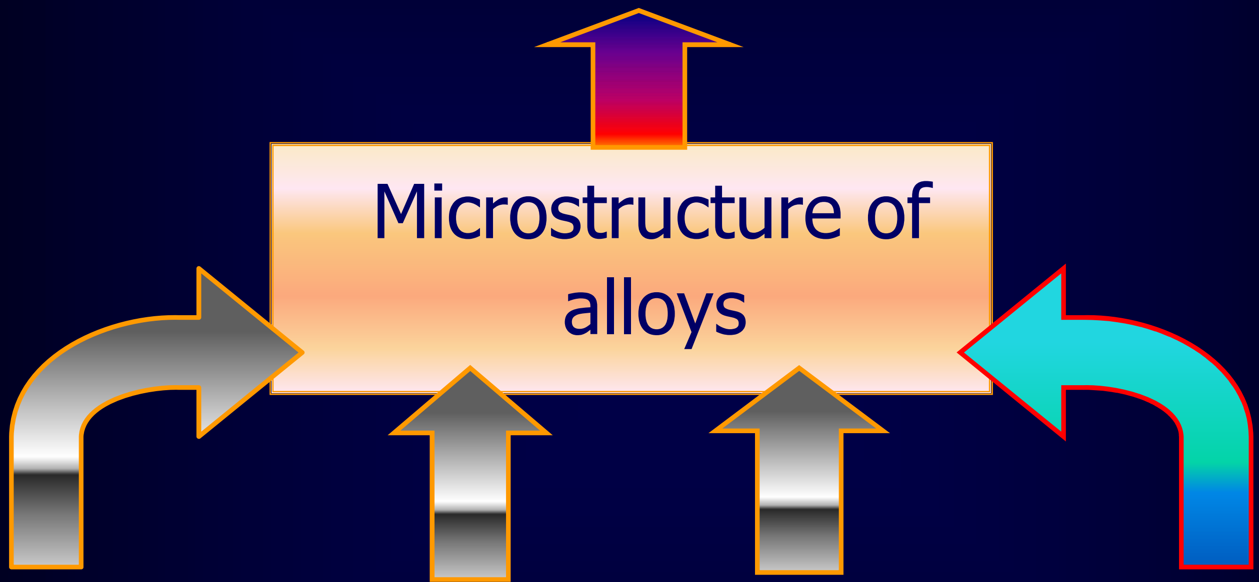
Artificial Neural Networks in Materials Science

N. S. Reddy

School of Materials Science and Engineering
Gyeongsang National University, Jinju,
Gyeongnam 660-701, Korea
nsreddy@gnu.ac.kr

Complex Materials System

Physical and Mechanical Properties



Mechanical Treatment

Alloy contents: C, Si, Mn, S, P, Ni, Cr, Mo….

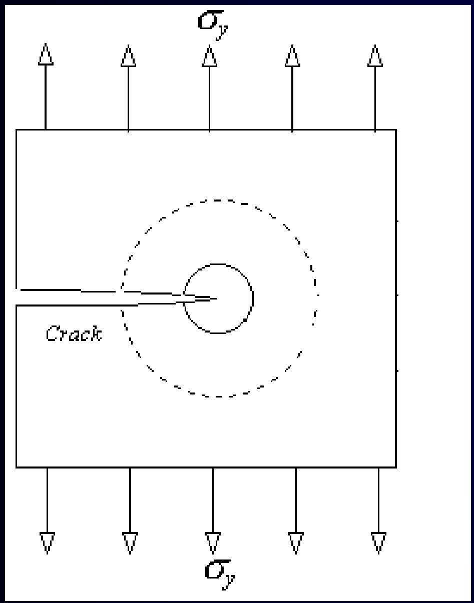
Solidification conditions

Heat treatment

Why

Charpy

Fatigue



Critical stress intensity

Corrosion

Tensile

Why

Axioms

- All properties can be measured.
- Measurements can be used in safe design.
- Measurements can be used in control.

Conclusions

- There are useful ways of expressing properties
- Limited models relating properties to independent variables
- No method for predicting properties in general

Thanks to Prof. HKDH Bhadeshia

Why

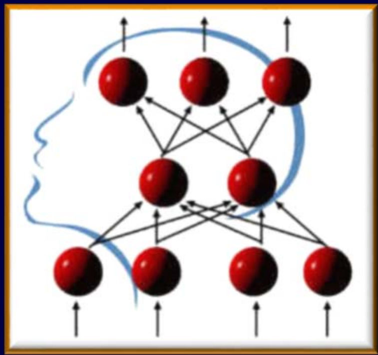
Pickering linear equations (1978)

$$\sigma_Y = 53.9 + 32.3W_{Mn} + 83.2W_{si} + 354.2(W_{Nf})^{0.5} + 17.4(d_\alpha)^{-0.5}$$

$$\sigma_U = 294.1 + 27.7W_{Mn} + 83.2W_{si} + 3.85(\%pearlite) + 7.7(d_\alpha)^{-0.5}$$

σ_Y is predicted yield strength in MPa and σ_U is predicted ultimate tensile strength in MPa, WMn, WSi and WNf are the contents of manganese, silicon and free nitrogen in weight percent respectively, and dα is the ferrite grain size in millimeters.

Contents Artificial Neural Networks (ANN)



- What
- Why
- When
- How
- Where

ANN Model Demonstration

Materials Science Problems

- Grain refinement in Al-7Si alloy (3 → 1)
- Phase volume fraction in Ti – 6Al – 4V alloy (6 → 2)
- Mechanical Properties in Steels (10 → 5)
- Estimation of Nano fiber diameter

Objectives

- To investigate the suitability of neural networks to complex Materials Systems.
- To predict properties/microstructure at new instances.
- To examine the Effect of Individual Elements on output parameters keeping other elements unaltered.
- To validate the model predictions with experiments

Why

A Brief history

- Early stages
  - 1943 McCulloch-Pitts: Neuron as computing element
  - 1949 Hebb: Learning rule
  - 1958 Rosenblatt: Perceptron
  - 1960 Widrow-Hoff: Least mean square algorithm
- Recession
  - 1969 Minsky-Papert: Limitations perceptron model
- Revival
  - 1982 Hopfield: Recurrent network model
  - 1982 Kohonen: Self-organizing maps
  - 1986 Rumelhart et. al.: Backpropagation

When

Where are ANN used?

- Recognizing and matching complicated, vague, or incomplete patterns
- Data is unreliable
- Problems with noisy data
  - Prediction
  - Classification
  - Data association
  - Data conceptualization
  - Filtering
  - Planning
  - **** …

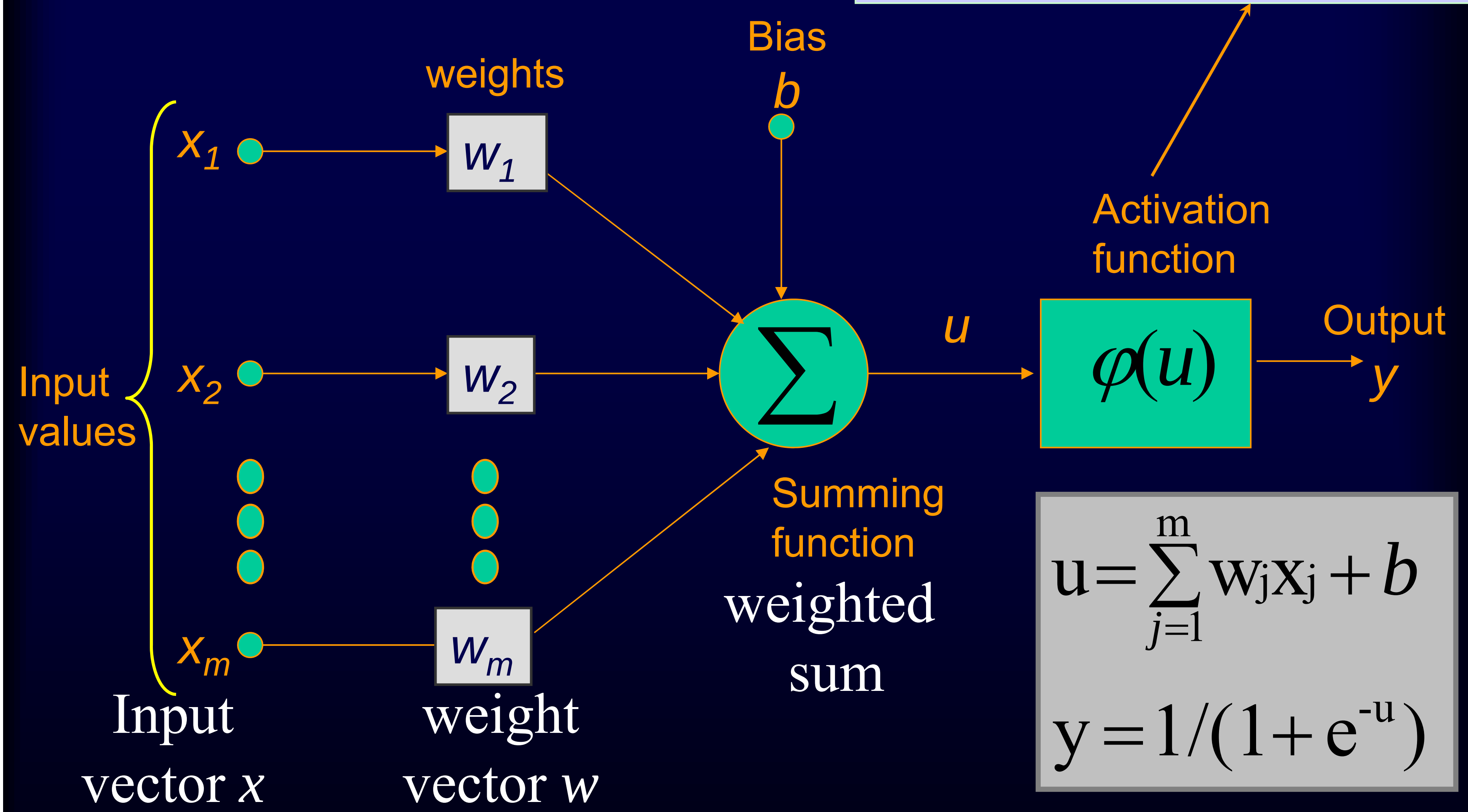
Neural Networks: Lessons from human brain

Artificial Neural networks (ANN) are modeling system, which mimics the human brain.

- Knowledge is acquired by the process of learning
- Storing the knowledge (like brain).
- Generalizations capability of the situation based on the acquired Knowledge

What

Simple Artificial Neuron



$$u = \sum_{j=1}^{m} w_j x_j + b$$

$$y = 1/(1+e^{-u})$$

The n-dimensional input vector x is mapped into variable y by means of the scalar product and a nonlinear function mapping

What

Feed Forward Neural Networks (FFNN)

- Neurons are arranged in layers.
- Each unit is linked only in the unit in next layer, no units are linked between the same layer, back to the previous layer or skipping a layer.
- Computations can proceed uniformly from input to output units.

<u>Parameters</u>

Learning Rate

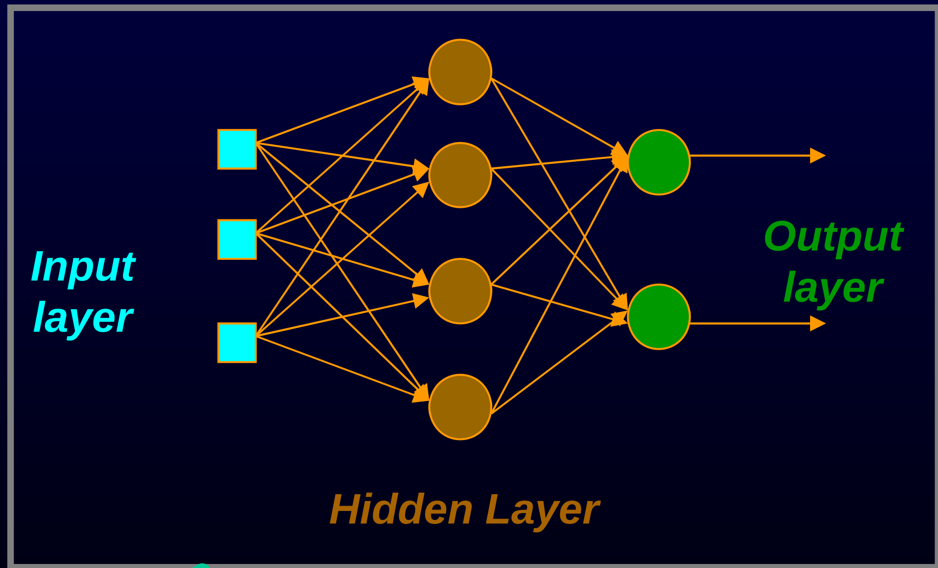
Momentum Term

Hidden Layers

Hidden Neurons

Iterations

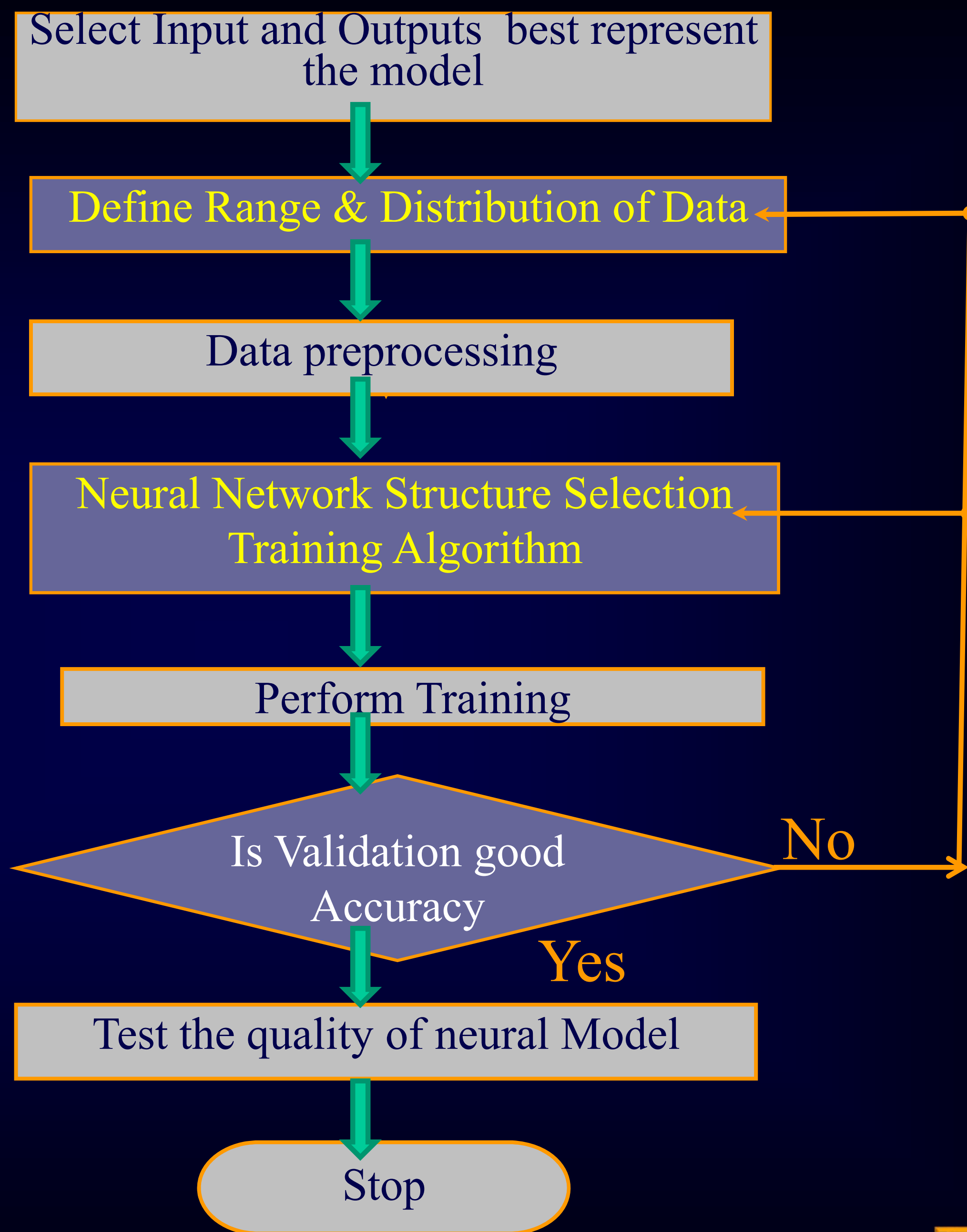
Information



Error Propagation

What

ANN Model Flowchart



$$MSE = \frac{1}{p}\sum_{j=1}^{p}[y_j^d - y_j^o]^2$$

y^d is the desired response, y^o is the output response from the ANN, and p is the of patterns presented

How

How does a neural network learn?

- A neural network learns by determining the relation between the inputs and outputs.
- By calculating the relative importance of the inputs and outputs the system can determine such relationships.
- Through trial and error the system compares its results with the expert provided results in the data until it has reached an accuracy level defined by the user.
  - With each trial the weight assigned to the inputs is changed until the desired results are reached.

How

Back propagation

- Desired output of the training examples
- Error = difference between actual & desired output
- Change weight relative to error size
- Calculate output layer error , then propagate back to previous layer
- Improved performance, very common!

How

Back Propagation Network

Algorithm

Step 1 : Initialize weights and offsets

Step 2 : Present Input and Desired Outputs

Step 3 : Calculate Actual Outputs

Step 4 : Adapt Weights

Step 5 : Repeat by going to Step 2, Until Convergence

Training Phase & Testing Phase

How

Learning in the BPN

- In the BPN, learning is performed as follows:

1. Randomly select a vector pair $(\mathbf{x}_p, \mathbf{y}_p)$ from the training set and call it $(\mathbf{x}, \mathbf{y})$.
2. Use $\mathbf{x}$ as input to the BPN and successively compute the outputs of all neurons in the network (bottom-up) until you get the network output $\mathbf{o}$.
3. Compute the error δ^o_{pk}, for the pattern p across all K output layer units by using the formula:

$$\delta^o_{pk} = (y_k - o_k) f'(net^o_k)$$

How

Learning in the BPN

4. Compute the error δ^{h}_{pj}, for all J hidden layer units by using the formula:

$$\delta_{pj}^{h} = f'(net_k^h)\sum_{k=1}^{K}\delta_{pk}^{o} w_{kj}$$

5. Update the connection-weight values to the hidden layer by using the following equation:

$$w_{ji}(t+1) = w_{ji}(t) + \eta\delta_{pj}^{h} x_i$$

How

Learning in the BPN

6. Update the connection-weight values to the output layer by using the following equation:

$$w_{kj}(t+1) = w_{kj}(t) + \eta \delta^{o}_{pk} f(net^{h}_{j})$$

Repeat steps 1 to 6 for all vector pairs in the training set; this is called a training **epoch**.

Run as many epochs as required to reduce the network error E to fall below a **threshold ε**:

$$E = \sum_{p=1}^{P} \sum_{k=1}^{K} (\delta^{o}_{pk})^2$$

How

Learning in the BPN

The only thing that we need to know before we can start our network is the **derivative** of our sigmoid function, for example, f'(net_k) for the output neurons:

$$f(\text{net}_k) = \frac{1}{1 + e^{-\text{net}_k}}$$

$$f'(\text{net}_k) = \frac{\partial f(\text{net}_k)}{\partial \text{net}_k} = o_k(1 - o_k)$$

How

Learning in the BPN

- **Now our BPN is ready to go!**
- If we choose the type and number of neurons in our network appropriately, after training the network should show the following behavior:
- If we input any of the training vectors, the network should yield the expected output vector (with some margin of error).
- If we input a vector that the network has never "seen" before, it should be able to generalize and yield a plausible output vector based on its knowledge about similar input vectors.

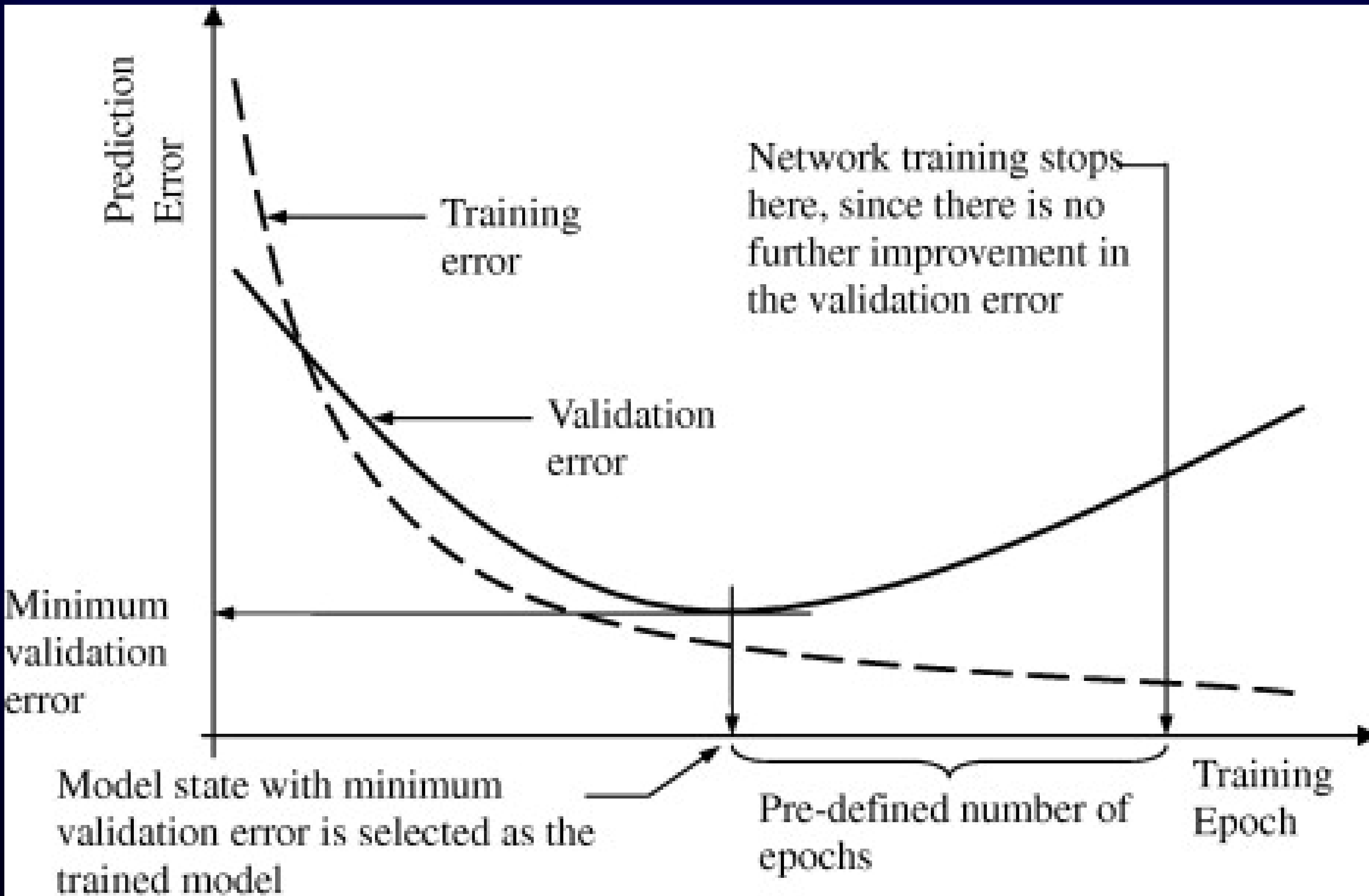
How

Criteria for model selection

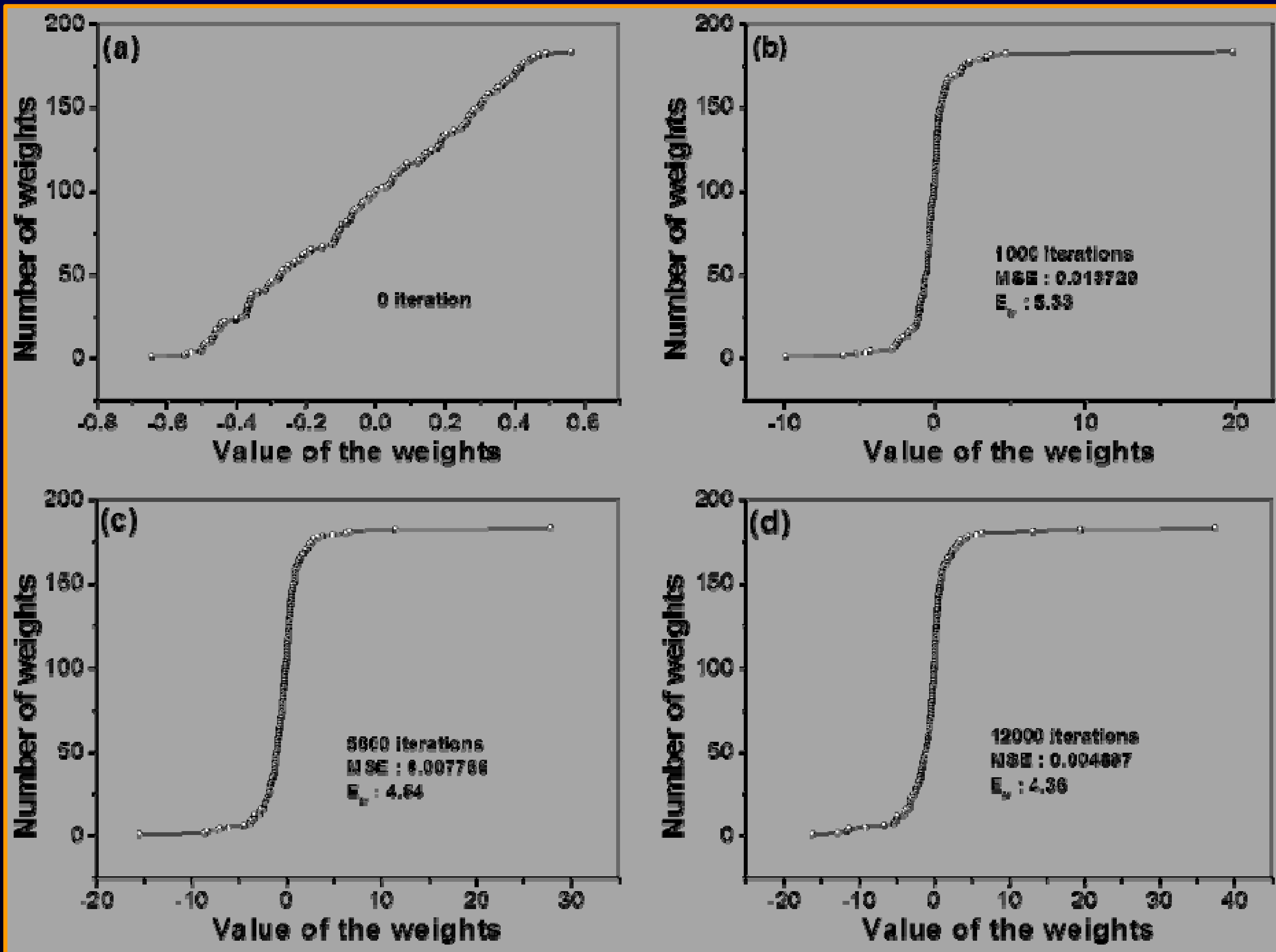
$$\mathrm{RMSE} = \frac{1}{p}\sum_{p}\sum_{i}(T_{\mathrm{ip}} - O_{\mathrm{ip}})^2$$

$$E_{\mathrm{tr}}(y) = \frac{1}{N}\sum_{i=1}^{N}\left|(T_i(y) - O_i(y))\right|$$

where Etr(y) = average error in prediction of training and testing data set for output parameter y, N = number of data sets, Ti(y) = targeted output, Oi(y) = output calculated.

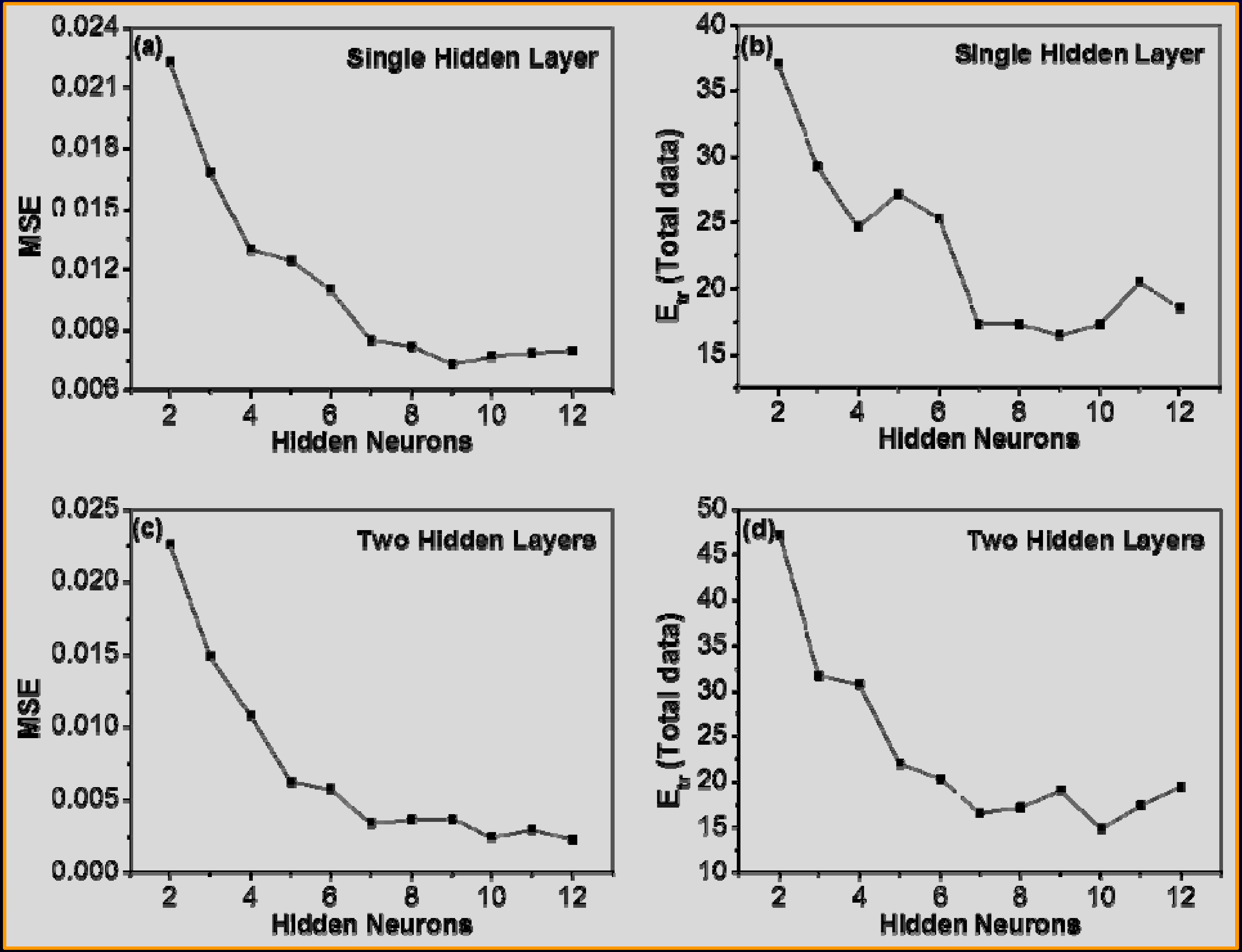


Prediction Error
Training error
Validation error
Network training stops here, since there is no further improvement in the validation error
Minimum validation error
Model state with minimum validation error is selected as the trained model
Pre-defined number of epochs
Training Epoch

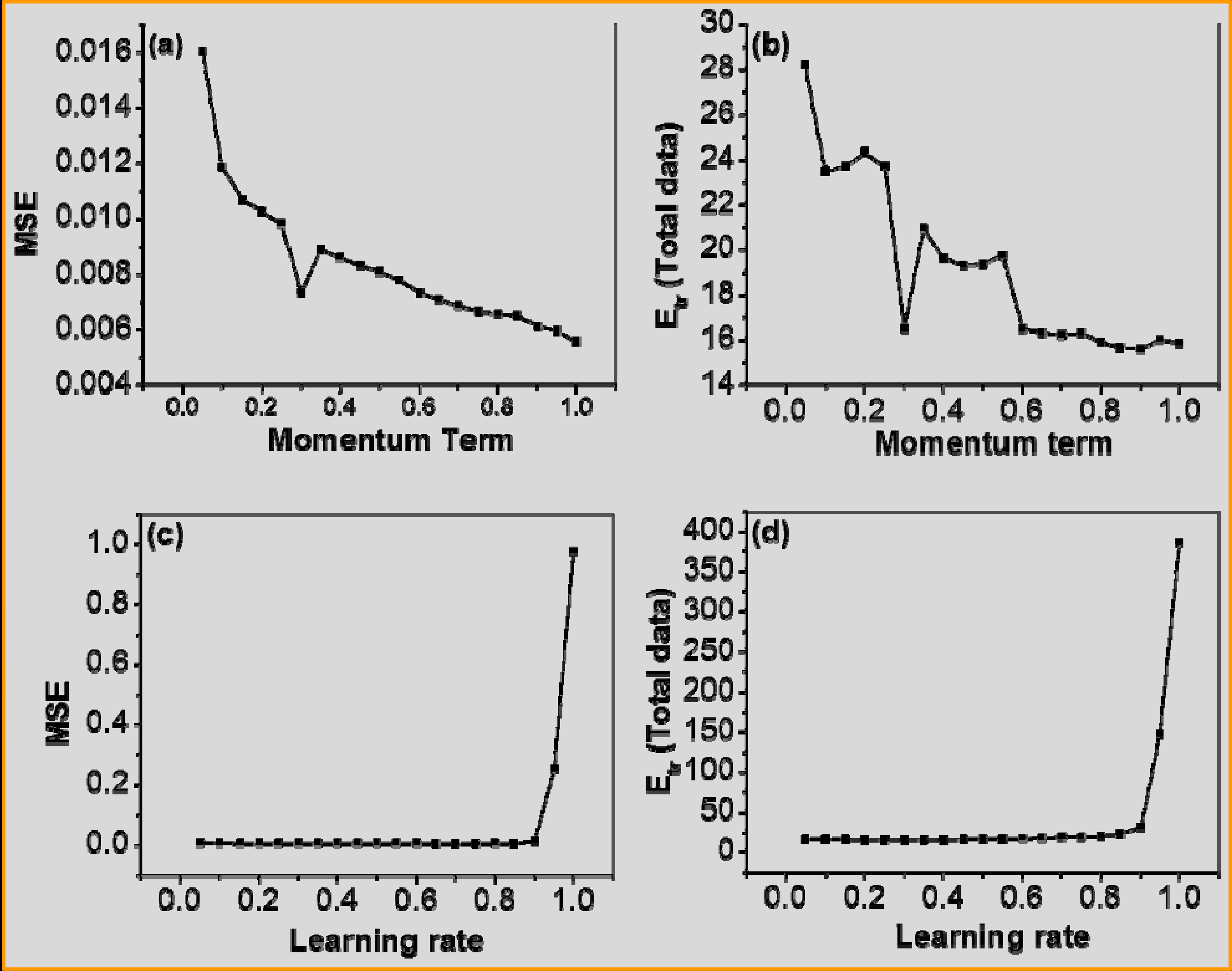


Variation of weights with varying iterations

Hidden layers



Momentum Term and Learning rate



WHAT ANN CAN DO?

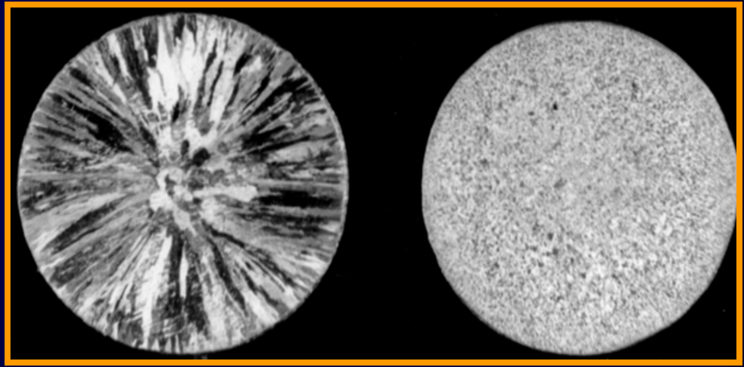
Demonstration of Neural Network Model

Example.1: Grain Refinement of Al-7Si Alloy

- Importance of Al-Si alloys

- Need of grain refinement

- How to achieve grain refinement



Master alloys for Grain refinement of Al –7Si alloy

Binary Master Alloys

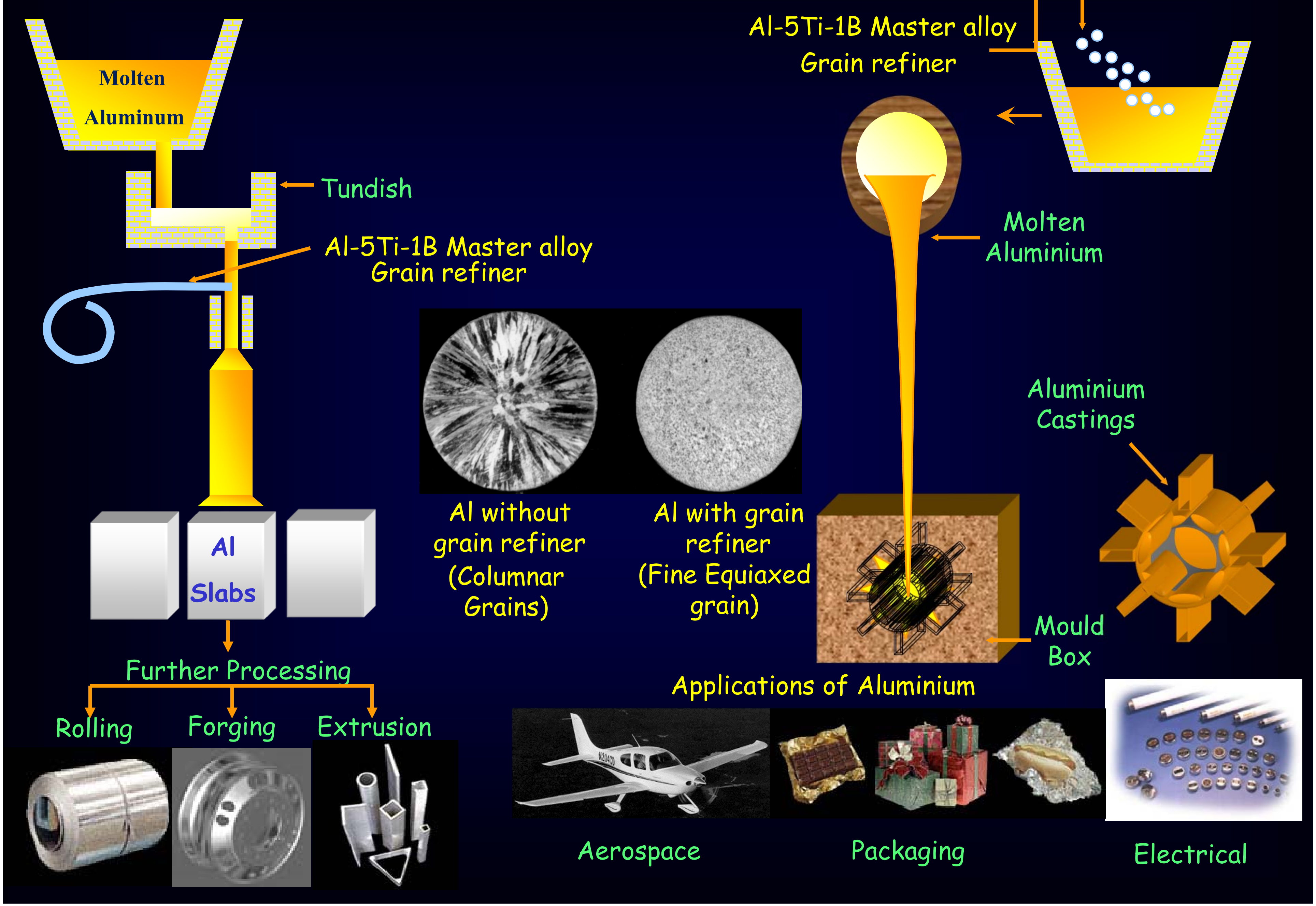
- Al – 3B
- Al – 3Ti

Ternary Master Alloys

- Al – 1Ti – 3B
- Al – 3Ti – 1B
- Al – 3Ti – 3B
- Al – 3Ti – 3B
- Al – 3Ti – 3B

Where

Example1. Grain refinement of Aluminum and its alloys



Example 1. Statistics of Grain refinement data

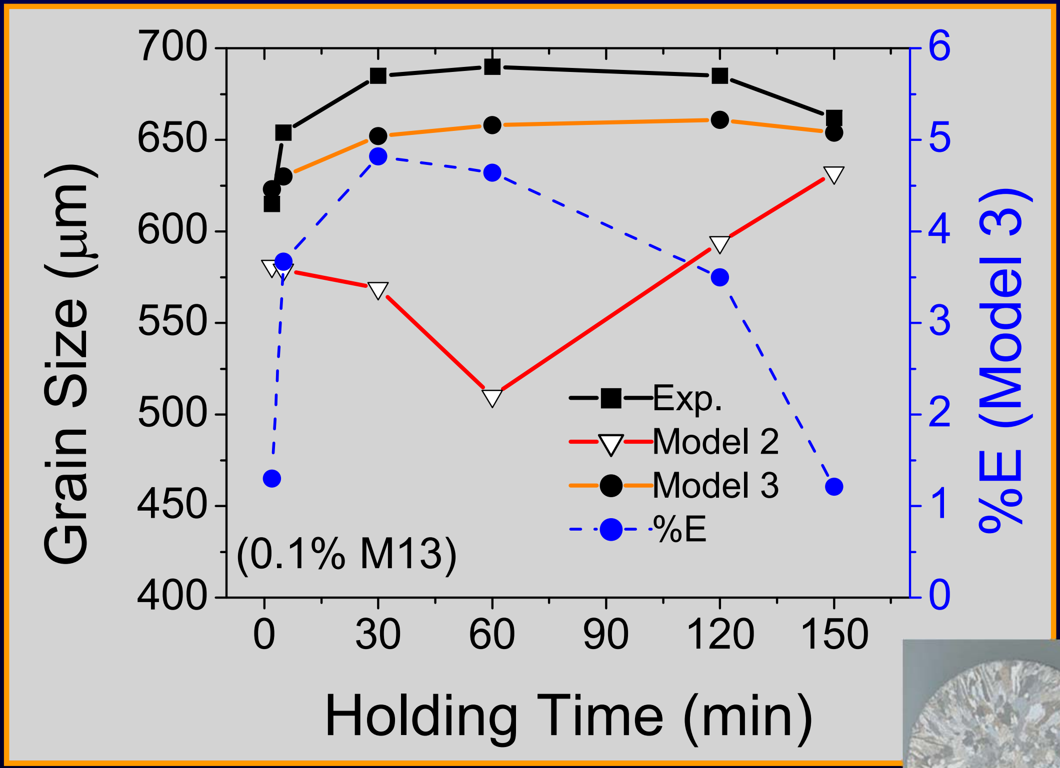
System (Train + Test) Sets	Variables	Minimum	Maximum	Mean	Standard Deviation
Binary Master alloy addition (48 +12)	Ti (%)	0	0.10	0.021	0.03
	B (%)	0	0.10	0.021	0.03
	Time (min)	0	120.00	40.69	43.86
	GS (μm)	98	610.00	276.08	158.16
Ternary Master alloy addition (120 + 30)	Ti (%)	0	0.10	0.029	0.028
	B (%)	0	0.10	0.029	0.028
	Time (min)	0	120	42.27	43.74
	GS (μm)	68	610.00	186.74	116.24
Total data (binary + Ternary alloy) (168 +42)	Ti (%)	0	0.10	0.023	0.029
	B (%)	0	0.10	0.022	0.028
	Time (min)	0	120.00	36.18	43.06
	GS (μm)	68	610.00	266.54	185.55

N.S. Reddy et al., "Materials Science and Engineering A", Vol. 391, 2005

Where

Validation of ANN model predictions with Al-1Ti-3B alloy

(0.1%) (B=0.003 & T=0.001)

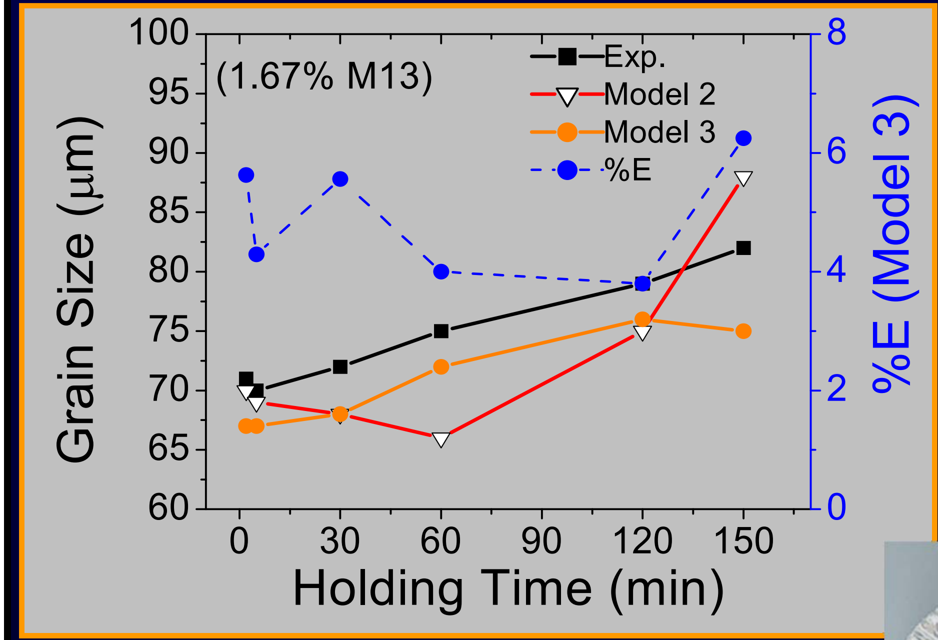


N.S. Reddy et al., "Materials Science and Engineering A", Vol. 391, 2005

Where

Validation of ANN model predictions with Al-1Ti-3B alloy

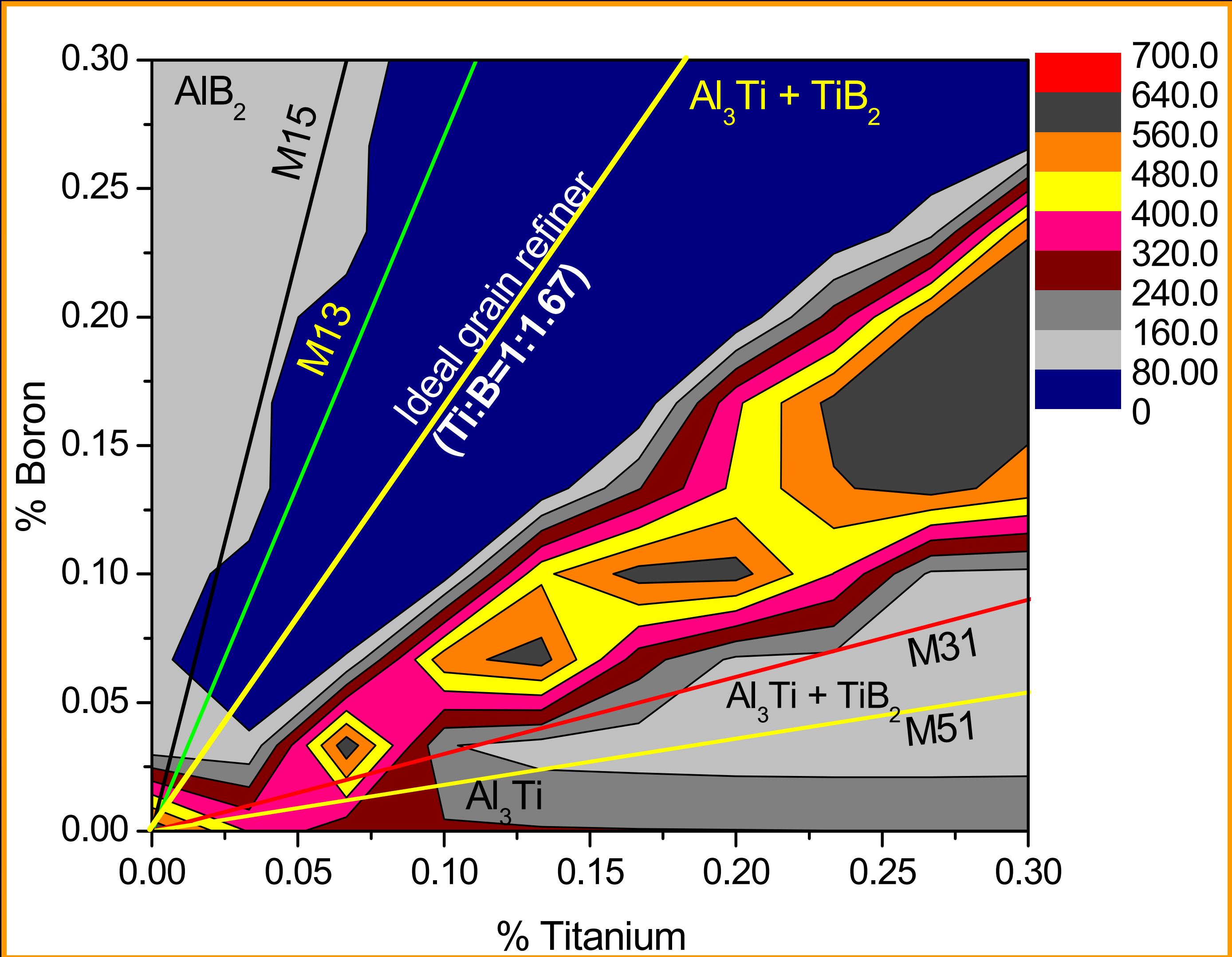
1.67% (B=0.05 & T=0.0167)



N.S. Reddy et al., "Materials Science and Engineering A", Vol. 391, 2005

Where

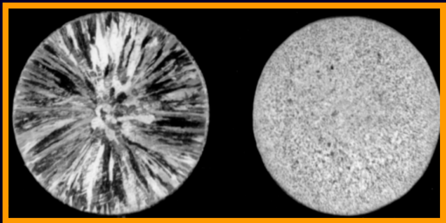
Predicted Grain refinement map of Al-7Si alloy



Color band numbers indicates grain size in µm

Master Alloys

Al – 1Ti – 3B (M13)
Al – 3Ti – 1B (M31)
Al – 3Ti – 3B (M33)
Al – 1Ti – 5B (M15)
Al – 5Ti – 1B (M51)



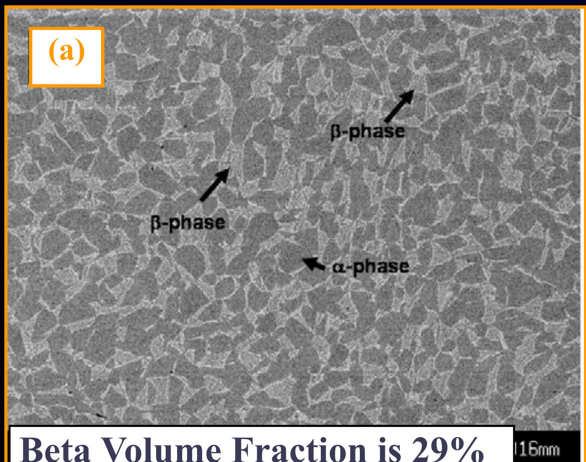
Ideal Grain Refiner

N.S. Reddy et al., Journal of Materials Performance, Vol 22, 2013

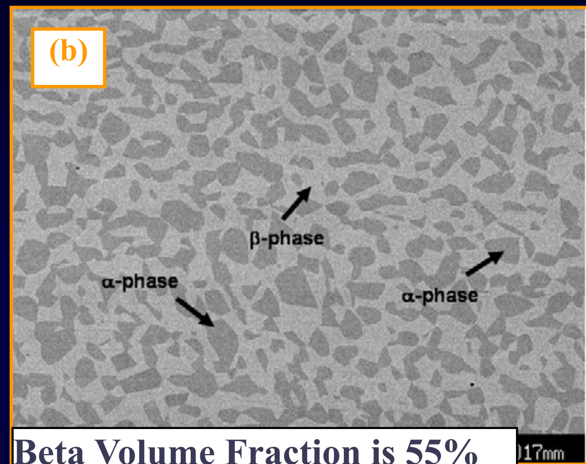
Where

Example 2. : Phase volume fraction in Ti-6Al-4V alloy

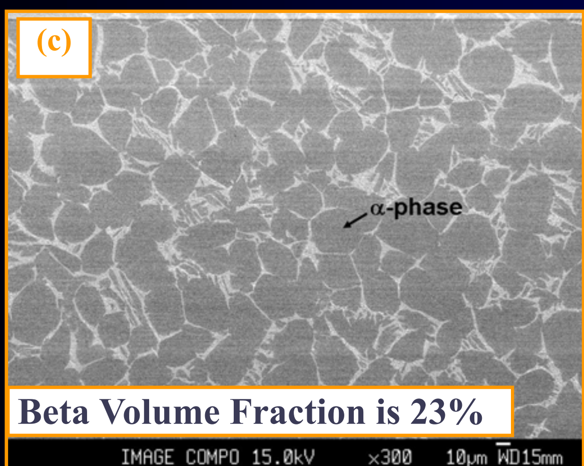
Ti-6.19Al-4.05V-0.19Fe-0.12O-0.01N

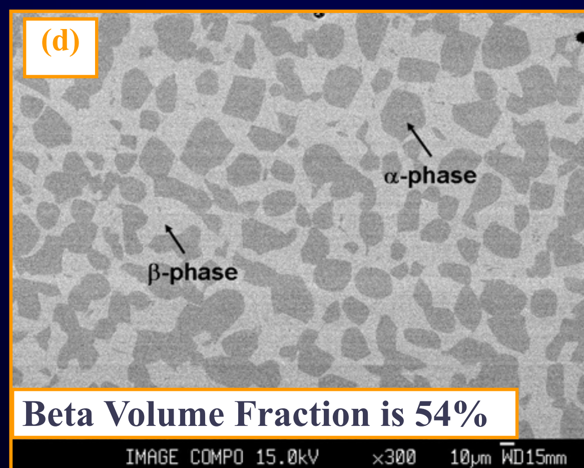


[750 ºC]



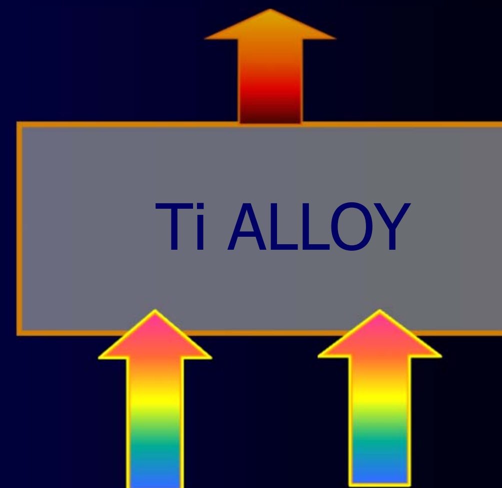
[900 ºC]





Ti-6.3Al-4.1V-0.21Fe-0.168O-0.005N

Volume fraction of α-β phases



Alloy contents: Al, V, Fe, O & N

Heat treatment

Where

Application of Ti-6Al-4V alloy

F-22 Raptor

Biomaterial

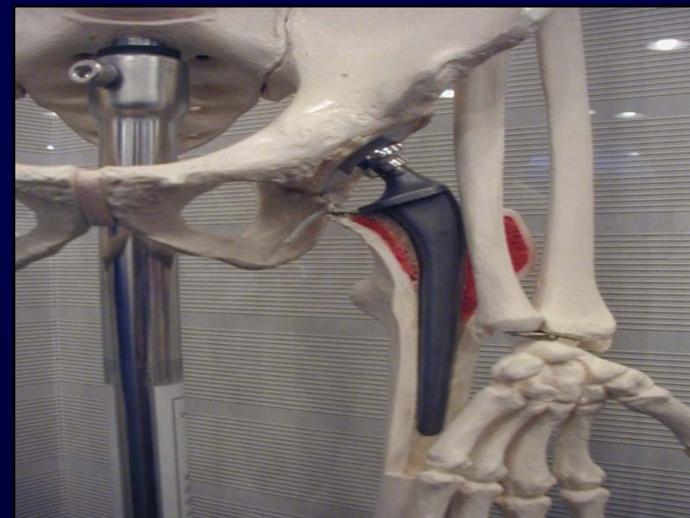
Guggenheim Museum

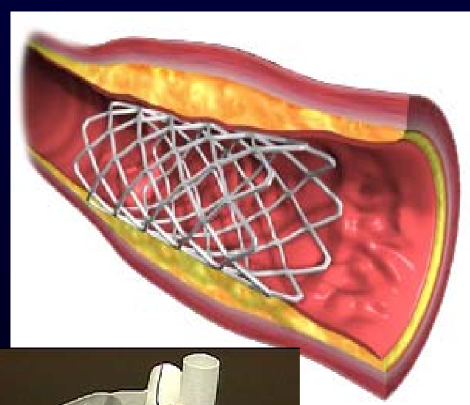
M2A3 Bradley

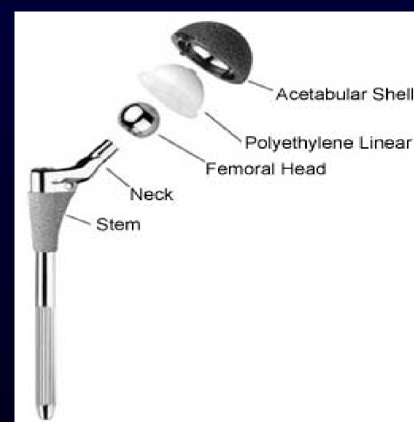
Tomahawk

Golf-head

Ti-6Al-4V alloy as a Biomaterial







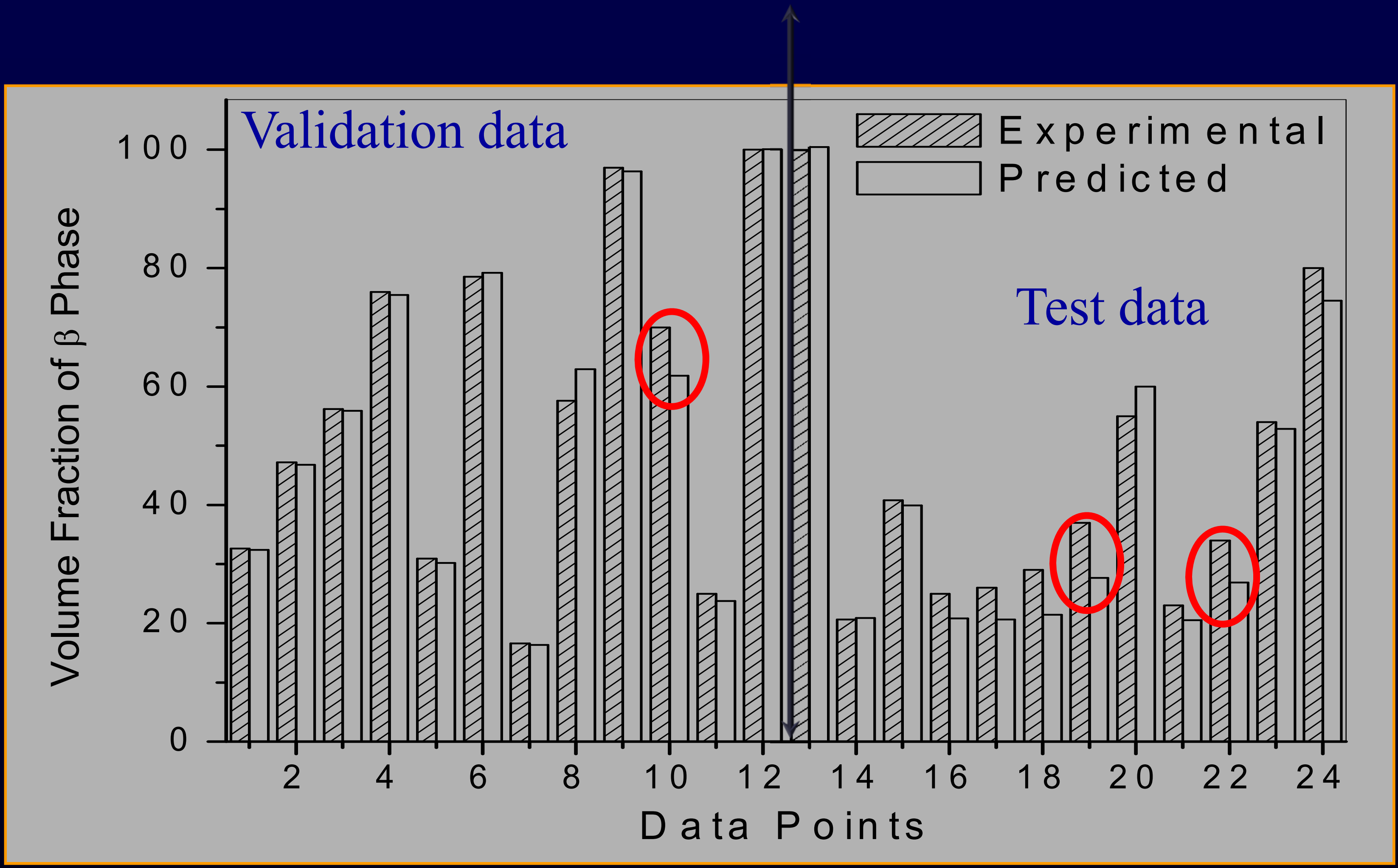
Where

Statistics of data used for modeling (Ti Alloys)

Experimental data	Input and output Variables	Minimum	Maximum	Mean	Standard deviation
99 Training + 35 test data sets	Al (%)	5.72	7	6.244	0.071
	V (%)	1.5	5	3.948	0.071
	Fe (%)	0.01	3.04	0.444	0.049
	O (%)	0.08	0.3	0.148	0.016
	N (%)	0.003	0.02	0.007	0.000
	Temperature (^{o}C)	600	1000.62	861	132.158
	α phase volume fraction (%)	0	100	52.7	42.002
	β phase volume fraction (%)	0	100	47.3	42.002

N.S. Reddy et al., "Computational Materials Science", Vol. 107 2015

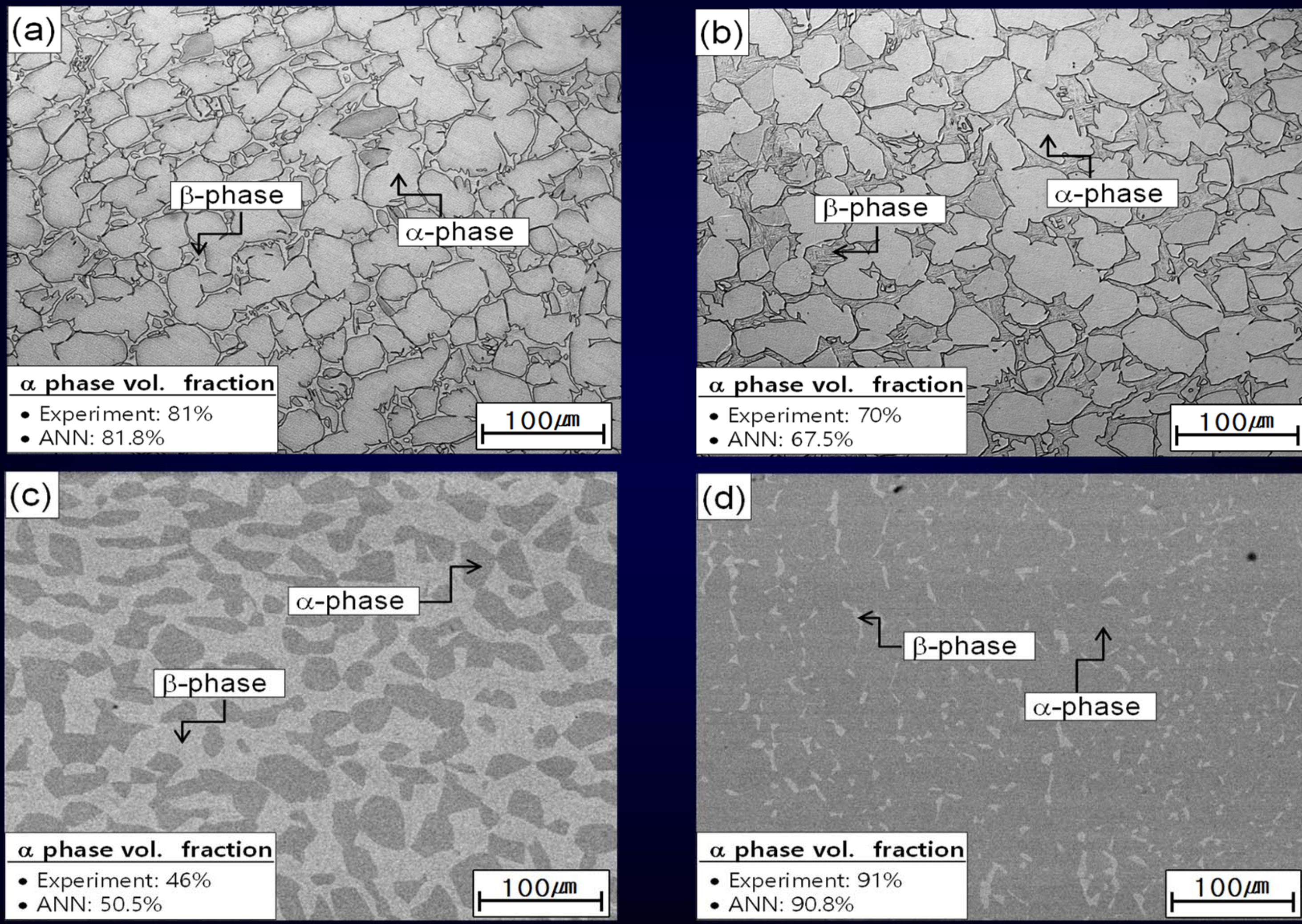
Performance of ANN Model : Validation



N.S. Reddy et al., "Materials Science and Engineering A", Vol. 434, 2006

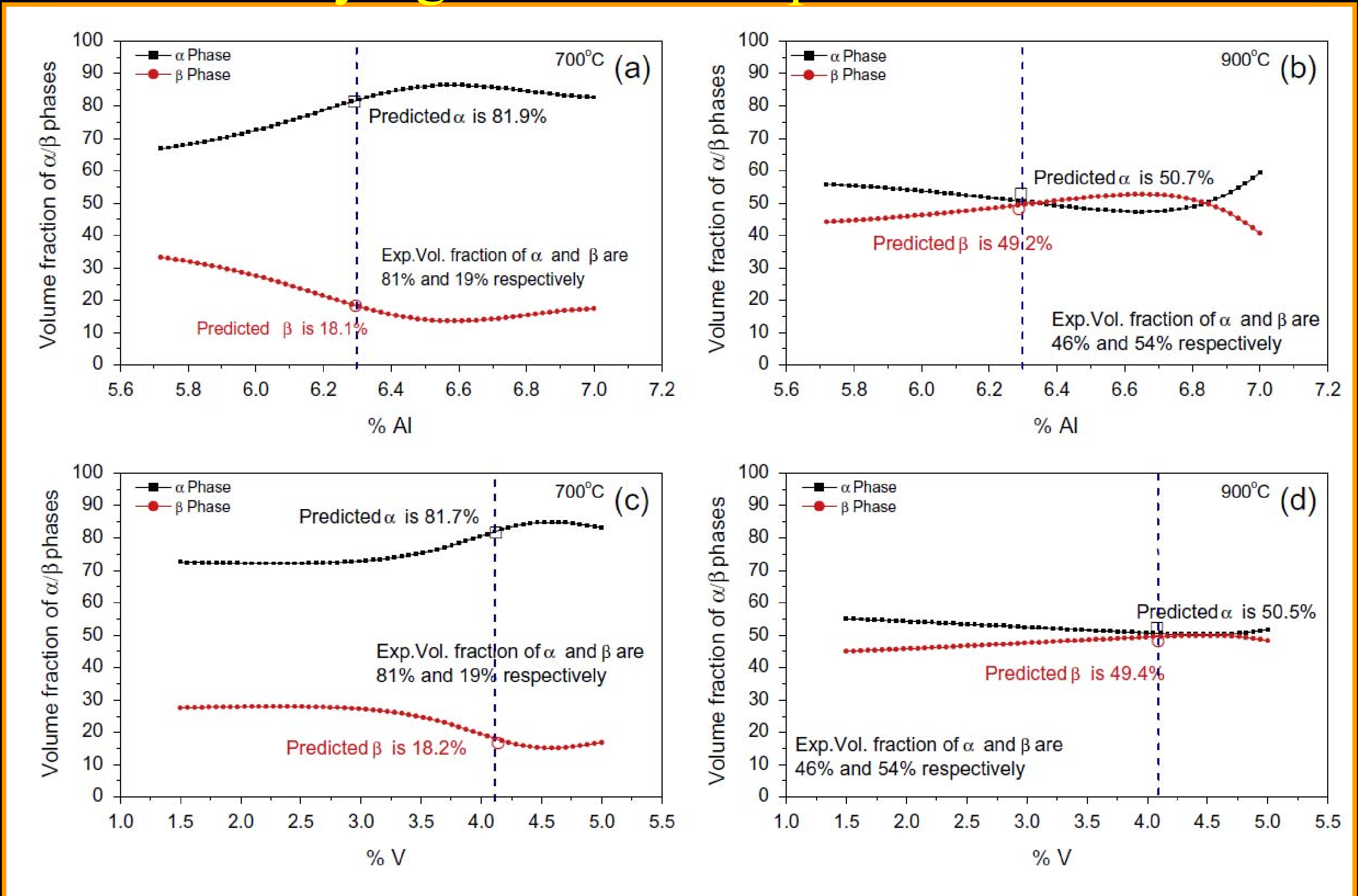
Where

Ti-6.3Al-4.1V-0.21Fe-0.17-0.005N alloy quenched at (a) 700 °C, (b) 815 °C, (c) 900 °C, and (d) in Ti-6.85Al-1.6V-0.13Fe-0.17-0.001N quenched at 900 °C.

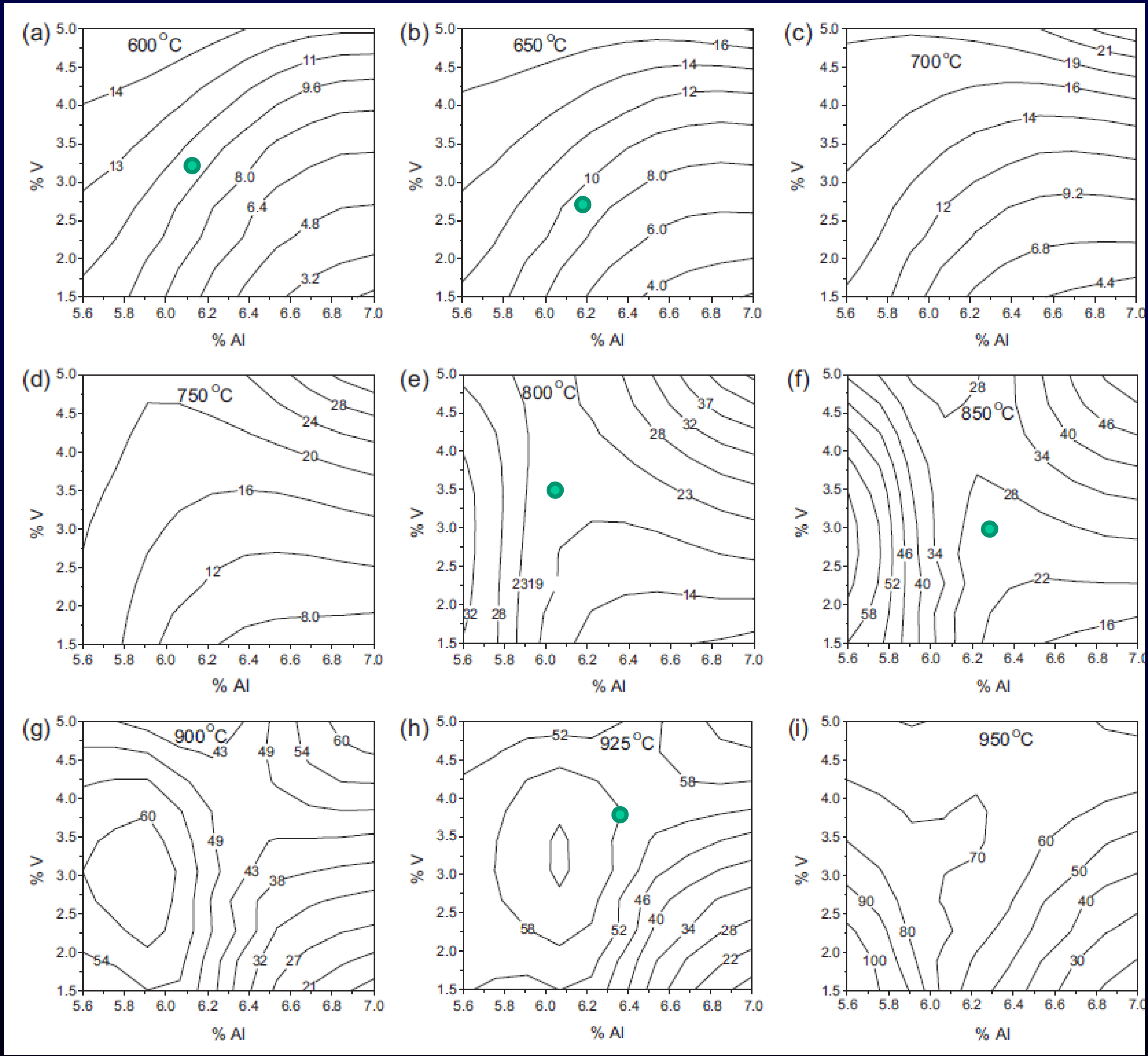


N.S. Reddy et al., "Computational Materials Science", Vol. 107 2015

Effect of alloying elements on phase volume fraction

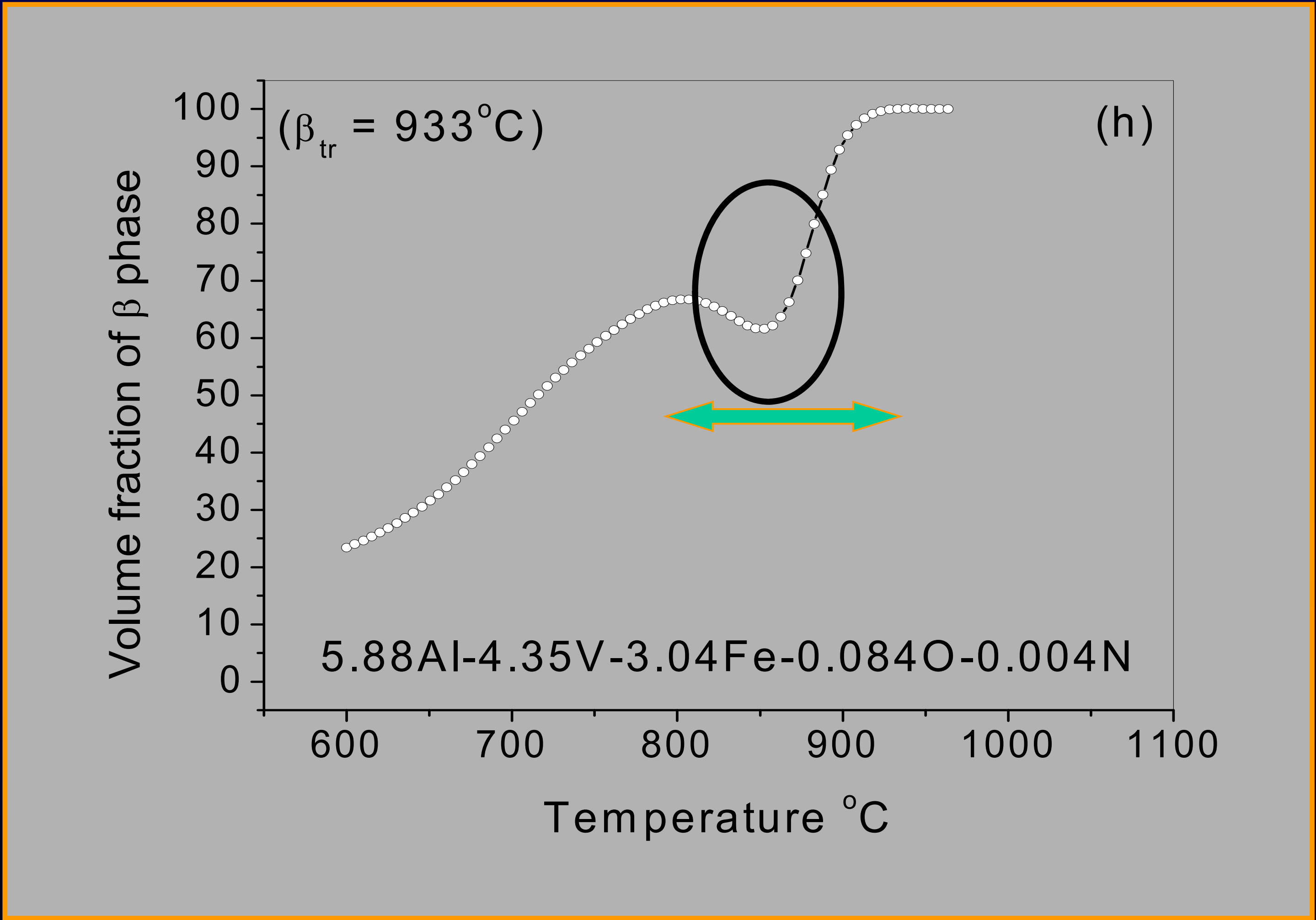


N.S. Reddy et al., "Computational Materials Science", Vol. 107, 2015



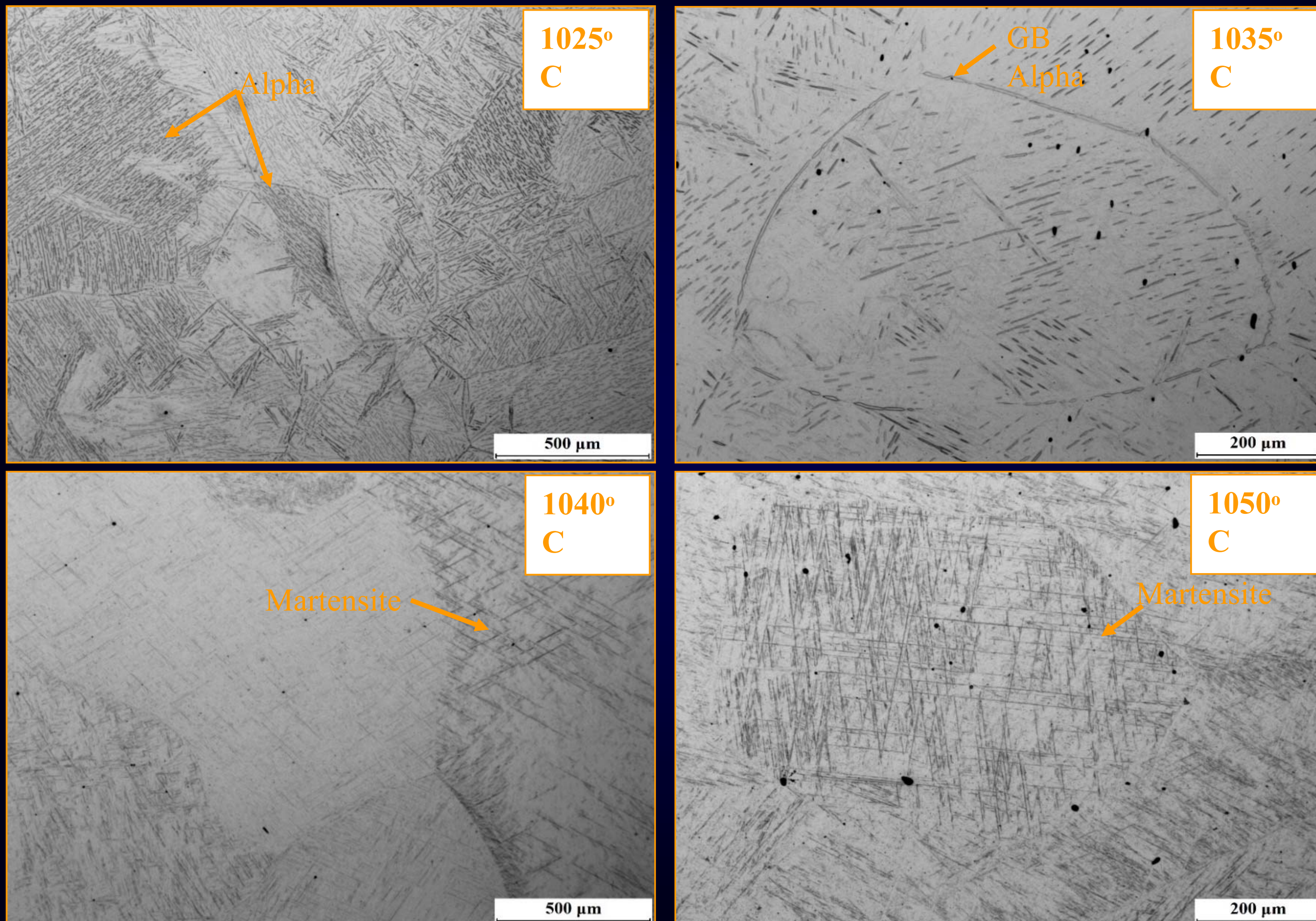
Volume Fraction of Beta map of Ti-Al- V at different temperatures

Unexpected Trend: Uncertainty



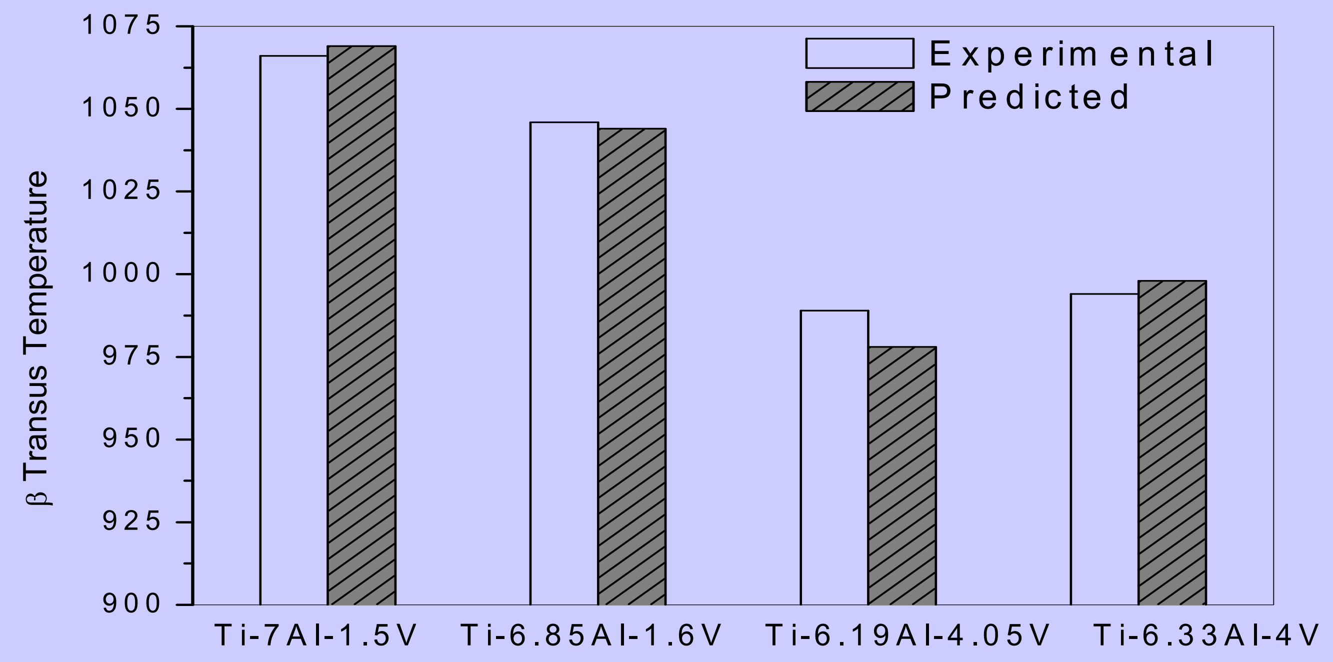
N.S. Reddy et al., "Materials Science and Engineering A", Vol. 434, 2006

Estimation of beta transus temperature
Experimental: 1040°C



➢ **ANN Prediction: 1039.88°C ; J Met Pro: 1076.77°C**

Validation of ANN model predictions



Comparison of predicted and measured beta-transus temperatures for single-phase-alpha, near-alpha, and alpha/beta titanium alloys.

N.S. Reddy et al., "Materials Science and Engineering A", Vol. 434, 2006

Where

Example 3.Low alloy steels data examples

Appendix A. Low alloy steels used in modeling: chemical composition, heat treatment parameters (CR: cooling rate in °C/s and TT: tempering temperature in °C) and mechanical properties.

Sl. no.	C	Si	Mn	P	Ni	Cr	Mo	Mn/S	CR	TT	YS	UTS	EL	RA	IS
1	0.32	0.23	1.28	0.028	0.85	0.46	0.16	38	38	580	934	1019	18.0	54.0	35
2	0.32	0.23	1.28	0.028	0.85	0.46	0.16	38	16	620	736	845	22.0	59.0	42
3	0.33	0.19	1.45	0.026	0.89	0.56	0.12	73	16	605	785	888	21.0	60.0	72
4	0.33	0.19	1.45	0.026	0.89	0.56	0.12	73	7	625	693	824	21.5	57.0	43
5	0.35	0.19	1.5	0.026	0.93	0.57	0.18	75	7	650	681	839	21.5	58.0	89
6	0.35	0.19	1.5	0.026	0.93	0.57	0.18	75	3	650	730	870	24.0	56.0	87
7	0.35	0.19	1.33	0.026	0.91	0.43	0.15	67	3	550	683	850	21.0	57.0	46
8	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	38	450	1163	1222	15.0	50.0	15
9	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	38	500	1069	1131	18.5	55.0	26
10	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	38	550	906	993	19.0	55.5	49
11	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	38	600	786	906	24.0	61.5	75
12	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	38	650	718	821	25.5	66.0	93
13	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	38	675	692	763	27.0	64.5	94
14	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	38	700	575	753	27.5	64.5	90
15	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	12	450	727	888	21.0	56.5	23
16	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	12	500	645	821	24.0	62.5	52
17	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	12	550	622	783	26.0	63.0	70
18	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	12	600	579	757	27.0	66.0	91
19	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	12	625	582	737	27.5	61.0	80
20	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	7	500	639	800	23.0	58.0	36
21	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	7	550	614	777	25.0	60.0	53
22	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	7	600	623	777	24.0	63.0	70
23	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	7	625	543	707	27.0	64.5	79
24	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	7	675	870	1008	14.0	46.5	93
25	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	3	400	683	850	21.0	57.0	19
26	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	3	500	661	829	23.0	57.5	47
27	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	3	575	620	772	24.0	60.0	74
28	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	3	600	588	754	25.0	61.5	61
29	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	3	625	558	713	27.0	64.5	75
30	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	3	675	750	903	18.5	52.0	82
31	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	3	700	667	835	22.0	56.5	20
32	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	118	400	941	1002	20.0	56.0	24
33	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	118	500	855	912	22.5	61.0	42
34	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	118	550	745	810	24.5	63.5	79
35	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	118	600	710	778	26.0	64.0	84
36	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	34	400	985	1078	17.0	52.5	21
37	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	34	500	813	920	21.0	61.0	45
38	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	34	550	742	853	24.0	63.0	55
39	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	34	600	663	800	25.0	[illegible]	[illegible]
40	0.35	0.25	0.33	0.032	0.91	0.43	0.15	8	34	625	605	766	26.0	[illegible]	[illegible]

N.S. Reddy et al., "Materials Science and Engineering A", Vol. 508, 2009

Where

The range of low alloy steels (EN100 steels)

Composition (in wt.%)	Min.	Max.
C	0.32	0.44
Si	0.19	0.37
Mn	0.33	1.51
S	0.01	0.042
P	0.02	0.038
Ni	0.56	1.08
Cr	0.21	0.57
Mo	0.11	0.25

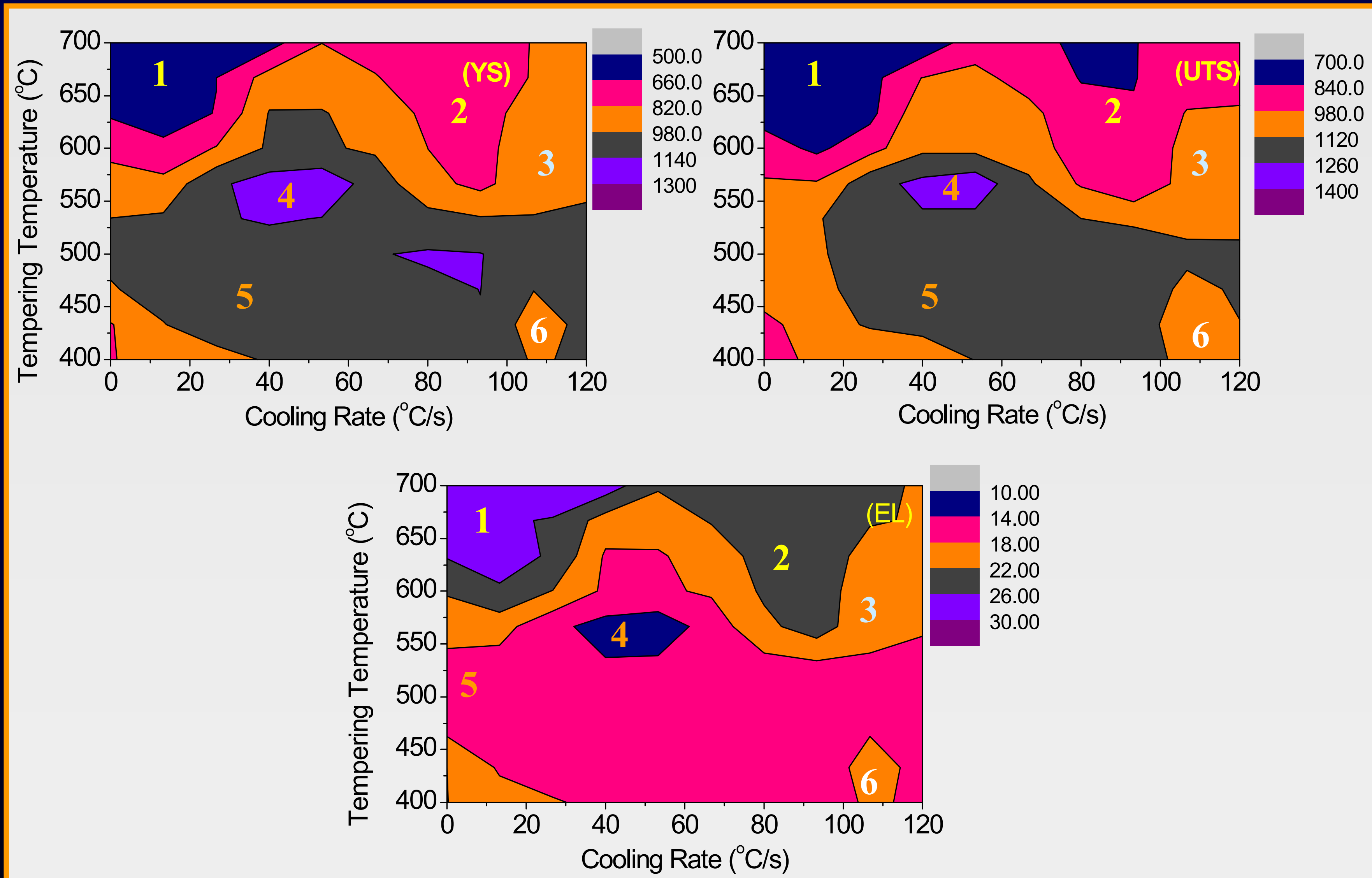
Heat treatment variables		
Cooling rate (^{o}C/S)	2.8	118
Tempering Temperature (^{o}C)	400	700
Mechanical properties		
Y.S (MPa)	542	1194
UTS (MPa)	707	1295
% El	13	29
% RA	31	67
Impact Strength (J)	15	94

YS: Yield strength, UTS: Ultimate tensile strength, %El: % Elongation, %RA: % Reduction in Area. Total data sets are 140 and 112 were used for training and 28 used for testing.

N.S. Reddy et al., "Materials Science and Engineering A", Vol. 508, 2009

Where

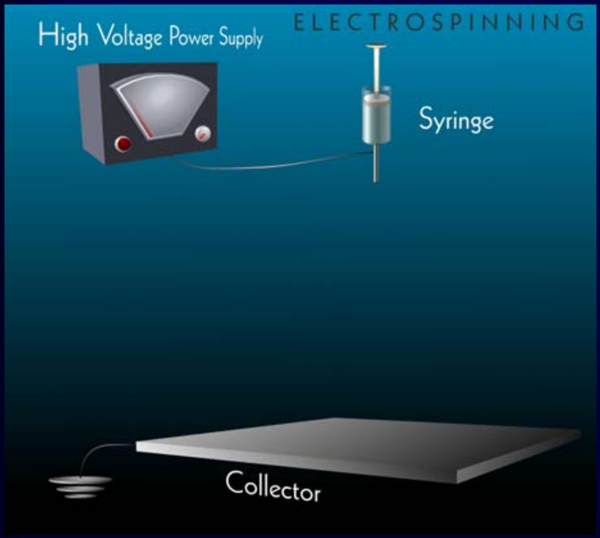
Simulated effect of heat treatment parameters

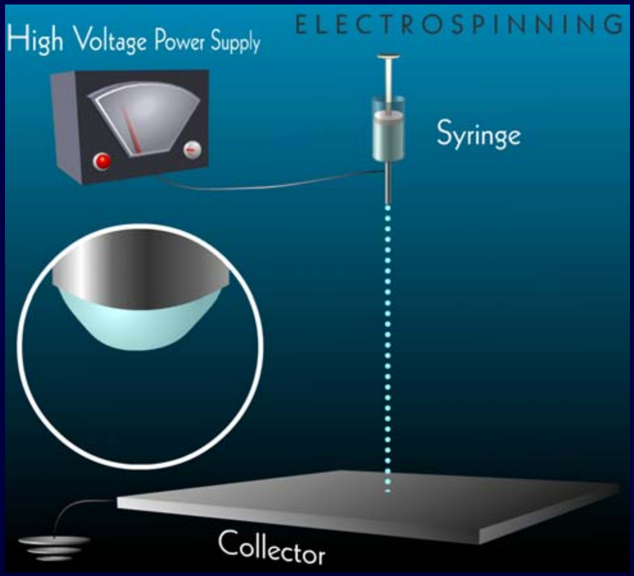


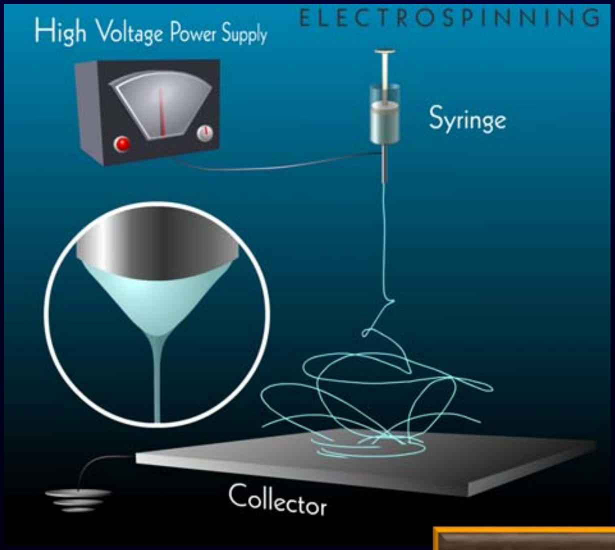
N.S. Reddy et al., "Materials Science and Engineering A", Vol. 508, 2009

Where

Example 4: Electrospinning



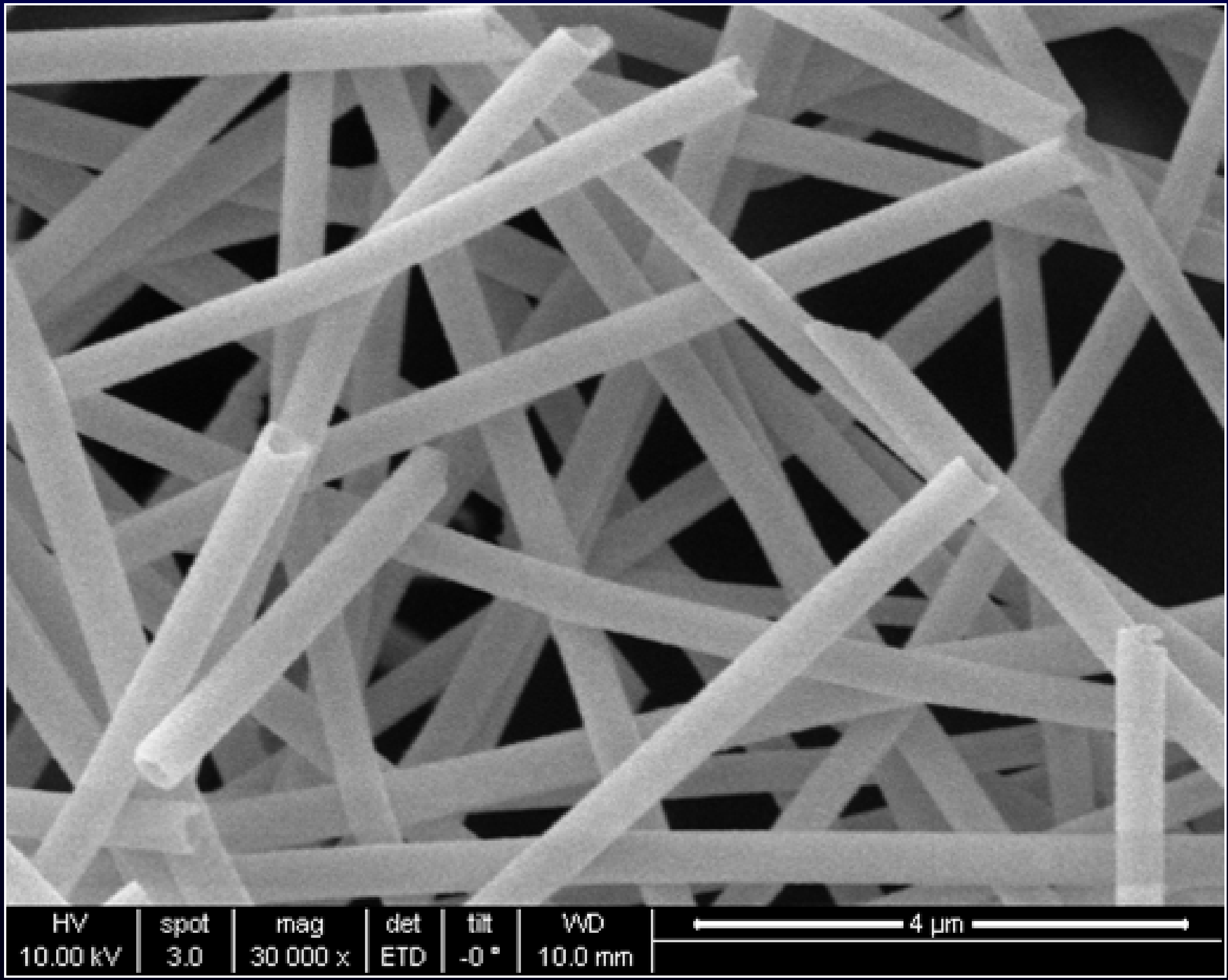


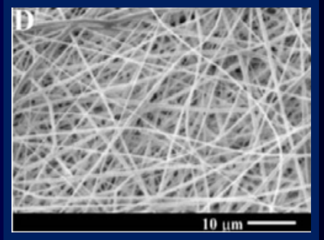


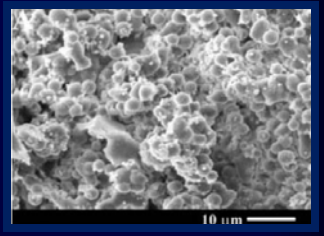
Energy

What

Nano fibers



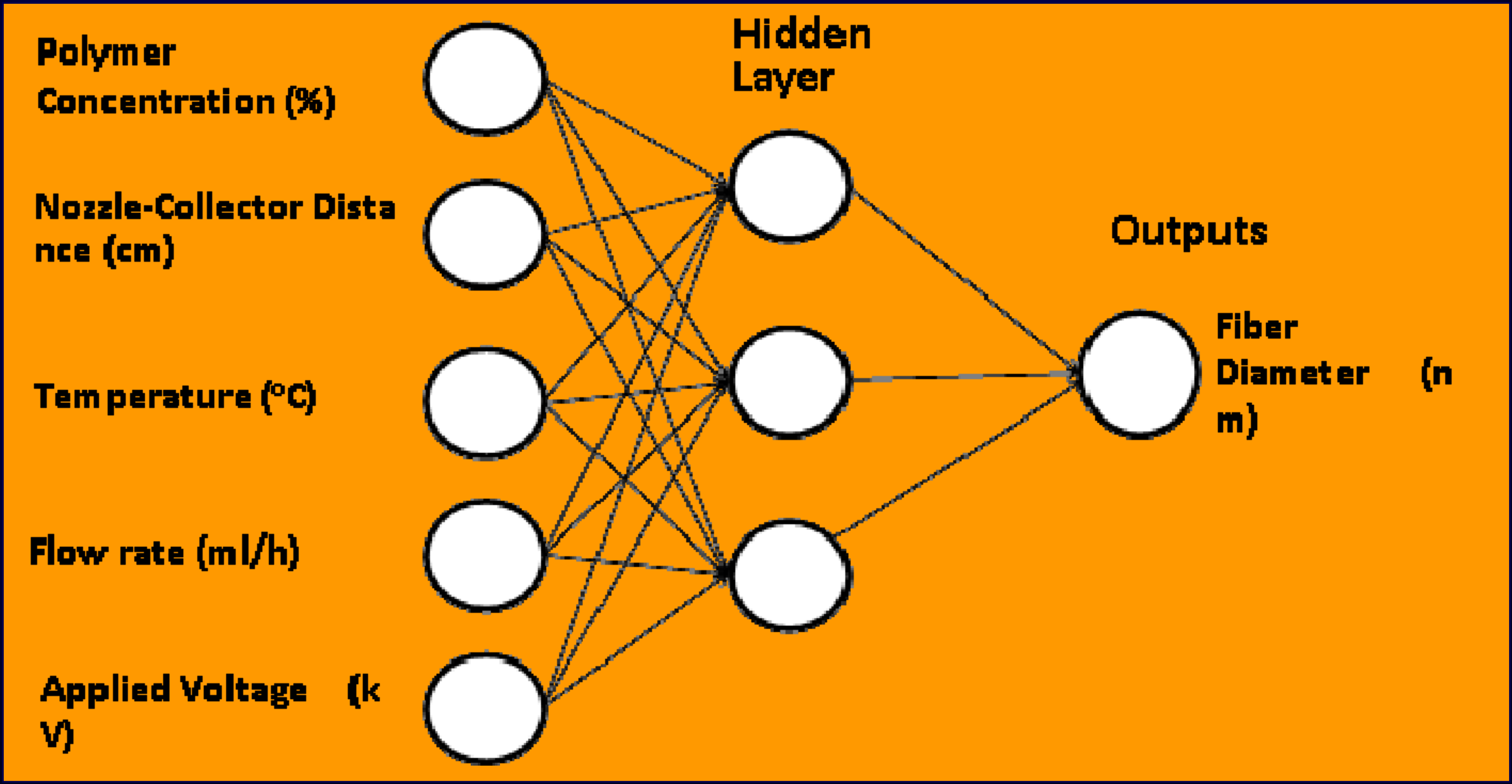




Energy

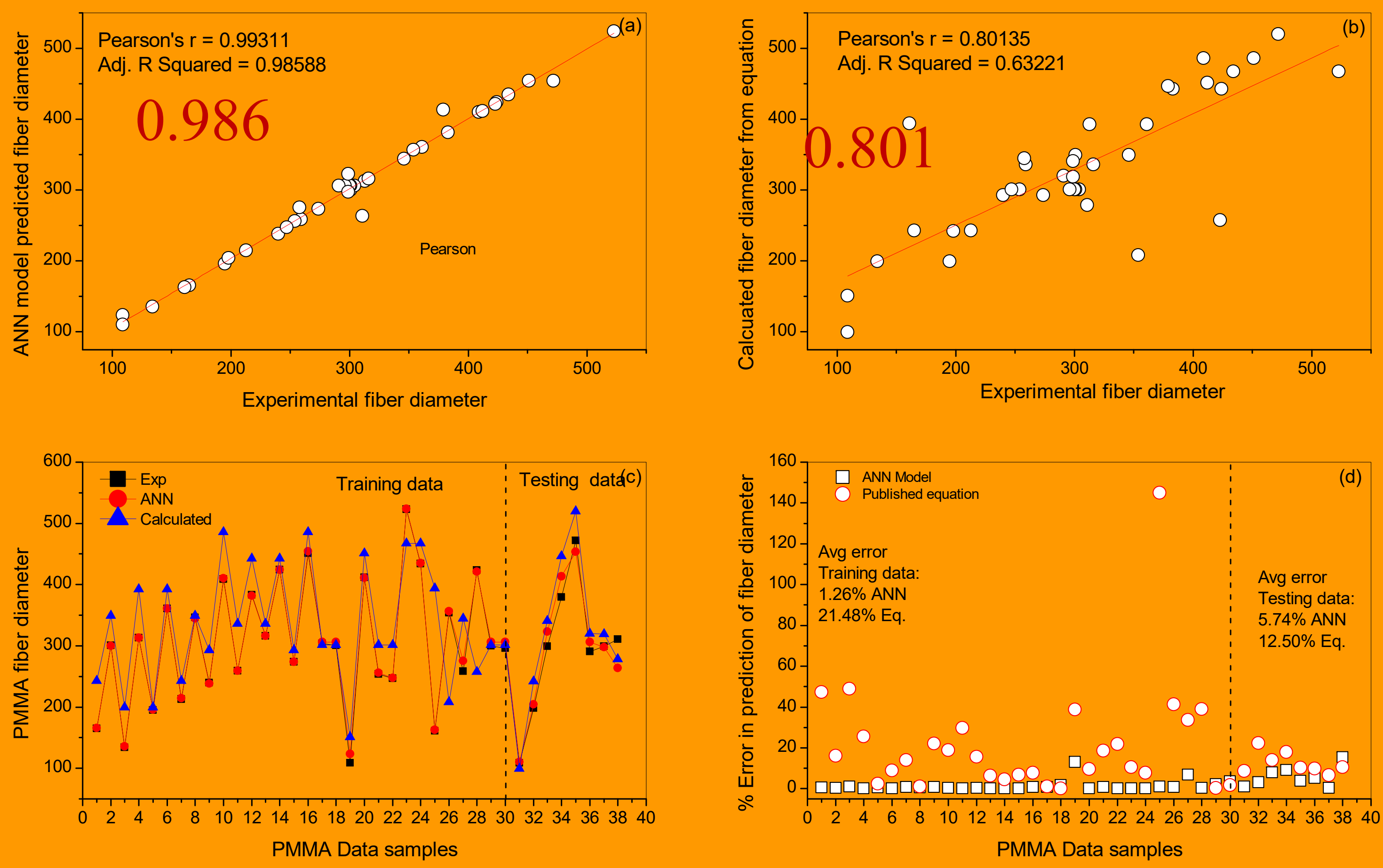
What

Modeling Electrospinning



Milan K. Sadan, Hyo-Jun Ahn, G. S. Chauhan, ***N.S. Reddy****, European Polymer Journal 74 (2016) 91.*

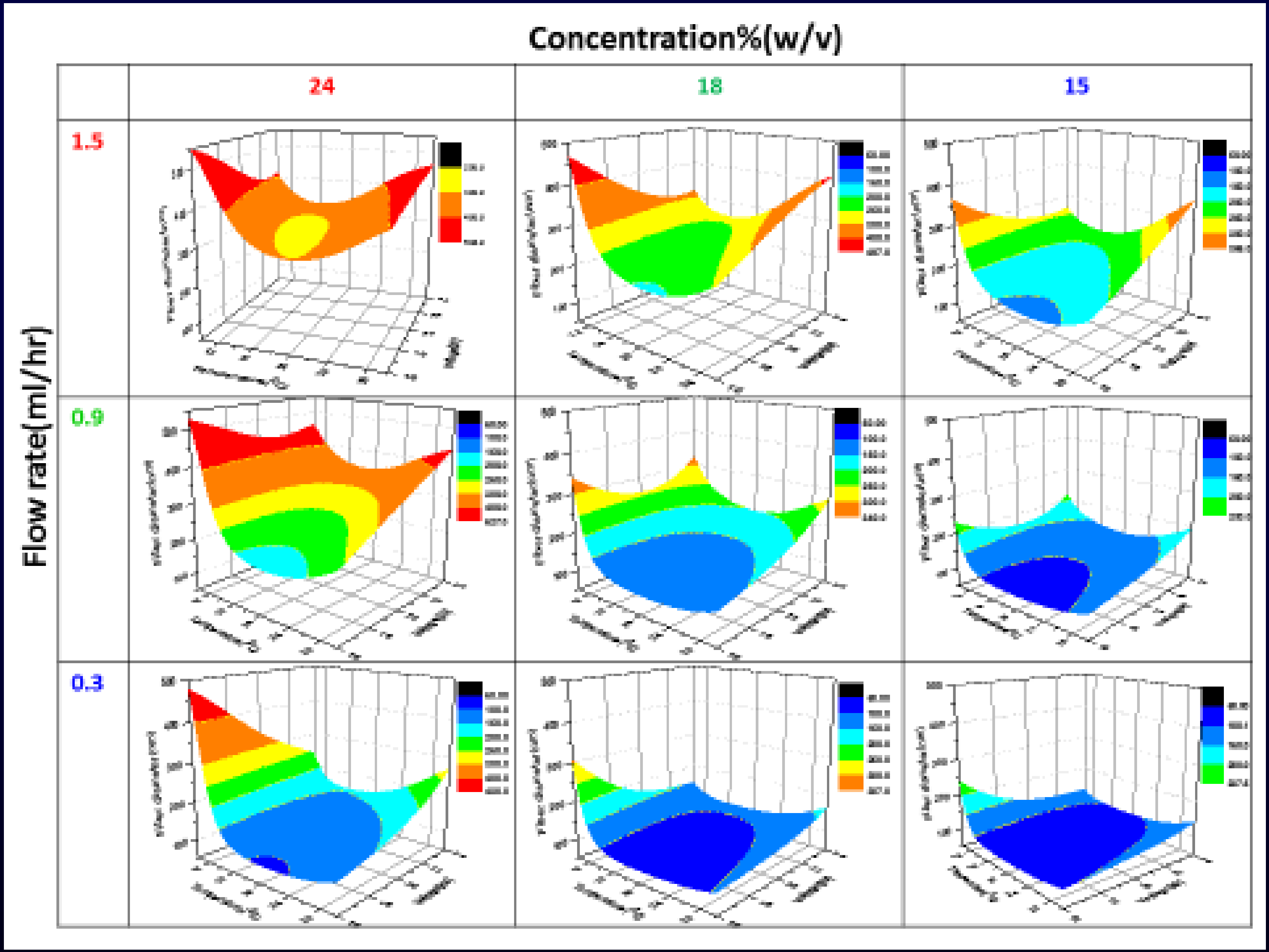
Performance of the model



*Milan K. Sadan, Hyo-Jun Ahn, G. S. Chauhan, **N.S. Reddy**, European Polymer Journal 74 (2016) 91.*

Energy

What



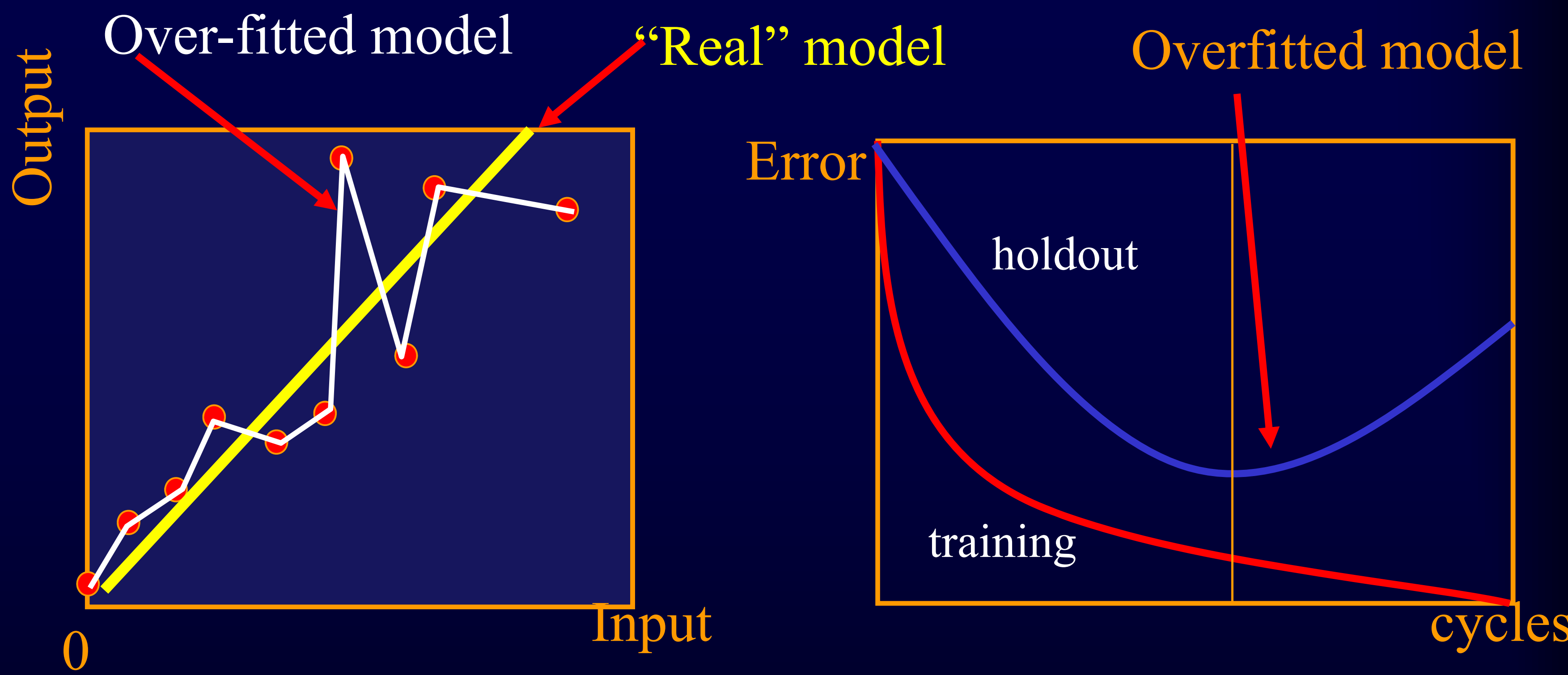
*Milan K. Sadan, Hyo-Jun Ahn, G. S. Chauhan, **N.S. Reddy**, European Polymer Journal 74 (2016) 91.*

Energy

Optimization of the parameters for required fiber dia.

What

Disadvantages of ANN modeling



Difficult to design: There are no clear design rules

Hard or impossible to train: When to stop and Over training

Difficult to evaluate internal operation: It is difficult to find out whether, and if so what tasks are performed by different parts of the net

Summary (broad)

- Ø Present model capability to map the complex nature of the metallurgical system has been demonstrated.
- Ø Sensitivity Analysis can be used to examine the effects of input variables on the output parameters, which is incredibly difficult to do experimentally.
- Ø Developed model is able to map the relation between the output parameters though this information is not fed to the model.
- Ø The present model helps in reducing the experiments required and there by saving a lot of money, material and manpower for designing the new alloys for desired properties

I would like to acknowledge

- Prof. C.S. Lee

  Department of Materials Science and Engineering, POSTECH, Korea

- Dr. Jeoung Han Kim, Dr. You Hwan Lee, and Dr. Chan Hee Park, Dr. Yeom, P L Narayana

  Former students of Department of Materials Science and Engineering, POSTECH, Korea

Artificial Neural Networks Literature

Main text books:

- "Neural Networks: A Comprehensive Foundation", S. Haykin (very good -theoretical)
- "Pattern Recognition with Neural Networks", C. Bishop (very good-more accessible)
- "Practical Neural Network Recipe's in C++"" T. Masters (emphasizing the practical aspects)

Review Articles:

- R. P. Lippman, "An introduction to Computing with Neural Nets"" IEEE ASP Magazine, 4-22, April 1987.
- A. K. Jain, J. Mao, K. Mohuiddin, "Artificial Neural Networks: A Tutorial"" IEEE Computer, March 1996' p. 31-44.
- H. K. D. H. Bhadeshia, "Neural Networks in Materials Science" ISIJ International, Vol. 39, 1999, 966-979
- H.K.D.H. Bhadeshia: *Performance of Neural Networks in Materials Science, MST*

Thank you …