*

- Technol.*, vol. 23, pp. 1062–1074, Mar. 2023, doi: 10.1016/j.jmrt.2023.01.021.
- [26] F. Yin, L. Hua, H. Mao, X. Han, D. Qian, and R. Zhang, “Microstructural modeling and simulation for GCr15 steel during elevated temperature deformation,” *Mater. Des.*, vol. 55, pp. 560–573, Mar. 2014, doi: 10.1016/j.matdes.2013.10.042.
- [27] S. Li, W. Chen, K. S. Bhandari, D. W. Jung, and X. Chen, “Flow Behavior of AA5005 Alloy at High Temperature and Low Strain Rate Based on Arrhenius-Type Equation and Back Propagation Artificial Neural Network (BP-ANN) Model,” *Materials*, vol. 15, no. 11, p. 3788, May 2022, doi: 10.3390/ma15113788.
- [28] R. Bobbili, V. Madhu, and A. K. Gogia, “Neural network modeling to evaluate the dynamic flow stress of high strength armor steels under high strain rate compression,” *Def. Technol.*, vol. 10, no. 4, pp. 334–342, Dec. 2014, doi: 10.1016/j.dt.2014.06.012.
- [29] G.-Z. Quan, Z. Zhang, J. Pan, and Y. Xia, “Modelling the Hot Flow Behaviors of AZ80 Alloy by BP-ANN and the Applications in Accuracy Improvement of Computations,” *Mater. Res.*, vol. 18, no. 6, pp. 1331–1345, Nov. 2015, doi: 10.1590/1516-1439.040015.
- [30] “LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S | Philosophical Transactions of the Royal Society of London.” <https://royalsocietypublishing.org/doi/abs/10.1098/rstl.1763.0053> (accessed Apr. 15, 2023).
- [31] R. van de Schoot *et al.*, “Bayesian statistics and modelling,” *Nat. Rev. Methods Primer*, vol. 1, no. 1, Art. no. 1, Jan. 2021, doi: 10.1038/s43586-020-00001-2.
- [32] D. E. Holmes and L. C. Jain, “Introduction to Bayesian Networks,” in *Innovations in Bayesian Networks: Theory and Applications*, D. E. Holmes and L. C. Jain, Eds., in Studies in Computational Intelligence. Berlin, Heidelberg: Springer, 2008, pp. 1–5. doi: 10.1007/978-3-540-85066-3_1.
- [33] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychol. Rev.*, vol. 65, pp. 386–408, 1958, doi: 10.1037/h0042519.
- [34] Š. Raudys, “Evolution and generalization of a single neuron: I. Single-layer perceptron as seven statistical classifiers,” *Neural Netw.*, vol. 11, no. 2, pp. 283–296, Mar. 1998, doi: 10.1016/S0893-6080(97)00135-4.
- [35] R. Kruse, S. Mostaghim, C. Borgelt, C. Braune, and M. Steinbrecher, “Multi-layer Perceptrons,” in *Computational Intelligence: A Methodological Introduction*, R. Kruse, S. Mostaghim, C. Borgelt, C. Braune, and M. Steinbrecher, Eds., in Texts in Computer Science. Cham: Springer International Publishing, 2022, pp. 53–124. doi: 10.1007/978-3-030-42227-1_5.
- [36] J. Li, J. Cheng, J. Shi, and F. Huang, “Brief Introduction of Back Propagation (BP) Neural Network Algorithm and Its Improvement,” in *Advances in Computer Science and Information Engineering*, D. Jin and S. Lin, Eds., in Advances in Intelligent and Soft Computing. Berlin, Heidelberg: Springer, 2012, pp. 553–558. doi: 10.1007/978-3-642-30223-7_87.
- [37] K. Koutroumbas and N. Kalouptsidis, “Generalized hamming networks and applications,” *Neural Netw.*, vol. 18, no. 7, pp. 896–913, Sep. 2005, doi: 10.1016/j.neunet.2005.02.003.
- [38] H. Ramsauer *et al.*, “Hopfield Networks is All You Need.” arXiv, Apr. 28, 2021. doi:

10.48550/arXiv.2008.02217.

- [39] Y. Song, M. Han, and J. Wei, “Stability and Hopf bifurcation analysis on a simplified BAM neural network with delays,” *Phys. Nonlinear Phenom.*, vol. 200, no. 3, pp. 185–204, Jan. 2005, doi: 10.1016/j.physd.2004.10.010.
- [40] G. Li *et al.*, “Understanding error propagation in deep learning neural network (DNN) accelerators and applications,” in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, in SC '17. New York, NY, USA: Association for Computing Machinery, Nov. 2017, pp. 1–12. doi: 10.1145/3126908.3126964.
- [41] J. Gu *et al.*, “Recent advances in convolutional neural networks,” *Pattern Recognit.*, vol. 77, pp. 354–377, May 2018, doi: 10.1016/j.patcog.2017.10.013.
- [42] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, “Recent Advances in Recurrent Neural Networks.” arXiv, Feb. 22, 2018. doi: 10.48550/arXiv.1801.01078.
- [43] J. R. Quinlan, “Induction of decision trees,” *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, Mar. 1986, doi: 10.1007/BF00116251.
- [44] J. R. Quinlan, “Improved Use of Continuous Attributes in C4.5,” *J. Artif. Intell. Res.*, vol. 4, pp. 77–90, Mar. 1996, doi: 10.1613/jair.279.
- [45] L. Breiman, “Bagging predictors,” *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996, doi: 10.1007/BF00058655.
- [46] M. Mehta, R. Agrawal, and J. Rissanen, “SLIQ: A fast scalable classifier for data mining,” in *Advances in Database Technology — EDBT '96*, P. Apers, M. Bouzeghoub, and G. Gardarin, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 1996, pp. 18–32. doi: 10.1007/BFb0014141.
- [47] B.-W. Zhang, F. Min, and D. Ciucci, “Representative-based classification through covering-based neighborhood rough sets,” *Appl. Intell.*, vol. 43, no. 4, pp. 840–854, Dec. 2015, doi: 10.1007/s10489-015-0687-5.
- [48] R. Wang, Y.-L. He, C.-Y. Chow, F.-F. Ou, and J. Zhang, “Learning ELM-Tree from big data based on uncertainty reduction,” *Fuzzy Sets Syst.*, vol. 258, pp. 79–100, Jan. 2015, doi: 10.1016/j.fss.2014.04.028.
- [49] Z. A. Sosnowski and Ł. Gadomer, “Fuzzy trees and forests—Review,” *WIREs Data Min. Knowl. Discov.*, vol. 9, no. 6, p. e1316, 2019, doi: 10.1002/widm.1316.
- [50] S. Suthaharan, “Decision Tree Learning,” in *Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning*, S. Suthaharan, Ed., in Integrated Series in Information Systems. Boston, MA: Springer US, 2016, pp. 237–269. doi: 10.1007/978-1-4899-7641-3_10.
- [51] J. Struyf and S. Džeroski, “Constraint Based Induction of Multi-objective Regression Trees,” in *Knowledge Discovery in Inductive Databases*, F. Bonchi and J.-F. Boulicaut, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2006, pp. 222–233. doi: 10.1007/11733492_13.
- [52] J. Gama, “Functional Trees,” *Mach. Learn.*, vol. 55, no. 3, pp. 219–250, Jun. 2004, doi: 10.1023/B:MACH.0000027782.67192.13.
- [53] T. K. Ho, “Random decision forests,” in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Aug. 1995, pp. 278–282 vol.1. doi:

- 10.1109/ICDAR.1995.598994.
- [54] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
- [55] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.
- [56] John C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," *Adv. Large Margin Classif.*, vol. 10, no. 3, pp. 61–74, 1999.
- [57] M. Allahviranloo and W. Recker, "Daily activity pattern recognition by using support vector machines with multiple classes," *Transp. Res. Part B Methodol.*, vol. 58, pp. 16–43, Dec. 2013, doi: 10.1016/j.trb.2013.09.008.
- [58] J. A. K. Suykens and J. Vandewalle, "Least Squares Support Vector Machine Classifiers," *Neural Process. Lett.*, vol. 9, no. 3, pp. 293–300, Jun. 1999, doi: 10.1023/A:1018628609742.
- [59] S. Koda, A. Zeggada, F. Melgani, and R. Nishii, "Spatial and Structured SVM for Multilabel Image Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5948–5960, Oct. 2018, doi: 10.1109/TGRS.2018.2828862.
- [60] A. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet, "More efficiency in multiple kernel learning," in *Proceedings of the 24th international conference on Machine learning*, in ICML '07. New York, NY, USA: Association for Computing Machinery, Jun. 2007, pp. 775–782. doi: 10.1145/1273496.1273594.
- [61] M. Awad and R. Khanna, "Support Vector Regression," *Effic. Learn. Mach.*, pp. 67–80, 2015, doi: 10.1007/978-1-4302-5990-9_4.
- [62] A. Ben-Hur, D. Horn, H. T. Siegelmann, and V. Vapnik, "Support vector clustering," *J. Mach. Learn. Res.*, vol. 2, no. Dec, pp. 125–137, 2001.
- [63] K. Bennett and A. Demiriz, "Semi-Supervised Support Vector Machines," in *Advances in Neural Information Processing Systems*, MIT Press, 1998. Accessed: Apr. 16, 2023. [Online]. Available: <https://proceedings.neurips.cc/paper/1998/hash/b710915795b9e9c02cf10d6d2bdb688c-Abstract.html>
- [64] K. Guo, Z. Yang, C.-H. Yu, and M. J. Buehler, "Artificial intelligence and machine learning in design of mechanical materials," *Mater. Horiz.*, vol. 8, no. 4, pp. 1153–1172, 2021, doi: 10.1039/D0MH01451F.
- [65] W. N. Chen, S. J. Li, K. S. Bhandari, S. Aziz, N. Kosimov, and D. W. Jung, "A Flow Stress Equation of AA5005 Aluminum Alloy Based on Fields-Backofen Model," *Mater. Sci. Forum*, vol. 1078, pp. 3–10, 2022, doi: 10.4028/p-95f697.
- [66] S. J. Li, W. N. Chen, K. S. Bhandari, N. Kosimov, and D. W. Jung, "Study on Flow Stress Model of AA5005 Material," *Solid State Phenom.*, vol. 335, pp. 107–112, 2022, doi: 10.4028/p-4t00fs.
- [67] W. N. Chen, S. J. Li, N. Kosimov, K. S. Bhandari, and D. W. Jung, "Research on High-Temperature Constitutive Relationship of Aluminum Alloy," *Solid State Phenom.*, vol. 335, pp. 101–106, 2022, doi: 10.4028/p-zr45qd.
- [68] W. R. Graff and D. C. Sargent, "A new grain-boundary etchant for aluminum alloys," *Metallography*, vol. 14, no. 1, pp. 69–72, Feb. 1981, doi: 10.1016/0026-0800(81)90021-5.

- [69] “Deformation microstructure and thermomechanical processing maps of homogenized AA2070 aluminum alloy,” *Mater. Sci. Eng. A*, vol. 834, p. 142619, Feb. 2022, doi: 10.1016/j.msea.2022.142619.
- [70] Y. Zhang, J. Jiang, Y. Wang, Y. Liu, and M. Huang, “Hot deformation behavior and microstructure evolution of hot-extruded 6A02 aluminum alloy,” *Mater. Charact.*, vol. 188, p. 111908, Jun. 2022, doi: 10.1016/j.matchar.2022.111908.
- [71] R. Sirichaivetkul, C. Limmaneevichitr, R. Tongstri, and J. Kajornchaiyakul, “Isothermal Investigation and Deformation Behavior during Homogenization of 6063 Aluminum Alloy,” *J. Mater. Eng. Perform.*, vol. 32, no. 2, pp. 638–650, Jan. 2023, doi: 10.1007/s11665-022-07139-y.
- [72] Y. V. R. K. Prasad, “Author’s reply: Dynamic materials model: Basis and principles,” *Metall. Mater. Trans. A*, vol. 27, no. 1, pp. 235–236, Jan. 1996, doi: 10.1007/BF02647765.
- [73] R. Wu *et al.*, “Study on hot deformation behavior and intrinsic workability of 6063 aluminum alloys using 3D processing map,” *J. Alloys Compd.*, vol. 713, pp. 212–221, Aug. 2017, doi: 10.1016/j.jallcom.2017.04.156.
- [74] X. Chen, Y. Si, R. Bai, X. Zhang, and Z. Li, “Hot Formability Study of Cr5 Alloy Steel by Integration of FEM and 3D Processing Maps,” *Materials*, vol. 15, no. 14, Art. no. 14, Jan. 2022, doi: 10.3390/ma15144801.
- [75] Y. Sun *et al.*, “3D processing map and hot deformation behavior of 6A02 aluminum alloy,” *J. Alloys Compd.*, vol. 742, pp. 356–368, Apr. 2018, doi: 10.1016/j.jallcom.2018.01.299.
- [76] X. Chen, Y. Du, T. Lian, K. Du, and T. Huang, “Hot Workability of Ultra-Supercritical Rotor Steel Using a 3-D Processing Map Based on the Dynamic Material Model,” *Materials*, vol. 13, no. 18, Art. no. 18, Jan. 2020, doi: 10.3390/ma13184118.
- [77] H. J. McQueen, E. Evangelista, N. Jin, and M. E. Kassner, “Energy dissipation efficiency in aluminum dependent on monotonic flow curves and dynamic recovery,” *Metall. Mater. Trans. A*, vol. 26, no. 7, pp. 1757–1766, Jul. 1995, doi: 10.1007/BF02670763.
- [78] H.-W. Son, T.-K. Jung, J.-W. Lee, and S.-K. Hyun, “Hot deformation characteristics of CaO-added AZ31 based on kinetic models and processing maps,” *Mater. Sci. Eng. A*, vol. 695, pp. 379–385, May 2017, doi: 10.1016/j.msea.2017.03.058.
- [79] G. Meng, B. Li, H. Li, H. Huang, and Z. Nie, “Hot deformation and processing maps of an Al–5.7wt.%Mg alloy with erbium,” *Mater. Sci. Eng. A*, vol. 517, no. 1, pp. 132–137, Aug. 2009, doi: 10.1016/j.msea.2009.03.068.
- [80] S. V. S. Narayana Murty, B. Nageswara Rao, and B. P. Kashyap, “Identification of flow instabilities in the processing maps of AISI 304 stainless steel,” *J. Mater. Process. Technol.*, vol. 166, no. 2, pp. 268–278, Aug. 2005, doi: 10.1016/j.jmatprotec.2004.09.089.
- [81] K. A. Marill, “Advanced Statistics: Linear Regression, Part II: Multiple Linear Regression,” *Acad. Emerg. Med.*, vol. 11, no. 1, pp. 94–102, 2004, doi: 10.1197/j.aem.2003.09.006.
- [82] C. J. Stone, “Additive Regression and Other Nonparametric Models,” *Ann. Stat.*, vol. 13, no. 2, pp. 689–705, Jun. 1985, doi: 10.1214/aos/1176349548.
- [83] T. Hastie and R. Tibshirani, “Generalized Additive Models: Some Applications,” *J. Am. Stat. Assoc.*, vol. 82, no. 398, pp. 371–386, Jun. 1987, doi: 10.1080/01621459.1987.10478440.
- [84] R. G. Brereton and G. R. Lloyd, “Support Vector Machines for classification and regression,”

- Analyst*, vol. 135, no. 2, pp. 230–267, Jan. 2010, doi: 10.1039/B918972F.
- [85] F. M. Javed Mehedi Shamrat, R. Ranjan, K. Md. Hasib, A. Yadav, and A. H. Siddique, “Performance Evaluation Among ID3, C4.5, and CART Decision Tree Algorithm,” in *Pervasive Computing and Social Networking*, G. Ranganathan, R. Bestak, R. Palanisamy, and Á. Rocha, Eds., in Lecture Notes in Networks and Systems. Singapore: Springer Nature, 2022, pp. 127–142. doi: 10.1007/978-981-16-5640-8_11.
- [86] W. Chen, S. Li, S. Aziz, K. S. Bhandari, X. Chen, and D.-W. Jung, “Flow Behavior Modeling Optimization and Activation Energy Analysis of Al-Mg Alloy Aided by Genetic Algorithm.” Rochester, NY, Oct. 05, 2022. doi: 10.2139/ssrn.4235225.
- [87] W.-Y. Loh, “Classification and regression trees,” *WIREs Data Min. Knowl. Discov.*, vol. 1, no. 1, pp. 14–23, 2011, doi: 10.1002/widm.8.
- [88] J. Elith, J. R. Leathwick, and T. Hastie, “A working guide to boosted regression trees,” *J. Anim. Ecol.*, vol. 77, no. 4, pp. 802–813, 2008, doi: 10.1111/j.1365-2656.2008.01390.x.
- [89] L. N. Kanal, “Perceptron,” in *Encyclopedia of Computer Science*, GBR: John Wiley and Sons Ltd., 2003, pp. 1383–1385.

Publications during the master's degree period

- [1] W. Chen, S. Li, K. S. Bhandari, S. Aziz, X. Chen, and D. W. Jung, "Genetic optimized Al-Mg alloy constitutive modeling and activation energy analysis," *Int. J. Mech. Sci.*, vol. 244, p. 108077, Apr. 2023, doi: 10.1016/j.ijmecsci.2022.108077.
- [2] "Supervised Shallow Machine Learning Strategies Aided Flow Behavior Modeling of Al-Mg-Si Alloy", prepared.
- [3] S. Li, W. Chen, K. S. Bhandari, D. W. Jung, and X. Chen, "Flow Behavior of AA5005 Alloy at High Temperature and Low Strain Rate Based on Arrhenius-Type Equation and Back Propagation Artificial Neural Network (BP-ANN) Model," *Materials*, vol. 15, no. 11, p. 3788, May 2022, doi: 10.3390/ma15113788.
- [4] W. N. Chen, S. J. Li, N. Kosimov, K. S. Bhandari, and D. W. Jung, "Research on High-Temperature Constitutive Relationship of Aluminum Alloy," *Solid State Phenom.*, vol. 335, pp. 101–106, 2022, doi: 10.4028/p-zr45qd.
- [5] W. N. Chen, S. J. Li, K. S. Bhandari, S. Aziz, N. Kosimov, and D. W. Jung, "A Flow Stress Equation of AA5005 Aluminum Alloy Based on Fields-Backofen Model," *Mater. Sci. Forum*, vol. 1078, pp. 3–10, 2022, doi: 10.4028/p-95f697.
- [6] W. Chen, S. Li, S. B. Krishna, N. Kosimov, and D. Jung, "Comparison of Shell Thickness Integration Rules and Element Types on Single Point Incremental Forming (SPIF) Simulation," *J. Adv. Technol. Eng. Res.*, vol. 7, no. 1, pp. 17–24, 2021, doi: <https://doi.org/10.20474/jater-7.1.3>.
- [7] S. J. Li, W. N. Chen, K. S. Bhandari, N. Kosimov, and D. W. Jung, "Study on Flow Stress Model of AA5005 Material," *Solid State Phenom.*, vol. 335, pp. 107–112, 2022, doi: 10.4028/p-4t00fs.
- [8] S. Li, W. Chen, N. Kosimov, S. B. Krishna, and D. Jung, "Reduction of Explicit Incremental Forming Simulation Time by Mass Scaling Parameters Adjusting and Symmetrical Setting," *Int. J. Technol. Eng. Stud.*, vol. 7, no. 1, pp. 27–34, doi: <https://dx.doi.org/10.20469/ijtes.7.10004-1>.
- [9] S. J. Li, W. N. Chen, S. Aziz, K. S. Bhandari, and D. W. Jung, "Modified Johnson-Cook model of AA5005 alloy for thermal tensile tests", accepted.
- [10] S. Aziz, J. Ali, K. S. Bhandari, W. Chen, S. Li, and D. W. Jung, "Reverse Offset Printed, Biocompatible Temperature Sensor Based on Dark Muscovado," *Sensors*, vol. 22, no. 22, Art. no. 22, Jan. 2022, doi: 10.3390/s22228726.
- [11] K. S. Bhandari, W. N. Chen, S. J. Li, and D. W. Jung, "Formability of Aluminum in Incremental Sheet Forming," *Solid State Phenom.*, vol. 330, pp. 77–81, 2022, doi: 10.4028/p-25oilk.
- [12] K. S. Bhandari, S. Aziz, W. N. Chen, S. J. Li, and D. W. Jung, "Forming Parameters and Optimization of A5052 in SPIF Process," *Mater. Sci. Forum*, vol. 1084, pp. 85–89, 2023, doi: 10.4028/p-248h81.
- [13] K. S. Bhandari, S. Aziz, W. N. Chen, S. J. Li, and D. W. Jung, "Deformation Evaluation of

- A5052 Sheet Metal in SPIF Process,” *Mater. Sci. Forum*, vol. 1084, pp. 91–95, 2023, doi: 10.4028/p-e6768o.
- [14] K. S. Bhandari, M. Murugesan, S. J. Li, W. N. Chen, J. Moon, and D. W. Jung, “Incremental Sheet Forming Simulation of Aluminum Alloys for Industry Level Production,” *Mater. Sci. Forum*, vol. 1054, pp. 129–133, 2022, doi: 10.4028/p-bgl4ij.
- [15] N. Kosimov, K. S. Bhandari, S. J. Li, W. N. Chen, J. Moon, and D. W. Jung, “Modeling Process of Manufacturing Parts Using Incremental Forming Process,” *Solid State Phenom.*, vol. 330, pp. 45–49, 2022, doi: 10.4028/p-gs3z36.
- [16] N. Kosimov, K. S. Bhandari, S. J. Li, W. N. Chen, and D. W. Jung, “Single Point Incremental Forming as a Cost-Effective Sheet Forming Process for Small Batch Production,” *Appl. Mech. Mater.*, vol. 908, pp. 169–175, 2022, doi: 10.4028/p-77k9g3.
- [17] N. Kosimov, K. S. Bhandari, S. J. Li, W. N. Chen, J. Moon, and D. W. Jung, “Developing Methods of Forming Sheet Materials Using Incremental Forming,” *Mater. Sci. Forum*, vol. 1054, pp. 123–127, 2022, doi: 10.4028/p-51vr3j.