Applications of Artificial Neural Networks (ANNs)

in

exploring materials property-property correlations

Xiaoyu Cheng

School of Engineering and Materials Science

Queen Mary, University of London

Submitted in partial fulfilment of the requirements of the Degree of

Doctor of Philosophy at University of London

February 2014

QUEEN MARY, UNIVERSITY OF LONDON

ABSTRACT

PHD THESIS

APPLICATIONS OF ARTIFICIAL NEURAL NETWORKS (ANNS) IN EXPLORING MATERIALS PROPERTY-PROPERTY CORRELATIONS

Xiaoyu Cheng

The discoveries of materials property-property correlations usually require prior knowledge or serendipity, the process of which can be time-consuming, costly, and labour-intensive. On the other hand, artificial neural networks (ANNs) are intelligent and scalable modelling techniques that have been used extensively to predict properties from materials' composition or processing parameters, but are seldom used in exploring materials property-property correlations. The work presented in this thesis has employed ANNs combinatorial searches to explore the correlations of different materials properties, through which, 'known' correlations are verified, and 'unknown' correlations are revealed. An evaluation criterion is proposed and demonstrated to be useful in identifying nontrivial correlations.

The work has also extended the application of ANNs in the fields of data corrections, property predictions and identifications of variables' contributions. A systematic ANN protocol has been developed and tested against the known correlating equations of elastic properties and the experimental data, and is found to be reliable and effective to correct suspect data in a complicated situation where no prior knowledge exists. Moreover, the hardness increments of pure metals due to HPT are accurately predicted from shear modulus, melting temperature and Burgers vector. The first two variables are identified to have the largest impacts on hardening. Finally, a combined ANN-SR (symbolic regression) method is proposed to yield parsimonious correlating equations by ruling out redundant variables through the partial derivatives method and the connection weight approach, which are based on the analysis of the ANNs weight vectors. By applying this method, two simple equations that are at least as accurate as other models in providing a rapid estimation of the enthalpies of vaporization for compounds are obtained.

CONTENTS

ABSTRACT ..... **i**

CONTENTS ..... **i**

LIST OF TABLE ..... **viii**

LIST OF FIGURE ..... **xii**

DECLARATION OF AUTHORSHIP ..... **xv**

ACKNOWLEDGEMENTS ..... **xvi**

DEFINITIONS AND ABBREVIATIONS ..... **xvii**

1. Introduction **1**

1.1 Aims and objectives 1

1.2 Thesis outline 3

1.3 Academic contribution 3

1.3.1 Journal article 3

1.3.2 Conference 4

1.4 Literature review 4

1.4.1 Property correlations of materials 5

1.4.2 Data mining 5

1.4.3 Artificial intelligence 6

1.4.4 Artificial neural networks (ANNs) 7

1) Introduction to ANNs 7

2) Types of ANNs 8

3) Learning rules and backpropagation algorithm 9

4) Application of ANNs 10

i. A brief history of ANNs 10

ii. General applications 11

iii. Applications in materials science 13

1.5 Capturing property correlations through a combinatorial ANN search 14

1.6 Verification of the elastic properties of the elements 16

1.6.1 Elementary definitions 19
1.6.2 Isotropic and anisotropic 21
1.6.3 The relationship between ***E***, ***G***, ***K*** and ***v*** 28
1.6.4 Static and dynamic measurements 31
1) Static approaches 31
i. Tension test 32
ii. Torsion test 33
iii. Flexure test 34
2) Dynamic approaches 36
i. Wave propagation methods 37
ii. Vibration methods 37
1.7 Capturing materials properties correlations using artificial neural networks: an example in hardening of pure metals by high pressure torsion 39
1.8 The discovery of materials properties correlations by artificial neural networks and symbolic regression 41
1.8.1 The enthalpy of vaporization 43
1.8.2 Genetic programming and symbolic regression 44
1.8.3 ANNs to identify the contributions of input variables 46
2. ANN methodology and configuration 50
2.1 Inputs and outputs of ANNs 54
2.2 Early stopping 55
2.3 Bayesian regularization 55
2.4 General evaluation criteria 56
3. Capturing property correlations through a combinatorial ANN search 58
3.1 Introduction 58
3.2 Experiment 58
3.2.1 Data collection 58
3.2.2 Input variables and output variables 59
3.2.3 Evaluation criteria 60

3.3 Results and discussion 63
3.3.1 Top binary order correlations 64
1) Cohesive energy and heat of vaporization at the normal boiling point 64
2) Atomic weight and specific heat capacity 70
3.3.2 Top ternary order correlations 73
1) Cohesive energy, boiling point and heat of vaporization 73
2) Heat of vaporization, surface energy and molar volume 76
3) Shear modulus, bulk modulus, and Poisson's ratio 78
3.3.3 Top quaternary order property correlations 81
1) Surface energy, thermal conductivity, lattice parameter ***a*** and work function 82
3.4 Conclusion 86
4. Verification of the elastic properties of the elements through ANNs 88
4.1 Introduction 88
4.2 Experiment 88
4.2.1 Data discrepancy in handbooks and databases 89
4.2.2 Data pre-treatment 89
1) Annotation removed 89
2) SI unit conversion 90
3) Data distribution information 90
4.2.3 ANN methodology 90
1) ANN constructions 94
2) ANNs simulations 96
3) Employ ANNs to verify data 104
4.3 Result and discussion 109
4.3.1 Validation of ANNs 109
1) Valid inputs for ANN constructions 109
2) Correlations captured by ANNs 114

4.3.2 Comparisons of elastic properties predicted by ANNs and the correlating equations....115
4.3.3 Original experimental value....120
1) Gd and Nd (Inconsistent ***G***)....121
2) Er and Ir (Inconsistent ***K***)....121
i. Erbium (Er)....121
i. Iridium (Ir)....122
3) Li, Pr and Tc (Inconsistent ***E*** and ***G***)....124
i. Lithium (Li)....124
ii. Praseodymium (Pr)....125
iii. Technetium (Tc)....126
4) Rh (Inconsistent ***E*** and ***K***)....127
5) Cs, Os, Re, Ru, Sc, Tm and Y (Inconsistent ***K*** and ***v***)....128
i. Cesium (Cs)....128
ii. Osmium (Os) and Ruthenium (Ru)....129
iii. Rhenium (Re)....131
iv. Scandium (Sc)....131
v. Thulium (Tm)....132
vi. Yttrium (Y)....133
6) Th (Inconsistent ***E*** and ***v***)....134
7) Ho and Zr (Inconsistent ***E***, ***K*** and ***v***)....135
i. Holmium (Ho)....135
ii. Zirconium (Zr)....136
8) Hf, Na and Pu (Inconsistent ***E***, ***G*** and ***v***)....137
i. Hafnium (Hf)....137
ii. Sodium (Na)....138
iii. Plutonium (Pu)....139
9) Ce, In, K and Tl (Inconsistent ***E***, ***G*** and ***K***)....140
i. Cerium (Ce)....140

ii. Indium (In) 141
iii. Potassium (K) 142
iv. Thallium (Tl) 143
10) Be, Cd, Eu, Ga, La, Lu, Rb, Sm, U and Yb (Inconsistent ***E***, ***G***, ***K*** and ***ν***) ..144
i. Beryllium (Be) 144
ii. Cadmium (Cd) 145
iii. Europium (Eu) 146
iv. Gallium (Ga) 147
v. Lanthanum (La) 147
vi. Lutetium (Lu) 150
vii. Rubidium (Rb) 150
viii. Samarium (Sm) 152
ix. Uranium (U) 152
x. Ytterbium (Yb) 153
4.3.4 Factors that influence elastic properties 154
1) Theoretical factors 154
2) Experimental factors 155
i. Purity 155
ii. Temperature 155
iii. Mechanical processing 156
iv. Static or dynamic measurements 156
4.4 Conclusion 157
5. Capturing materials properties correlations using artificial neural networks: an example in hardening of pure metals by high pressure torsion 159
5.1 Introduction 159
5.2 Methodology 159
5.2.1 Data collection 159
5.2.2 The inputs and output 161
5.2.3 Neural network analysis method 162

5.2.4 Knowledge extraction 163
5.3 Results and discussion 164
5.3.1 Comparison of ANN curves 164
5.3.2 Modelling with a limited data supply 169
5.3.3 Factors that affect the accuracy in the prediction 171
5.3.4 Underlying physical principle of parameters extracted by ANNs 174
5.3.5 Comparison with the physical model 174
5.3.6 The applicability of the forward selection method 176
5.4 Conclusion 176
6. The discovery of materials properties correlations through artificial neural networks and symbolic regression 178
6.1 Introduction 178
6.2 Experiment 178
6.2.1 The contribution of the different variables in ANNs 179
1) The 'PaD' method 179
2) The 'CW' method 180
6.2.2 SR modelling to obtain the mathematical expression 181
6.3 Results and discussion 182
6.3.1 ANN combinatorial search 182
6.3.2 The 'PaD' 'and 'CW' method 183
6.3.3 The important input variables - T_b, T_c and P_c 187
6.3.4 SR model analysis 189
1) SR models using all five input variables 189
2) Mathematical expression using variables selected by ANNs 191
6.4 Conclusion 193
7. General conclusion and future work 195
7.1 General conclusion 195
7.2 Original contribution of the thesis 197
7.3 Future work 199

References **201**

Appendix **231**

Appendix I 231

Appendix II 237

Appendix III 239

Appendix IV 241

Appendix V 243

Appendix VI 246

Appendix VII 248

Appendix VIII 251

Appendix IX 255

Appendix X 258

Appendix XI 261

Appendix XII 264

Appendix XIII 266

Appendix XIV 270

Appendix XV 275

Appendix XVI 276

Appendix XVII 277

Appendix XVIII 279

LIST OF TABLE

Table 1-1 Summary of stiffness matrixes with independent elastic constants, and simplified VRH formulas for the typical pure metals at room temperature..... 24

Table 1-2 Relations between the elastic properties. 30

Table 1-3 Annotations for the correlating equations listed in Table 1-2..... 30

Table 1-4 Relative merits of the dynamic and static approaches. 36

Table 3-1 The 24 properties used in the ANN combinatorial search. 60

Table 3-2 The binary and ternary order correlations between cohesive energy, boiling point and heat of vaporization. 74

Table 3-3 The binary and ternary order correlations between heat of vaporization, surface energy and molar volume. 78

Table 3-4 The binary and ternary order correlations between shear modulus, bulk modulus and Poisson's ratio. 80

Table 3-5 The binary, ternary and quaternary order correlations between work function, surface energy, thermal conductivity and lattice parameter ***a*** for the 37 elements. 86

Table 4-1 Unit conversion. 90

Table 4-2 Systematic methodology for error corrections. 91

Table 4-3 Dataset used to build the ANNs includes the consistent values of 5 elements (Co, Dy, Fe, Np, Ta and Tb) and the most common values of 22 elements (Ag, Al, Au, Ba, Bi, Ca, Cr, Cu, Mg, Mn, Mo, Nb, Ni, Pb, Pd, Pt, Sn, Sr, Ti, V, W and Zn)..... 95

Table 4-4 A summary of ANN simulations of the elastic properties correlations that are constituted by ***E***, ***G***, ***K*** and ***v***..... 96

Table 4-5 Correlation for error checking in the source pool. 104

Table 4-6 The systematic methodology to resolve the discrepancy among different ANN predictions. 107

Table 4-7 Comparison of elastic properties of 27 pre-defined consistent elements between the values used to train ANNs (denoted as Anp.), and the values calculated from single crystal elastic constants (denoted as Exp.). See text for a discussion for Mn and Ca. 110

Table 4-8 The elastic property data for the 27 pre-defined consistent elements. Values are calculated from single crystal elastic constants at 295±5 K, except for Mn and Ca. 112

Table 4-9 Comparison of the elastic property values yield by ANNs and the Equation 1-21. 117

Table 4-10 Elastic property of gadolinium (Gd) and neodymium (Nd). 122

Table 4-11 Elastic property of erbium (Er). 123

Table 4-12 Elastic property of iridium (Ir). 123

Table 4-13 Elastic property of lithium (Li). 124

Table 4-14 Elastic property of praseodymium (Pr). 125

Table 4-15 Elastic property of technetium (Tc). 126

Table 4-16 Elastic property of rhodium (Rh). 127

Table 4-17 Elastic property of cesium (Cs). 128

Table 4-18 Elastic property of osmium (Os). 130

Table 4-19 Elastic property of ruthenium (Ru). 130

Table 4-20 Elastic property of rhenium (Re). 131

Table 4-21 Elastic property of scandium (Sc). 132

Table 4-22 Elastic property of thulium (Tm). 133

Table 4-23 Elastic property of yttrium (Y). 134

Table 4-24 Elastic property of thorium (Th). 135

Table 4-25 Elastic property of holmium (Ho). 136

Table 4-26 Elastic property of zirconium (Zr). 137
Table 4-27 Elastic property of hafnium (Hf). 137
Table 4-28 Elastic property of sodium (Na). 138
Table 4-29 Elastic property of plutonium (Pu). 139
Table 4-30 Elastic property of cerium (Ce). 140
Table 4-31 Elastic property of indium (In). 141
Table 4-32 Elastic property of potassium (K). 142
Table 4-33 Elastic property of thallium (Tl). 143
Table 4-34 Elastic property of beryllium (Be). 144
Table 4-35 Elastic property of cadmium (Cd). 145
Table 4-36 Elastic property of europium (Eu). 146
Table 4-37 Elastic property of lanthanum (La). 148
Table 4-38 Elastic property of gallium (Ga). 149
Table 4-39 Elastic property of lutetium (Lu). 150
Table 4-40 Elastic property of rubidium (Rb). 151
Table 4-41 Elastic property of samarium (Sm). 152
Table 4-42 Elastic constants of uranium (U). 153
Table 4-43 Elastic property of uranium (U). 153
Table 4-44 Elastic property of ytterbium (Yb). 154
Table 5-1 Vickers hardness data of 17 elements before and after HPT ($\boldsymbol{H_{Vs}}$ and $\boldsymbol{H_{Vi}}$) and the absolute increment of hardness ($\Delta\boldsymbol{H_V}$). 160
Table 5-2 The 14 properties used in ANNs: $\Delta\boldsymbol{H_V}$ is the target output, while the input variables are chosen from the rest 13 properties. 161
Table 5-3 Comparison of 1st order ANN models and linear regression results using 13 properties. 165

Table 5-4 The 12 ANNs correlations to predict $\Delta \boldsymbol{H_V}$ using 2 properties in the forward selection results. 167

Table 5-5 Top 12 ANNs correlations to predict $\Delta \boldsymbol{H_V}$ using 3 properties in the forward selection results. 168

Table 5-6 Hardness at steady state after HPT in different sources, * indicated the value used for each element. 172

Table 6-1 ANN models in the prediction of $\boldsymbol{\Delta H_{vb}}$ using different input variables.183

Table 6-2 Comparison of the calculation accuracy of the Kistiakowsky's equation and the equation obtained by SR (indicated by *) using one input variable ($\boldsymbol{T_b}$) for 175 compounds. 192

Table 6-3 Comparison of the various equations available in literature and the equation obtained by SR (indicated by *) using three input variables, i.e. $\boldsymbol{T_b}$, $\boldsymbol{T_c}$ and $\boldsymbol{P_c}$. 193

LIST OF FIGURE

Figure 1-1 Schematic arrangement of causation in materials science. 6

Figure 1-2 An illustration of how an artificial neuron functions. 8

Figure 1-3 A typical three-layer feed-forward neural network with the back propagation learning method..... 10

Figure 1-4 Element subjected to shear and associated complementary shear. 28

Figure 1-5 Direct stresses due to shear. 29

Figure 1-6 Classification of the methodologies employed to determine the elastic properties of solids. 32

Figure 1-7 a) Schematic representation of the apparatus used to conduct tensile stress–strain tests; b) A typical standard specimen; and c) A typical stress-strain curve. 33

Figure 1-8 Schematic representation of torsional deformation..... 34

Figure 1-9 Schematic diagrams of (top left) four-point flexure using machine displacement measurement, (top right) strain gauges applied to both sides of the test-piece, and (lower) direct measurement of deflection between fixed points on the test-piece surface using a single linear displacement transducer. 35

Figure 1-10 Three vibration modes using electromagnetic-acoustic transducers: a) longitudinal, b) torsional, and c) flexural. 38

Figure 1-11 Block diagram of a typical thread suspension flexural vibration test apparatus. 38

Figure 1-12 The principle of high pressure torsion. 40

Figure 1-13 Flowchart of the genetic programming paradigm. 47

Figure 1-14 A tree-like hierarchically structured computer program corresponding to the equation: (y/x)*y-y-cos(x+y)..... 48

Figure 1-15 Illustration of GP operations: a) reproduction, b) mutation, and c) crossover. 49

Figure 1-16 Major preparatory steps of the basic GP process. 49

Figure 2-1 A flowchart showing how the desired Neural network with one hidden layer was achieved.. 53

Figure 2-2 Result of ANNs in predicting the hardness increment due to HPT from shear modulus and melting temperature for 17 elements..... 57

Figure 3-1 An example of a ternary order correlation-group that is constituted by three different correlations utilizing the property X, Y and Z.. 62

Figure 3-2 The correlation between the cohesive energy of the solid (E_{coh}) and the heat of vaporization at normal boiling point (ΔH_{vb}) obtained by the ANN combinatorial search for 37 metallic elements: a) Predict ΔH_{vb} from E_{coh}; b) Predict E_{coh} from ΔH_{vb}. 69

Figure 3-3 Linear correlation between the cohesive energy of the solid (E_{coh}) and the heat of vaporization at normal boiling point (ΔH_{vb}). 70

Figure 3-4 The correlation between the atomic weight (A_r) and specific heat capacity (C_p) obtained by the ANN combinatorial search for 37 elements: a) Predict A_r from C_p; b) Predict C_p from A_r. 71

Figure 3-5 The correlation between cohesive energy, heat of vaporization, and boiling point, which is obtained through the ANN combinatorial search: a) Predict E_{coh} from ΔH_{vb} and T_b; b) Predict ΔH_{vb} from E_{coh} and T_b; c) Predict T_b from E_{coh} and ΔH_{vb}..... 76

Figure 3-6 Plots of the atomic number of a metal versus its: a) work function, b) surface energy, c) thermal conductivity, and d) lattice parameter a. 84

Figure 4-1 ANN correlations of property group (E, K and ν): a) Predict ν from E and K, b) Predict K from E and ν, c) Predict E from K and ν. 99

Figure 4-2 ANN correlations of property group (G, K and ν): a) Predict ν from K and G, b) Predict G from K and ν, c) Predict K from G and ν. 101

Figure 4-3 ANN correlations of property group (***G***, ***E*** and ***ν***): a) Predict ***G*** from ***E*** and ***ν***, b) Predict ***E*** from ***G*** and ***ν***. 102

Figure 4-4 ANN correlations of property group (***G***, ***E*** and ***K***): a) Predict ***K*** from ***E*** and ***G***, b) Predict ***E*** from ***K*** and ***G***, c) Predict ***G*** from ***E*** and ***K***. 103

Figure 4-5 Comparison of the ANN model and the correlating equation in predicting ***E*** from ***ν*** and **G** in 3D dimensions. 115

Figure 5-1 The result of ANNs in predicting ΔH_V from ***G*** and T_m. 164

Figure 5-2 Result of ANNs in predicting ΔH_V from ***G***, T_m and ***b***. 166

Figure 5-3 The property importance revealed in the top 12 correlations of 2nd and 3rd order in predicting ΔH_V from the ANN combinatorial search..... 169

Figure 5-4 Modelling uncertainty: predicting ΔH_V from ***G*** and T_m from 30 independent ANNs. 171

Figure 6-1 The property importance revealed in the top 20% correlations of 1st, 2nd, 3rd and 4th order in predicting ΔH_{vb} from the ANN combinatorial search 185

Figure 6-2 Results of ANNs in predicting ΔH_{vb} from a) T_b; b) T_b, T_c, P_c; and c) M_w, T_b, T_c. 187

Figure 6-3 The contributions of the 5 variables accessed by the 'PaD' method and the 'CW' method in the prediction of ΔH_{vb}. 188

Figure 6-4 The contributions of the 5 variables accessed by Discipulus (SR method). 191

Figure A-1 Ternary order correlations between elastic properties are reflected in 3D dimensions.. 284

DECLARATION OF AUTHORSHIP

I, Xiaoyu Cheng

declare that the thesis entitled

APPLICATIONS OF ARTIFICIAL NEURAL NETWORKS (ANNS) IN EXPLORING MATERIALS PROPERTY-PROPERTY CORRELATIONS

and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

Signed: ……..

Date:…….

ACKNOWLEDGEMENTS

I would like to thank Dr. Shoufeng Yang for his extraordinary supervision and generous help in the development of this PhD research.

Special thanks to Professor Marco Starink, who suggested me to study variable contributions that would benefit to the construction of a physical model. He opens a new idea of employing ANNs in the field of high pressure torsion. Thanks are extended to Professor Mike Reece, Dr. Asa Barber, and the research administrator Jonathon Hills, for their kindness help and support.

I am heartily thankful to my wonderful friends, namely, Yiming Zhang, Ying Chen, Xiaomei Ren and Binyan He, for sharing good times and dividing pain.

My deepest gratitude goes to my parents for their caring, loving and understanding.

Finally, I am grateful to the School of Engineering and Materials Science of Queen Mary, University of London and China Scholarship Council (CSC), for providing Queen Mary - China Scholarship Council Joint PhD Scholarship for me. I would like to thank the Faculty of Engineering and the Environment of University of Southampton for supporting me as a visiting research student since 2010.

DEFINITIONS AND ABBREVIATIONS

a	Lattice parameter, a
A	Prediction accuracy
AI	Artificial intelligence
A_n	Atomic number
ANNs	Artificial neural networks
A_r	Atomic weight
b	Burgers vector
bcc	Body-centred cubic
BE/A	Binding energy per nucleon
b_j	The amount of external/environmental input to neuron j
BPANNs	Feed forward backpropagation neural networks
b_s	Bias
b_t	Test-piece width
C	Elastic constants matrix
CES	A database named as CES EduPack
C_{ij}	Elastic constants /stiffness
C_m	Molar heat capacity
C_p	Specific heat capacity measured at constant pressure
C_v	Specific heat capacity measured at constant volume
CW	The Connection Weight approach
d_i	The partial derivatives of the output Y with respect to the input X_i
DM	Data mining
D_m	Dipole moment
$Dminmaxp$	(Max-min)/min
E	Young's modulus
E_c	Correlation error
E_{coh}	Cohesive energy
E_D	The sum of squared errors between the network predicted value and the target output
E_w	The sum of squares of the network weights
F	Load

fcc	Face-centred cubic
G	Shear modulus
GP	Genetic programming
h	Test-piece thickness
H	The maximum number of neurons in hidden layer
hcp	Hexagonal close-packed
H_{fus}	Heat of fusion
HPT	High pressure torsion
H_{Vi}	Vickers hardness at annealed state
H_{Vs}	Vickers hardness at the saturated level
H°_{T}	The standard-state enthalpy at temperature ***t***
I	The ranking index
IW	The weight vector between the input layer and the hidden layer
K	Bulk modulus
L	Length
LW	The weight vector between the hidden layer and the output layer,
MLP	Multilayer perceptron network
MSE	Mean squared error
M_w	Molecular weight
N_{hid}	The number of neurons in the hidden layer
N_{inp}	The number of input variables
N_{out}	The number of output parameters
PaD	The partial derivatives method
P_c	Critical pressure
Q	The number of ANNs training
Q_h	Heat flux
r	Atomic radius
R	Gas constant, 8.314 J mol^{-1} K^{-1} (1.986 cal K^{-1} mol^{-1})
RBF	Radial basis function network
R_l	Regression coefficient
RMSE	Root mean square of errors
S	Elastic coefficient /compliance matrix
S_l	The slope of the linear regression line

SOM	Self-organizing map
SR	Symbolic regression
T	Temperature
T_b	Boiling point
T_{br}	Reduced boiling point which equals to T_b/T_c
T_c	Critical point
T_m	Melting point
T_r	T-dependence of resistivity
T_t	The applied torque
V_a	Atomic volume
V_L	The longitudinal velocities
V_m	Molar volume
$Vmax$	The maximum value
$Vmean$	The mean value
$Vmedian$	The median value
$Vmin$	The minimum value
$Vmode$	The value that has the highest occurrence
VRH	Voigt-Reuss-Hill approximation method
V_S	The shear velocities
$Vvar$	The variance
W	Work function
W_t	Weights
X_i	The output of neuron i and the input signal for neuron j;
Z_o	The optimal number of neurons in hidden layer
α_L	Thermal expansion coefficient
γ	Shear strain
δ	ANN model generalization ability
ΔE	The internal energy
ΔH_V	Vickers hardness increment after HPT
ΔH_{vap}	The enthalpy of vaporization
ΔH_{vb}	Heat of vaporization at the normal boiling point
$\Delta V/V$	Volumetric strain
ε	Tensile or compression strain

θ	The angle
θ_D	Debye temperature
λ	Thermal conductivity
ν	Poisson's ratio
ρ	Density
ρ_e	Electrical resistivity
σ	Tensile or compression stress
σ_A	Neutron absorption cross section (0.025 ev)
σ_m	Hydrostatic stress
σ_S	Neutron scattering cross section (0.025 ev)
τ	Shear stress
Υ	Surface energy (liquid)
χ	Compressibility
χ_m	Magnetic susceptibility
ψ	The evaluating parameter of ANN modelling performance
ω_{ij}	The weight assigned to the connection of neuron i to neuron j;

1. Introduction

1.1 Aims and objectives

Materials properties correlations are compact summaries that provide "a ready means of access to the information represented by a body of discrete values or the equivalent" [1]. Correlations can be employed to generate new information, and are very useful for calculations involving materials properties [2]. The advances in computing power, coupled with the computational modelling, namely, artificial neural networks (ANNs), and the readily available materials properties databases, enable the author to apply this modelling approach, which is efficient in terms of both time and cost, to solve diverse problems by exploring materials property-property correlations from databases.

Artificial neural networks are currently one of the most powerful data mining techniques that have been widely applied in many fields, including marketing strategy, chemistry, biology, materials science, and pattern recognition [3-8]. The general aim of this work is to use ANNs to explore correlations that might exist between different properties in materials, and employ such correlations to solve problems that are hardly accessible to conventional methods. Because correlating equations may be of greater interest as a source of property correlation information than graphical correlations or digital models [1], there is also a need to interpret the correlations captured by ANNs into explicit correlating equations.

Thus, the general purpose of this work is to capture materials property-property correlations and apply such correlations to: i) the prediction of materials properties, ii) the detection of errors in materials properties databases, and iii) the identification of important input variables for the establishing of physical models and the construction of explicit correlating equations. In total, four distinct examples of ANNs applications are presented.

The above aim implies the following specific objectives:

1) To demonstrate that ANNs are capable of capturing meaningful property-property correlations without any prior knowledge or any assumptions of the form of the relationship made in advance.

2) To test whether the ANN combinatorial search is feasible to explore property-property correlations from a discrete, irregularly distributed database that is subject to unknown error.
3) To model the binary, ternary, and quaternary order correlations between 24 properties of 37 metals through an ANN combinatorial search, and provide an evaluation criterion to rank the correlation based on the model performance, a factor that implies how strong the correlation is.
4) To illustrate that the ANN combinatorial search is a way of illuminating the facts by analysing the typical types of binary, ternary and quaternary order correlations captured in the search.
5) To build the most reliable database of Young's modulus, shear modulus, bulk modulus, and Poisson's ratio for 68 pure metallic elements, because large discrepancies exist in literature.
6) To extend the application of ANNs in detecting and correcting errors in handbooks and databases, and access the reliability of the method by comparing the results that are obtained from ANNs with the results that are generated by the physically established correlating equations, and the experimental values as well.
7) To employ the ANN method in a situation where the size of data set has an inherent limitation. For example, only a relatively small number of pure metals have been processed through high pressure torsion.
8) To predict the increments of Vickers hardness of pure metals due to high pressure torsion, and identify the properties that contribute most to the changes of hardness through two ANN approaches: the combinatorial search and the forward selection.
9) To develop correlating equations through a combined approach of ANNs and symbolic regression (SR). This is exemplified by determining the enthalpies of vaporization for 175 compounds.
10) To perform a comparison study of the Partial Derivatives method and the Connection Weight approach in accessing the contributions of input variables. Both methods utilize ANNs weight vectors.

1.2 Thesis outline

The thesis is divided into seven chapters. Chapter 1 describes the motivation for the work performed in this thesis. It also includes a literature review of ANNs and an introduction for each application of ANNs in the exploration of materials property-property correlations. The methodology used in the design of ANNs is discussed in Chapter 2. In Chapter 3, an ANN combinatorial search method is employed to capture property correlations hidden in the database. Chapter 4 presents a method of employing ANNs to detect and correct the errors of elastic properties of elements. Chapter 5 uses an example in the hardening of pure metals by HPT to describe a general method to identify important variables based on the correlations captured by ANNs. In Chapter 6, variable contributions are identified by the Partial Derivatives method and the Connection Weight approach. Thermal property-property correlations are not only captured but also obtained in analytic forms through a combined ANN-SR method. Chapter 7 summarises the main conclusions and presents an outlook for future work. Appendixes listed in the final chapter give extra information including all the numeric data used in the modelling.

1.3 Academic contribution

1.3.1 Journal article

So far, parts of this work have been published or to be submitted as:

1. Marco Starink, **Xiaoyu Cheng**, Shoufeng Yang (2012) Hardening of pure metals by high pressure torsion: a physically-based model. Acta Materialia 61(1), 183-192.
2. **Xiaoyu Cheng**, Marco Starink, Shoufeng Yang (2014) Capturing materials properties correlations using artificial neural networks: an example in hardening of pure metals by high pressure torsion. *Submitted to Acta Materialia.*
3. **Xiaoyu Cheng**, Shoufeng Yang (2014) Selected values of the elastic properties for elements. *In preparation*.
4. **Xiaoyu Cheng**, Shoufeng Yang (2014). The discovery of materials properties correlations through artificial neural networks and symbolic regression. *To be submitted*.

5. **Xiaoyu Cheng**, Shoufeng Yang (2014) Capturing materials properties correlations through the combinatorial neural network search. *In preparation*.

1.3.2 Conference

Xiaoyu Cheng. An application of artificial neural networks in finding elastic properties of rare earth elements. The 3rd International Symposium on Rare Earth Resource Utilization (ISRERU-3) & The 3rd Special Symposium on Advances in Functional Materials (AFM-3). Changchun, China, 9-13 December 2012. **Best oral presentation award**.

Xiaoyu Cheng. Capturing materials properties correlations using ANNs: an example in hardening of pure metals by HPT [Poster]. Queen Mary, University of London, SEMS Graduate Research Show. London, 14 November 2012.

Xiaoyu Cheng. Selected values of the boiling points and the enthalpies of vaporization of the elements [Poster]. University of Southampton Postgraduate Conference 2011. Hilton hotel, Southampton, 28 September 2011.

Xiaoyu Cheng. The discovery of materials properties correlations through artificial neural networks and genetic programming: an example in determining the enthalpy of vaporization. The 17th Joint Annual Conference of CSCST & SCI. Oxford University, UK, 17 October 2010.

1.4 Literature review

Materials Science and Engineering is a study that investigates the relationships that exist between the composition, structure, processing, and properties of materials [9]. For many years, the methodological framework for materials science has been following the composition-processing-structure-property causal pathway [10]. Material property values are usually obtained via a huge amount of experiments based on the understanding of materials' composition, processing, and structures. Advances in modern data analytic tools and statistical algorithms allow materials properties to be estimated by employing property interrelationships.

1.4.1 Property correlations of materials

As Ashby points out, material properties "derive ultimately from the way in which the atoms or molecules are arranged in space (structure) and the nature of the intermolecular forces that hold them together (bonding)" [2]. Therefore, the properties of a material are correlated in varying degrees, see Figure 1-1 [10]. The property-property correlations are strong when properties derive directly from the nature of the atomic bonding and structure, and are relatively weak when defects in the structure are introduced or environmental interactions are involved [2]. Moore and Notz [11] summarised the correlations between two variables into three types: 1) a direct causation, 2) common response, and 3) confounding. Such summary still holds good for correlations of higher order (i.e. correlations that are constituted with more variables).

Compared to the study of structure-property relationships, which has many successes in materials science, the idea of exploring property-property correlations is less unconventional for binary order correlations, especially when both properties arise from the same atomic structural level. The relationship between thermal conductivity and electric conductivity for materials with more or less freely moving electrons is a well-known example. A number of binary order correlations (for instance, the correlation between specific heat and atomic mass, and the correlation between boiling point and latent heat of vaporization) have been established, partially because the relationships can be easily obtained from conventional, simple regressions. However, complicated correlations between several numbers of properties require a scalable and sustainable predictive modelling technique, e.g. data mining.

1.4.2 Data mining

To deal with an information poor but data rich situation [12], and provide solutions where no practical experience exists, data mining (DM) is emerging as an efficient approach to derive implicit information and knowledge from databases. It is particularly useful for incomplete, noisy and fuzzy data [13]. DM has become an important research area that draws wide attention [14-16].

The underlying knowledge discovered by data mining is potentially useful information (such as knowledge rules, constrains, and regularities) [17], corresponding to general or

domain-specific problems [18]. Trying to interpret such knowledge into a solid physical model may not always be possible. However, it is still of great practical value to extract novel or interesting information on the premise of no clear assumption, as opposed to the requirement of traditional statistical linear approach [19]. Predictions made from interpolations or extrapolations of the captured correlations are expected to be reliable if similar patterns always exist in the observed data.

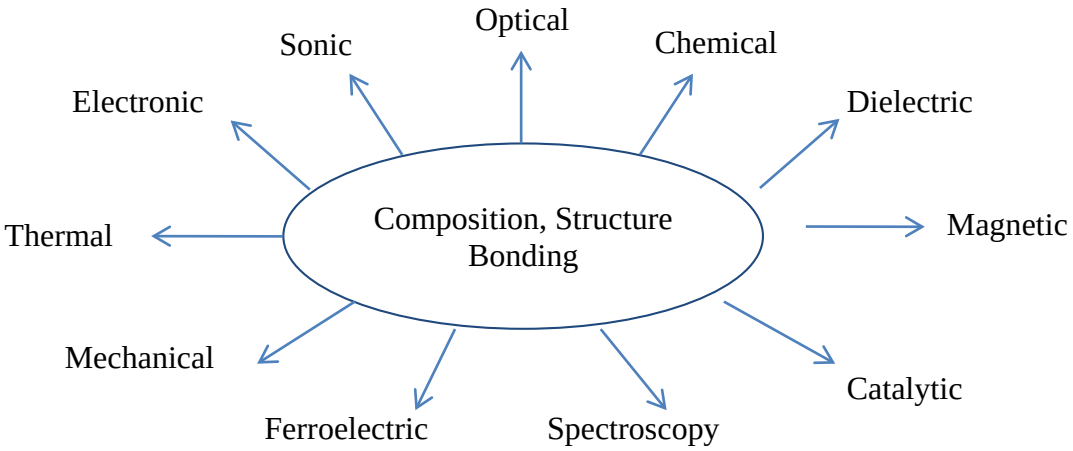


Figure 1-1 Schematic arrangement of causation in materials science (from Ref. [10]).

1.4.3 Artificial intelligence

The most influential concept of artificial intelligence (AI) was described as the famous Turing's problem [20]. Though the question "Can machines think" is much debated, AI has been gradually accepted as a reliable new technology that helps human beings to better understand the world through the use of computational models [21]. AI has been used in a wide range of fields including medical diagnosis [22], bioinformatics [23], manufacturing [24], property evaluation [25], modelling [26], process controlling [27], and scientific discovery [28].

Artificial neural networks (ANNs) are generally considered as one of the most efficient AI approaches to incorporate and process qualitative knowledge derived from data [21]. Despite the fact that ANNs and AI share the same goal of simulating human intelligence, some researchers argue that ANNs are different from AI approaches, because the ANN

methods apply inductive reasoning while the AI approaches refer to deductive reasoning [29].

1.4.4 Artificial neural networks (ANNs)

1) Introduction to ANNs

Artificial neural networks (ANNs) attempt to simulate biological brain functions to perform parallel computations for data processing and knowledge extraction [29]. A typical structure of ANN comprises:

a) Interconnected adaptive artificial neurons, which sometimes are referred as processing elements or nodes. A biased neuron, which is an additional input with constant value normally assumed equal to one, may be introduced in a ANN model to shift the threshold of the activation function [30];

b) Layers. An ANN model normally includes an input layer, an output layer, and additional hidden layers. Each layer contains a number of neurons;

c) Weights (value between 0 and 1), which are assigned to every connection of neurons.

Figure 1-2 illustrates a generic artificial neuron with connections. The net input to the artificial neuron j is shown in Equation 1-1 [29, 31, 32]. The output of neuron j ($\boldsymbol{Y_j}$) is computed through an activation function $f(\cdot)$ before it is sent to other neurons. Such activation function could be the hyperbolic tangent sigmoid transfer function ($\frac{2}{1+e^{-2net_j}}-1$) [33], the logistic sigmoid function ($\frac{1}{1+e^{-net_j}}$) [34], the piecewise function or the linear function.

$$net_j = \sum_{i=1}^{N} \omega_{ij} X_i + b_j \qquad \textit{Equation 1-1}$$

where:

$\boldsymbol{\omega_{ij}}$ — the weight assigned to the connection of neuron i to neuron j;

$\boldsymbol{X_i}$ — the output of neuron i and the input signal for neuron j;

b_j — the amount of external/environmental input to neuron j, which is also known as the weight value of bias neuron.

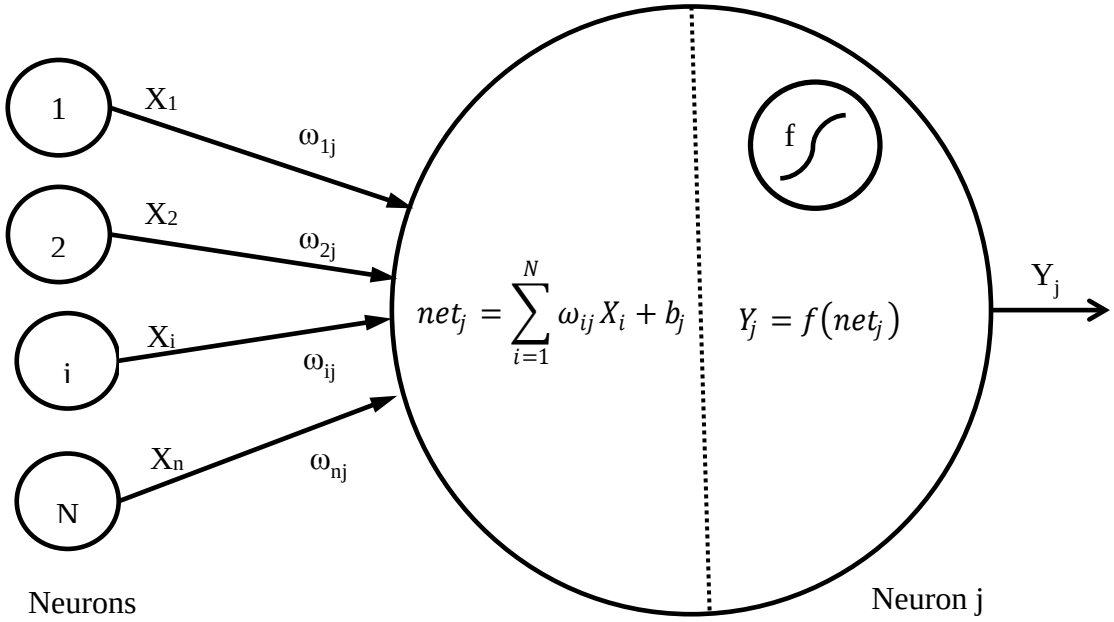


Figure 1-2 An illustration of how an artificial neuron functions (from Ref. [31]).

2) Types of ANNs

The frequently used types of ANNs are discussed as below:

1. Multilayer perceptron network (MLP): as illustrated in Figure 1-3, it is usually a feed-forward neural network comprising an input layer, one or more hidden layers and an output layer. In MLP, the connection goes through the neurons in each layer sequentially. Usually, in this kind of network, the supervised learning process is the backpropagation algorithm. MLP is versatile and can be used in many fields such as data modelling, classification, forecasting, control, image compression and pattern recognition [34].
2. Radial basis function network (RBF): it consists of an input layer, one hidden and an output layer, and uses a Gaussian kernel function to calculate the activations of the neurons. It is a special case of MLP network, which trains faster but is less versatile [35].
3. Hopfield network: it is a fully connected single layer network that acts as a nonlinear associative memory. After an input pattern is presented, the network

will converge by means of a state update rule to a stable pattern [36]. A Hopfield network is especially efficient in solving optimisation problems [37].

4. Self-organizing map (SOM): it is an unsupervised training technique based on competitive learning. Instead of mapping the input pattern to a target output, such network learns to form its own classification of training data by identifying the common features shared by the input patterns. Kohonen network is one particular kind of SOM [38].

3) Learning rules and backpropagation algorithm

ANNs are trained to modify the weights and biases to perform a certain task through supervised learning, unsupervised learning or reinforcement learning [39]. In supervised learning, both the inputs and the corresponding target outputs are provided, and weights and biases are adjusted to make the network outputs closer to the target outputs. For an unsupervised learning, no training sets (inputs and corresponding target outputs) are provided. ANNs learn to discover common features in input patterns and classify input data into appropriate categories without feedback provided from the environment to improve the mapping [29]. When a competitive learning rule is used, the output neurons compete amongst themselves to be activated and only the winner produces an output signal. In contrast to supervised learning and unsupervised learning, evaluative feedback ('right' or 'wrong') is given as reinforcement signal in reinforcement learning, and ANNs are modified to maximize the reward (feedback is 'right') [39].

In the present work, both the inputs and the corresponding target outputs are provided, so supervised learning rule is used. One of the most popular methods for performing supervised learning is the backpropagation algorithm (BP), which involves calculating errors backward through the network [40]. In feed forward backpropagation neural networks (BPANNs), a random set of initial weights is assigned to start mapping the input to the corresponding target output in a chosen input-output pair. With the fixed threshold units and activation function, neuron activation is propagating from the input layer through the hidden layers and then passing on to the output layer, while errors are computed and weights for the neurons are adjusted from the output layer to each hidden layer accordingly. As shown in Figure 1-3, this process is repeated until an epoch is completed, when all the input-output pairs have been presented to the network. The

network will stop learning if the desired number of epochs has reached or the desired level of output accuracy has been obtained.

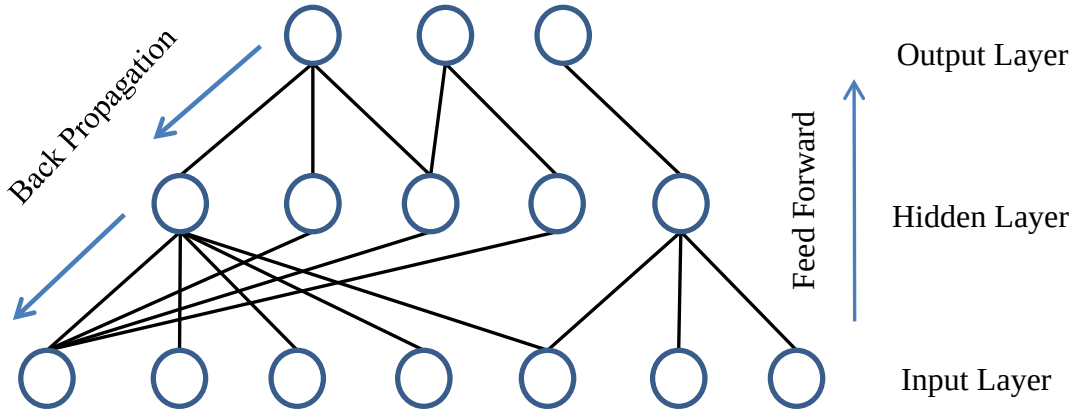


Figure 1-3 A typical three-layer feed-forward neural network with the back propagation learning method (from Ref. [29]).

4) Application of ANNs

Phenomenological models and empirical models are the two dominant analysis classes to characterise the properties of a given material system: in phenomenological models, the quantitative relationships between property variables are derived from physical theories, while in most empirical models the relationships are built on a priori postulate [41]. Compared to phenomenological models, which have complexities in designing and difficulties in generalizing, empirical models have more applications that are extensively used in practice. ANNs are one of the most wildly used empirical models emerged in recent years that can substantially reduce experimental testing needed in property evaluation and provide relative reliable data without having to perform an actual experimental test [42]. The great predictive capability of ANNs, especially for non-linear problems, is generally considered to surpass the performance of traditional statistical regression models [43].

i. A brief history of ANNs

The earliest work on neural networks started in 1943 when McCulloch and Pitts [44] modeled the activity of biological neurons as a logic circuit. In 1958, a single layer

network, i.e. the perceptron, was developed by Rosenblatt [45], which showed initial successes in learning certain classifications by using trial-and-error method to randomly adjust connection weights. Such weight changing method is later modified into randomly choosing a direction vector, a process became known as descending on the gradient [46], and then was improved to be a gradient search method based on minimizing the squared error [47], i.e. Least Mean Squares. However, in 1969, the perceptron was demonstrated to be not able to solve non-linear classification problems, such as exclusive-or (XOR) logic [48]. Such limitation led to the decline of research in the field of neural networks until the 1980s that backpropagation algorithm using a differentiable sigmoidal function rather than signum function as the activation function for multi-layered networks, was introduced to solve this problem [49]. New ANN models that were developed by Grossberg [50], Kohonen [38], Klopf [51], and Hopfield [37], have also been applied to solve real life problems.

In 1985, the first annual meeting on Neural Networks for Computing was hosted by the American Institute of Physics, and in 1987 the Institute of Electrical and Electronic Engineers (IEEE) first International Conference on Neural Networks drew the world's attention [52]. Ever since then, The International Conference on Artificial Intelligence and Neural Networks and The International Joint Conference on Neural Networks have become the two premier conferences for the presentation of new advances and research results. To date, researches in the field of ANNs continue to show the great advantages for pattern classification, clustering, function approximation, prediction, optimization and non-linear controlling [53, 54].

ii. General applications

The general ANNs applications can be grouped into the following categories:

Pattern recognition

As summarized by Bishop [54], there are two pattern recognition problems: regression and classification. The first one is concerned with predicting the output values from the input variables (prediction/forecasting) with a suitable set of functions found by ANNs (function approximation), while the second one is to assign input patterns to a set of discrete classes.

Solving regression problems is possibly the most popular application of ANNs. Because ANNs are quantitative treatments that can rapidly approximate non-linear functions, it is appropriate to use ANNs to model a large quantity of data that broadly distribute over a wide range. ANNs are particularly useful when the relationship between input and target output is nonlinear and so complex that not yet accessible to physical modelling [55]. Unlike traditional statistical forecasting tools, i.e. linear least squares models, ANNs can be applied without any assumption on the forms of relationships made in advance.

For classification problems, if the classes in the training set are pre-specified, it is referred as pattern classification; if no class labels are assigned to the training set, it is referred as unsupervised pattern classification, i.e. clustering [53]. Important applications of classification problem are including using SOM or RBF for speech and text recognition [56, 57], face recognition [58, 59], structure predicting [6] and image processing [60].

Control

In addition to the wide applications in data analysis, ANNs are considerably promising techniques in non-linear control. Surveys on the applications of a variety of neural network architectures in control have been proposed by Hunt *et al.* [61] and Hagan *et al.* [62]. The applications of ANNs in control systems include: a) real time positional control of welding equipment [63]; b) high precision motion control in the presence of large friction [64]; c) springback control in an air bending process [65]; d) automated control of an induction hardening process [66]; e) laser surface strengthening parameters control [67]; f) damping force control of magnetorheological fluid dampers [68]; g) surface roughness control [69]; h) drug release control from matrix tablets [70]; and i) steel strip deviation control in electro-hydraulic servo system [71].

Optimisation

ANNs can also be applied to solve optimisation problems by finding a solution that satisfies a set of constraints in order to maximize or minimize an objective function, for example, the travelling salesman problem [53]. To deal with real-world problems, Rodemerck *et al.* [72] used ANNs to search for new catalytic composition. Song and Zhang [73] used ANNs to search the optimum heat treatment technique. Ootao *et al.* [74]

used ANNs to find the optimal material composition that minimize thermal stress distribution under thermal loading. Recently, ANNs have also been employed by Somashekhar *et al.* [75] to optimize material removal rate in a micro-electric discharge machining process, by Sun *et al.* [76] to predict flow stress of alloys during hot deformation, and by Elsayed [77] to model the pressure drop and the cut-off diameter for cyclone separators.

iii. Applications in materials science

ANNs have been used extensively in materials science with varying success [4, 5, 78, 79], but most applications fall within the scope of establishing composition-processing-structure-property-performance relationships. In most cases, materials compositional information and processing parameters are used as the main input variables to produce a desired result, although other descriptors sometimes are added to the models to help improve performance. For example, ANNs developed by Qian *et al.* [80] to predict tensile strength based on materials' composition and microstructures, showed better precision and generalization ability than multiple statistical analysis. A critical review carried out by Zhang and Friedrich [81] suggests the ANN approach is a promising technique in predicting mechanical properties of polymer composites from materials composition. It is demonstrated by Abendroth and Kuna [82] that ANNs are suitable for the determination of load displacement of small punch test from the geometry of specimens and the loading parameters. Accurate predictions of the relative permittivity and oxygen diffusion properties of ceramic materials were obtained by Scott *et al.* [83], using compositional information as the core of the ANN input data. The ANN models built by Wen *et al.* [84] successfully captured the relationship between processing parameters and the bending strength and microhardness of compound materials. The potential of using ANNs to predict materials tribological properties as a function of materials' composition and testing conditions was first investigated by Jones *et al.* [42], and extended by a number of researchers [85-96].

Compared to the wide applications of ANNs utilizing compositional information and processing parameters as inputs, the number of studies in exploring materials property-property correlations (excluding compositional information and processing parameters) has been very limited so far. Homer *et al.* [97] used physicochemical properties to

predict viscosity, density, heat of vaporization and boiling point for pure organic liquid hydrocarbons. Zhang *et al.* [8, 98] used ANNs to predict solid solubility limits in metallic systems with inputs including the atomic size parameter, the valence parameter, the electrochemical parameter and the structure parameter. They also derived the correlation between heat of vaporization and boiling point from ANNs, which is useful in correcting errors in handbooks [7].

1.5 Capturing property correlations through a combinatorial ANN search

Combinatorial approaches are high-throughput screening methods that systematically investigate material composition-structure-property relationships [99]. It significantly reduces experimental time at a relatively lower cost, compared to traditional trial-and-error approaches which normally test one sample at a time [100]. The modern idea of combinatorial approach was introduced in 1960s by Kennedy *et al.* [101] to determine isothermal sections of ternary-alloy diagrams using samples with composition gradients. It was further developed by Hanak *et al.* [100] to include high-throughput property screening to find superconductors. The development of the combinatorial approach has attracted tremendous interests in the search for new materials or material property optimizations [102, 103].

Although there is some scepticism criticizing the combinatorial method as "an engineering shortcut, a route to invention that circumvents true understanding" [104], the advent of computation power has further accelerated the applications of combinatorial approach. The synthesis and analysis of a large number materials that either has continuous composition gradients or discrete composition 'libraries' [105], are extended from the studies of functional materials, such as superconductors [100] and catalysts [102], to the fields of pharmaceutical and biotechnology [104].

A large number of studies on the topic of combinatorial approach have become available. Amis *et al.* [104] ascribed the prevalence of combinatorial approach in materials science to the works published by a group of Berkeley researchers leaded by Xiang [103, 106]. Chen *et al.* [107] summarised the four basic methods in creating 'libraries': 1) thin-film methods, 2) solution-based methods, 3) inkjet printing methods,

and 4) dry powder mixing methods. Barber and Blamire [105] discussed the influence of a range of parameters during thin film deposition, and pointed out that the property characterization process needs to be localised, sensitive, high-throughput and rapid to assure the efficiency of combinatorial approach.

As can be seen, most combinatorial methods applied to materials science aim at mapping out the composition-structure-property correlations. A combinatorial approach has seldom been applied to correlate materials' properties. Historically, the discoveries of material correlations are in part similar to the discoveries of new materials, which are the results of deductive reasoning from known principles or serendipity [105]. In addition, material correlations can be derived from the time-consuming data analysis that involves substantial human effort by plotting properties against each other or by fitting them into hypothetical mathematical equations. The difficulty in discovering material correlations will be exacerbated with the increase of complexity that exists in the relationships.

The study of property-property correlations is an emerging part in materials science. Property correlations can be utilized to identify materials confined to desired properties. Meanwhile, when a target property is difficult to be measured experimentally, property correlations can be employed to make fast estimation through the usage of the other known properties. Therefore, the pertinent question might be how to accelerate the process in the exploration of property correlations.

Furthermore, the research community and industries have generated a tremendous amount of property data from scientific measurements. Deriving knowledge from those data is a longstanding issue and calls for human endeavour. The advanced automatic techniques, which allow huge datasets to be collected and stored, create ever larger challenges in this field. It is desirable to employ an advanced data mining method in capturing the complex property correlations hidden in the data with minimum or no prior knowledge.

ANNs have become increasingly aware of the ability to search correlations; hence, ANN models can be built in an effort to understand property data. However, preparing input variables for modelling is a fundamental, crucial choice that needs to be decided before the modelling. As May *et al.* [108] pointed out, an exhaustive search that

evaluates all of the possible combinations of input variables is the only method that is guaranteed to determine the optimal set of input variables for a given ANN model, i.e. the most interesting property correlation.

Inspired by the current combinatorial approach studies, the purpose of Chapter 3 is to apply the combinatorial approach as the exhaustive search method to produce possible property correlations for the subsequent ANN analysis. Instead of creating composition 'libraries', property 'libraries' are prepared and analysed. For example, if 20 properties for elements in the Periodic Table are considered, roughly, 3400 ternary order correlations will be produced, representing all possible combinations of three properties.

It should be emphasised that ANNs are capable of evaluating correlations in parallel, thus, the advantages of combinatorial approach that guarantee to find the optimal correlations in the shortest possible time are retained. By searching through the largely unexplored universe of binary, ternary, quaternary, or even higher order property correlations, it is possible to use ANNs to predict unknown data, test theoretical models and hypotheses, and finally enrich the understanding of the fundamental material properties.

1.6 Verification of the elastic properties of the elements

It is evident that the use of data is incredibly important in materials-based activities such as calculating, modelling and designing, whose performance clearly depends on the quality of input data [109]. Ensuring the quality of data has been a continuing concern due to the possible large cost of using inaccurate data [110]. Though handbooks and databases engaged in providing sufficient and authoritative data to satisfy the general needs of users, serious anomalies still exist. Data gaps and inconsistencies may arise and require clarifications, when the user finds different handbooks presenting different values.

Such discrepancy is not accidental. Apart from the occurrences of typographical errors, inaccuracy could be introduced when subjective judgement is made in selecting the most practical value or using the average value of multiple differing reported values in literature [111]. It is also recognized that years may elapse between the first published data in literature and the most recent data. Improved technology of measurement,

increasing knowledge, and deeper understanding of materials, normally lead to a more accurate value, sometimes even disprove the previously recorded data.

Small discrepancies between different handbooks and databases may be inescapable, but large discrepancies should be treated as errors that need to be verified. Correcting errors and updating the data are always important principles, under which new editions and newly written handbooks and databases are published [112]. However, such correction and updating seldom happen to the properties data of elements. Collections in handbooks or databases for elements are usually assembled/transcribed from earlier sources. As a result, existing errors could pass on like genetic mutations into subsequent editions [7]. This situation should now be changed.

Young's modulus (***E***), shear modulus (***G***), bulk modulus (***K***) and Poisson's ratio (***ν***) of pure metals are the important and fundamental mechanical properties in materials research and manufacturing applications. Despite the fact that elastic properties are relatively less sensitive to the purity of metals than other mechanical properties associated with plastic deformation [113], and have been measured for many years, the scatter of the measurement results has drawn very little attention [114, 115]. The accuracy of elastic properties is associated with the accuracy of many other properties in the calculation and modelling, such as hardness [116], specific heat capacity [113], melting point [117], and Debye temperature [118]. The ever-growing demand for accurate data in industry and academia necessitates the verification of the four elastic properties and the clarification of discrepancy presented by the major handbooks and databases.

Little work has been developed over the previous years to solve the problem of identifying outliers and errors in elastic properties of elements. To resolve the large discrepancy in handbooks and databases, it either needs large amounts of extra information in order to make subjective judgment that based on expert's knowledge [119] or must utilize the established mathematical equations between ***E***, ***G***, ***K*** and ***ν*** [114]. Like most statistical methods [2, 109] and outlier detections [120-123], such methods cannot be extended to a general situation where it is difficult or impossible to know the correlations or explicit mathematical function beforehand.

Artificial neural networks have been widely used to capture linear or nonlinear relationships with or without prior knowledge in diverse phenomena [124-126]. Recently, ANNs have been used to detect and correct errors of boiling points and enthalpies of vaporization of elements in handbooks [7]. However, Zhang *et al.* [7] only employed this method to a binary order correlation with data collected from five handbooks. Thus, it is desirable to know if ANNs are capable of detecting and correcting errors in a more complicated situation (i.e. a situation involving ternary order correlations). It is also desirable at this time to present a comprehensive review of single crystal elastic constants and polycrystalline elastic data for pure metals.

It is well known that the values of four elastic data of pure metals heavily depend on specimen purity [127-129], experimental temperature [130-132], and experimental pressure [133, 134]. Considering the time needed for data collection and the possible heavy work load caused by the size of problems, the majority of Chapter 4 focuses on the data (normally within 99.99% purity) at ambient pressure and room temperature (295±5 K). A general idea of how experimental parameters affect elastic property is discussed in Ref. [135].

As more and more single crystal and polycrystalline data have become available in recent years, a few compilations of elastic properties for pure metals have been prepared. Scott [119] made an excellent compilation for rare earth elements in 1978, but lutetium and thulium were studied in a limited way at that time. Elastic properties of six platinum metals were surveyed by Darling [136]. Sisodia and Verma [137] calculated polycrystalline elastic moduli in terms of single crystal elastic constants of some hexagonal and tetragonal metals. For cubic metals, single crystal elastic constant data can be found in Ref. [138, 139].

To the author's knowledge, the most completed compilation of elastic properties of elements was first prepared by Koster [140-142] and followed by Swamy and Narayana [115], and Gale [143]. However, some doubt has been cast on Koster's values [144], given the fact that his data were collected from static measurements. In addition, Gale's compilation [143] only included 52 pure metals, in which only 35 metals list bulk moduli. Even though Swamy and Narayana [115] intended to compile the data for a larger number of metals, most of their data were calculated from single crystal elastic

constants, providing no experimental polycrystalline data acquired from dynamic measurements. They also failed to disclose the original literature of single crystal data used for each element.

1.6.1 Elementary definitions

Elasticity is a physical property of materials that exhibits shape and volume changes to some extent when external stresses are applied, and will return to the original undeformed state after the stress removed [145]. Within the linear elastic regime, strain (***ε***) is proportional to stress (***σ***). The relationship between stress and strain is defined by Hooke's law in terms of elastic coefficient /compliance ***S*** or elastic constants /stiffness ***C***:

$$\varepsilon = S\sigma\ ,\ \sigma = C\varepsilon \qquad \textit{Equation 1-2}$$

When stresses and strains are acting in three-dimensions, and a subscript combination of 1, 2, or 3 correlates the three-dimensional surface and direction the stress or strain acting upon, Hooke's law may be written in a component form (with suffixes) as Equation 1-3 [146]:

$$\varepsilon_{ij} = S_{ijkl}\sigma_{kl}\ ,\ \sigma_{kl} = C_{ijkl}\varepsilon_{ij} \qquad \textit{Equation 1-3}$$

where ***i***, ***j***, ***k***, ***l*** is coordinate index (1, 2, or 3)

Because stress and strain tensors are both symmetric tensors [146],

$$S_{ijkl} = S_{jikl} \qquad \textit{Equation 1-4}$$

An abbreviation can be introduced to elastic constants $\boldsymbol{C_{ijkl}}$ [146]:

$$\begin{aligned} ii &\rightarrow i(i = j) \\ kk &\rightarrow k(k = l) \\ ij &\rightarrow 9 - i - j(i \neq j) \\ kl &\rightarrow 9 - k - l(k \neq l) \end{aligned}$$

For elastic coefficient $\boldsymbol{S_{ijkl}}$ [146]:

$$\begin{aligned} S_{ij} &\rightarrow S_{iijj} \\ S_{i,9-k-l} &= 2S_{iikl} \\ S_{9-i-j,9-k-l} &= 4S_{ijkl} \end{aligned}$$

Thus, tensor notation can be expanded to the form of compliance matrix ***S*** as the following:

$$\begin{Bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \varepsilon_4\\ \varepsilon_5\\ \varepsilon_6\end{Bmatrix}=\begin{bmatrix}S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16}\\ S_{21} & S_{22} & S_{23} & S_{24} & S_{25} & S_{26}\\ S_{31} & S_{32} & S_{33} & S_{34} & S_{35} & S_{36}\\ S_{41} & S_{42} & S_{43} & S_{44} & S_{45} & S_{46}\\ S_{51} & S_{52} & S_{53} & S_{54} & S_{55} & S_{56}\\ S_{61} & S_{62} & S_{63} & S_{64} & S_{65} & S_{66}\end{bmatrix}\begin{Bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\\ \sigma_4\\ \sigma_5\\ \sigma_6\end{Bmatrix},\varepsilon=S\sigma$$

Its inverse matrix is the elastic constant matrix ***C***:

$$\begin{Bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\\ \sigma_4\\ \sigma_5\\ \sigma_6\end{Bmatrix}=\begin{bmatrix}C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16}\\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26}\\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36}\\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46}\\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56}\\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66}\end{bmatrix}\begin{Bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \varepsilon_4\\ \varepsilon_5\\ \varepsilon_6\end{Bmatrix},\sigma=C\varepsilon$$

The relationship between the elastic coefficient and elastic constant is that the compliance matrix ***S*** can be obtained by taking the reciprocal of the elastic constant matrix ***C*** [146]:

$$S_{qp}=\frac{(-1)^{p+q}A_{pq}}{D} \qquad \textit{Equation 1-5}$$

where $\boldsymbol{A_{pq}}$ is the subdeterminant of the stiffness matrix ($\boldsymbol{C_{pq}}$) after eliminating the $\boldsymbol{p^{th}}$ row and the $\boldsymbol{q^{th}}$ column, and ***D*** is the determinant of ($\boldsymbol{C_{pq}}$).

In general, there are 36 matrix components. Since elastic constants represent the second derivatives of the energy density with respect to strain (see Equation 1-6) [147], and because the order of differentiation is irrelevant, the elastic constant matrices and elastic coefficient matrices should be symmetric. Therefore, only 21 stiffness components are actually independent for a anisotropic solid. The number of independent components of elastic constants can be further reduced considering the solid symmetry [148].

$$C_{qp}=\frac{1}{V}\left(\frac{\partial^2 U}{\partial\varepsilon_p\partial\varepsilon_q}\right) \qquad \textit{Equation 1-6}$$

In mechanical engineering, it is applicable to describe elastic deformation in solids with four elastic properties: Young's modulus (***E***), shear modulus, which is also known as rigidity modulus (***G***), bulk modulus (***K***), and Poisson's ratio (***ν***). ***E*** is defined as the ratio of stress to corresponding strain in a material under tension or compression, ***G*** and ***K*** are the ratio of stress to strain in a material subjected to shear stress and volumetric stress respectively, ***ν*** is the absolute value of ratio of lateral strain to corresponding axial (longitudinal) strain in the load direction [149].

$$E = \sigma/\varepsilon \qquad \textit{Equation 1-7}$$

$$G = \tau/\gamma \qquad \textit{Equation 1-8}$$

$$K = \sigma_m/(\Delta V/V) \qquad \textit{Equation 1-9}$$

$$\nu = -\varepsilon_y/\varepsilon_x \qquad \textit{Equation 1-10}$$

where $\boldsymbol{\sigma}$ is tensile or compression stress, $\boldsymbol{\varepsilon}$ is tensile or compression strain, τ is shear stress, $\boldsymbol{\gamma}$ is shear strain, $\boldsymbol{\sigma_m}$ is the mean or hydrostatic stress, and $\boldsymbol{\Delta V/V}$ is volumetric strain.

The Second Law of Thermodynamics requires the following holding limits on the elastic moduli: for any engineering materials, ***E***>0, ***G***>0, ***K***>0, and -1<***ν***<0.5 [150]. Recently, Mott and Roland [151] theoretically limited the lower bound of Poisson's ratio in isotropic materials to 0.2. But the experimental minimum value 0.039 has been determined for beryllium [152]. Materials with an extremely low value of Poisson's ratio also include cork and concrete, while the upper limit of ***ν*** (0.5) corresponds to incompressible elastic materials, i.e. soft rubber. For most crystalline metals, ***ν*** normally lies in the range 1/4 to 1/3 [153], and the ratio of shear modulus to Young's modulus generally equals to 3/8 [154].

1.6.2 Isotropic and anisotropic

Most metals and metallic alloys are considered as isotropic materials, which have a homogenous internal organization and display the same mechanical properties in all directions at an arbitrary point regardless of which surface and direction a force is applied [155]. Isotropic materials require two independent components, i.e. $\boldsymbol{C_{11}}$ and $\boldsymbol{C_{12}}$, to specify their elastic constants matrix ***C***.

In contrast, the mechanical properties of anisotropic materials depend on the directions in measurements. Without any planes of symmetry, a fourth rank tensor with 21 terms is required to describe the elasticity by relating the second rank tensor of stress and strain. Other material's behaviour is between isotropic and anisotropic, which needs between 2 and 21 independent elastic constants to describe the strain-stress relationship. For example, 9 independent elastic constants are needed to specify the fourth rank tensor for orthotropic materials, while for materials with cubic symmetry, only 3 independent components are needed. Single crystal solid is normally anisotropic, while polycrystalline aggregates are usually treated as isotropic. A summary of crystal symmetry, corresponding elastic constant matrix with independent elastic constants, and simplified equations to calculate the bounds for the aggregate properties from the single crystal elastic constants (whenever available), are listed in Table 1-1 for most pure metals at room temperature.

$$\begin{bmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{C_{11}-C_{12}}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{C_{11}-C_{12}}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{C_{11}-C_{12}}{2} \end{bmatrix}$$

The elastic properties of crystal aggregates can be derived from the elastic constants of single crystals by different approximations. One of the most widely used approximation methods is the Voigt-Reuss-Hill average (VRH) [156, 157], which is obtained by calculating the arithmetic mean of the Voigt bound ($\boldsymbol{G_V}$, $\boldsymbol{K_V}$) and the Reuss bound ($\boldsymbol{G_R}$, $\boldsymbol{K_R}$) of the elasticity based on the assumption of a homogeneous strain field and homogeneous stress field, respectively [158]. The VRH average can be determined from the single crystal elastic constants $\boldsymbol{C_{ij}}$ by employing Equation 1-11 to Equation 1-16 [159], where H denotes the Hill average value, and V and R denote the Voigt and Reuss bounds, respectively. Poisson's ratio and Young's modulus can be then computed from the values of shear modulus and bulk modulus determined by the VRH averaging method through the relationships described in Section 1.6.3.

Voigt approximation:

$$K_V = \frac{A_1+2A_2}{3} \qquad \textit{Equation 1-11}$$

$$G_V = \frac{A_1-A_2+3A_3}{5} \qquad \textit{Equation 1-12}$$

$$3\mathrm{A}_1 = C_{11} + C_{22} + C_{33}$$

$$3\mathrm{A}_2 = C_{23} + C_{31} + C_{12}$$

$$3A_3 = C_{44} + C_{55} + C_{66}$$

Reuss approximation:

$$K_R = \frac{1}{3a_1+6a_2} \qquad \textit{Equation 1-13}$$

$$G_R = \frac{5}{4a_1-4a_2+3a_3} \qquad \textit{Equation 1-14}$$

$$3\mathrm{a}_1 = S_{11} + S_{22} + S_{33}$$

$$3\mathrm{a}_2 = S_{23} + S_{31} + S_{12}$$

$$3a_3 = S_{44} + S_{55} + S_{66}$$

Hill approximation:

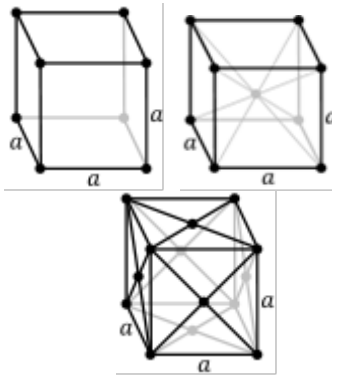
$$K_H = \frac{K_V+K_R}{2} \qquad \textit{Equation 1-15}$$

$$G_H = \frac{G_V+G_R}{2} \qquad \textit{Equation 1-16}$$

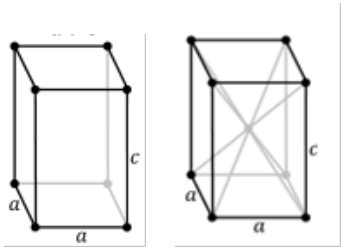
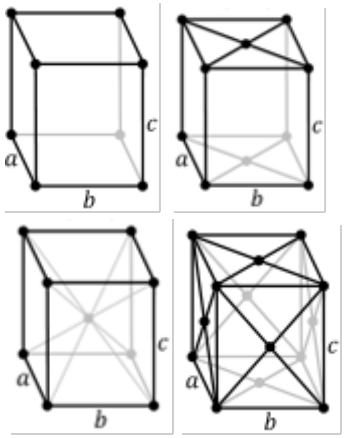
Because the Voigt-Reuss approximation provides the lowest upper bounds and the highest lower bounds, and the Hill average is an empirical estimation, it is argued that the true elastic properties of aggregate crystals may lie anywhere between the Voigt bound and the Reuss bound [160].

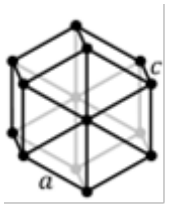
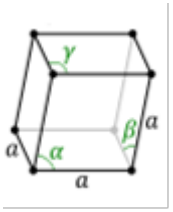
The Voigt-Reuss-Hill average is also known as the first-order bounds. The second-order bounds based on elastic energy expression have been introduced by Hashin and Shtrikman [161]. Though it is claimed to have higher accuracy than Voigt and Reuss bounds, and is suitable for various symmetries except triclinic [137, 162], the VRH average is still the most widely accepted method in literature regardless crystal symmetry. Therefore, as part of the present work, aggregates elastic properties calculated from single crystal elastic constants are all based on the VRH method.

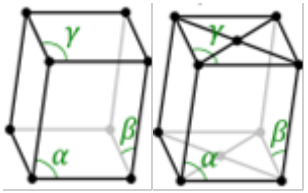
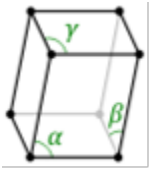
Table 1-1 Summary of stiffness matrixes with independent elastic constants, and simplified VRH formulas for the typical pure metals at room temperature.

Crystal System	Axial Distances & Axial Angles	Bravais Lattices	Simplified stiffness matrix [148, 163]	Simplified Formulas for VRH Averaging [162]	Metals
Cubic	a = b = c α = β = γ = 90°		$\begin{bmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{bmatrix}$	$K_V = K_R = \frac{C_{11} + 2C_{12}}{3}$ $G_V = \frac{C_{11} - C_{12} + 3C_{44}}{5}$ $G_R = \frac{5(C_{11} - C_{12})C_{44}}{4C_{44} + 3(C_{11} - C_{12})}$	Au, Al, Ag, Ba, Ca, Cr, Cu, Fe, Mo, Mn[1], Nb, Ni, Pb, Pd, Pt, Sr, Ta, V, W, Ir, Li, Rh, Cs, Th, Zr, Na, Ce, K, Eu, Rb, Yb

[1] Mn has the complex cubic structure.

Crystal System	Axial Distances & Axial Angles	Bravais Lattices	Simplified stiffness matrix [148, 163]	Simplified Formulas for VRH Averaging [162]	Metals
Tetragonal	a = b ≠ c α = β = γ = 90°		$\begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & C_{16} \\ C_{12} & C_{11} & C_{13} & 0 & 0 & -C_{16} \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ C_{16} & -C_{16} & 0 & 0 & 0 & C_{66} \end{bmatrix}$		Sn, In
Orthorhombic	a ≠ b ≠ c α = β = γ = 90°		$\begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{22} & C_{23} & 0 & 0 & 0 \\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix}$		Ga, U

Crystal System	Axial Distances & Axial Angles	Bravais Lattices	Simplified stiffness matrix [148, 163]	Simplified Formulas for VRH Averaging [162]	Metals
Hexagonal	a = b ≠ c α = β = 90°, γ = 120°		$\begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix}$ $C_{66} = \frac{C_{11} - C_{12}}{2}$	$K_V = \frac{2(C_{11} + C_{12}) + C_{33} + 4C_{13}}{9}$ $C^2 = (C_{11} + C_{12})C_{33} - 2C_{13}^{\ 2}$ $M = C_{11} + C_{12} + 2C_{33} - 4C_{13}$ $G_V = \frac{12C_{66} + 12C_{44} + M}{30}$ $K_R = \frac{C^2}{M}$ $G_R = \frac{5}{2}\left[\frac{C_{44}C_{66}C^2}{(C_{44} + C_{66})C^2 + 3K_V C_{44}C_{66}}\right]$	Co, Dy, Mg, Tb, Ti, Zn, Gd, Nd, Er, Pr, Tc, Os, Ru, Re, Sc, Th, Y, Ho, Hf, Tl, Be, Cd, La, Lu,
Rhombohedral (trigonal)	a = b = c α = β = γ ≠ 90°		$\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & 0 & 0 \\ C_{12} & C_{11} & C_{13} & C_{14} & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ C_{14} & C_{14} & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & C_{14} \\ 0 & 0 & 0 & 0 & C_{14} & C_{66} \end{bmatrix}$ $C_{66} = \frac{C_{11} - C_{12}}{2}$	$K_V = \frac{2(C_{11} + C_{12}) + C_{33} + 4C_{13}}{9}$ $C^2 = (C_{11} + C_{12})C_{33} - 2C_{13}^{\ 2}$ $M = C_{11} + C_{12} + 2C_{33} - 4C_{13}$ $K_R = \frac{C^2}{M}$ $G_V = \frac{12(C_{66} + C_{44}) + M}{30}$ $G_R = \frac{5C^2(C_{44}C_{66} - C_{14}^{\ 2})}{6K_V(C_{44}C_{66} - C_{14}^{\ 2}) + 2C^2(C_{44} + C_{66})}$	Bi, Sm

Crystal System	Axial Distances & Axial Angles	Bravais Lattices	Simplified stiffness matrix [148, 163]	Simplified Formulas for VRH Averaging [162]	Metals
Monoclinic	$a \neq b \neq c$ $\alpha = \gamma = 90°,$ $\beta \neq 90°$		$\begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & C_{16} \\ C_{12} & C_{22} & C_{23} & 0 & 0 & C_{26} \\ C_{13} & C_{23} & C_{33} & 0 & 0 & C_{36} \\ 0 & 0 & 0 & C_{44} & C_{45} & 0 \\ 0 & 0 & 0 & C_{45} & C_{55} & 0 \\ C_{16} & C_{26} & C_{36} & 0 & 0 & C_{66} \end{bmatrix}$		Pu
Triclinic	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma \neq 90°$		$\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{bmatrix}$		

1.6.3 The relationship between *E*, *G*, *K* and *v*

Provided that the solid is isotropic, a relationship exists between any three of the four elastic moduli, which means two independent elastic properties are sufficient to describe all modes of deformation [149]. Figure 1-4 illustrates the strained shape of a cube material subjected to the action of the shear and complementary shear forces [145].

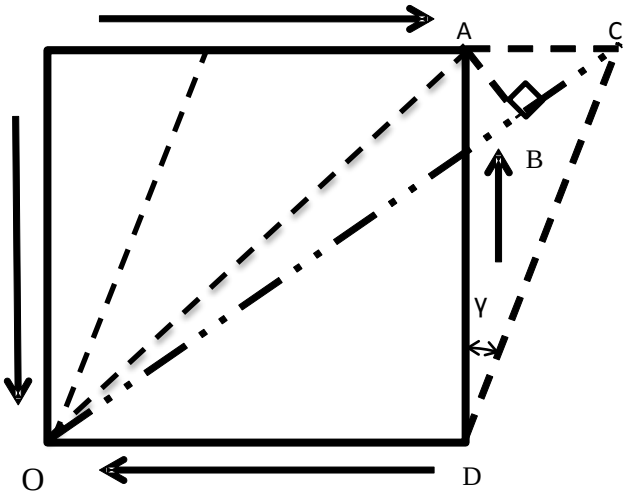


Figure 1-4 Element subjected to shear and associated complementary shear (from Ref. [145]).

Assuming that the strains are so small that the angle ACB may be taken as 45°:

$$\varepsilon_d = \frac{BC}{OA} \approx \frac{AC\cos 45^\circ}{l\sqrt{2}} = \frac{AC}{2l} = \frac{l\gamma}{2l} = \frac{\gamma}{2}$$

$$\because \gamma = \frac{\tau}{G}$$

$$\therefore \varepsilon_d = \frac{\tau}{2G} \qquad \textit{Equation 1-17}$$

where $\boldsymbol{\varepsilon_d}$ is the strain on diagonal, $\boldsymbol{l}$ is the length of one side of the cube, $\boldsymbol{\gamma}$ is the angle of distortion or shear strain.

The shear stresses system is equivalent to the direct stress system as shown in Figure 1-5, which can be represented by one compressive set and one tensile set, each at 45° to the original shear directions, and equals in magnitude to the applied shear [145].

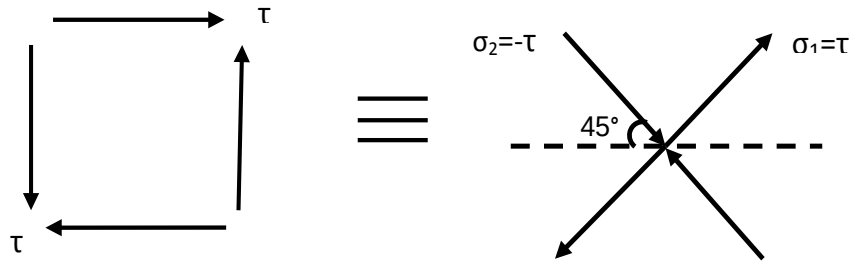


Figure 1-5 Direct stresses due to shear (from Ref. [145]).

$$\varepsilon_d = \frac{\sigma_1}{E} - \nu\frac{\sigma_2}{E} = \frac{\tau}{E} - \nu\frac{-\tau}{E} = \frac{\tau}{E}(1+\nu) \qquad \textit{Equation 1-18}$$

Combining Equation 1-17 and Equation 1-18, the correlating equation between Young's modulus, shear modulus, and Poisson's ratio is obtained:

$$E = 2G(1+\nu) \qquad \textit{Equation 1-19}$$

Consider a cube subjected to three equal stresses $\boldsymbol{\sigma}$ [145]:

$$\varepsilon_t = \frac{\sigma}{E} - \nu\frac{\sigma}{E} - \nu\frac{\sigma}{E} = \frac{\sigma}{E}(1-2\nu)$$

$$\because \varepsilon_v = \varepsilon_x + \varepsilon_y + \varepsilon_z = 3\varepsilon_t = \frac{3\sigma}{E}(1-2\nu)$$

$$\because \varepsilon_v = \frac{\sigma}{K}$$

$$\therefore E = 3K(1-2\nu) \qquad \textit{Equation 1-20}$$

where $\boldsymbol{\varepsilon_t}$ is the total strain along one edge, and $\boldsymbol{\varepsilon_v}$ is the volumetric strain.

As shown in Equation 1-19 and Equation 1-20 each elastic property can be expressed in terms of any other two properties. There are total 12 correlating equations as summarized in Table 1-2 [155]. Their corresponding annotations are listed in Table 1-3. For the sake of convenience, the correlating equations used in Chapter 4 refer to the equations listed in Table 1-2 unless otherwise specified.

Table 1-2 Relations between the elastic properties [155].

	Young's modulus ***E***	**Shear modulus** ***G***	**Bulk modulus** ***K***	**Poisson's ratio** ***v***
E, ***G***			$\frac{GE}{3(3G-E)}$	$\frac{E-2G}{2G}$
G, ***v***	$2G(1+\nu)$		$\frac{2G(1+\nu)}{3(1-2\nu)}$	
G, ***K***	$\frac{9KG}{3K+G}$			$\frac{3K-2G}{2(3K+G)}$
E, ***v***		$\frac{E}{2(1+\nu)}$	$\frac{E}{3(1-2\nu)}$	
E, ***K***		$\frac{3EK}{9K-E}$		$\frac{3K-E}{6K}$
K, ***v***	$3K(1-2\nu)$	$\frac{3K(1-2\nu)}{2(1+\nu)}$		

Table 1-3 Annotations for the correlating equations listed in Table 1-2. Equation 1-21 is used to represent all the correlating equations for convenience.

	Young's modulus ***E***	**Shear modulus** ***G***	**Bulk modulus** ***K***	**Poisson's ratio** ***v***
E, ***G***			Equation 1-21-2	Equation 1-21-3
G, ***v***	Equation 1-21-1		Equation 1-21-4	
G, ***K***	Equation 1-21-5			Equation 1-21-6
E, ***v***		Equation 1-21-7	Equation 1-21-8	
E, ***K***		Equation 1-21-9		Equation 1-21-10
K, ***v***	Equation 1-21-11	Equation 1-21-12		

1.6.4 Static and dynamic measurements

The values of the elastic properties ***K***, ***G***, ***E*** and ***ν*** can be experimentally measured by a large number approaches based on static tests or dynamic tests, see Figure 1-6 [164]. The terms of static and dynamic refer to the strain rate and amplitude: static testing introduces large elastic strain with slow strain rates and is isothermal in nature, while dynamic testing involves small elastic strain with high strain rates and is inherently adiabatic process [165]. Sometimes, compressibility (χ), which is usually determined by static methods, is used instead of bulk modulus (***K***). As discussed in Ref. [119], isothermal values (χ_t) can be converted to adiabatic values (χ_s) via the following Equation 1-22.

$$\chi_S = \chi_t - \frac{9\alpha_L{}^2 V_a T}{C_p} \qquad \textit{Equation 1-22}$$

where α_L is linear thermal expansion coefficient, V_a is atomic volume, which equals to the atomic weight divided by the density, ***T*** is absolute temperature (in unit of K), and ***Cp*** is isobaric specific heat.

Because the difference between the isothermal and adiabatic value for a metal at room temperature is only a few percentage or less [166], and handbooks / databases usually do not specify whether isothermal or adiabatic values are given in their compilation, no distinction is made in the present work.

1) Static approaches

The static approaches are mostly based on the direct measurements of stresses and strains, such as tension tests, torsion tests and flexure tests. Although many standards [167-170] have been provided as guidance to determine elastic properties for metals, static approaches often yield poor results due to the strain from material creep or deflection of the test machine [171].

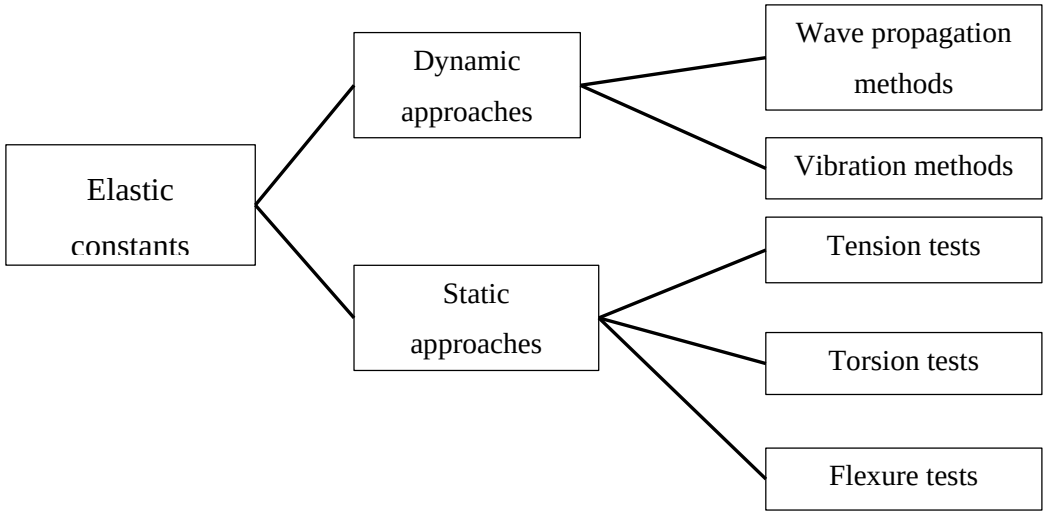


Figure 1-6 Classification of the methodologies employed to determine the elastic properties of solids (from Ref. [164]).

i. Tension test

The Young's modulus can be measured from a simple uniaxial tension test. As shown in Figure 1-7 [172], a 'dogbone' specimen is elongated at a constant rate by the moving crosshead of the tensile testing machine. The increasing applied load $\boldsymbol{F}$ and the elongation ($\boldsymbol{L}$-$\boldsymbol{L_0}$) is simultaneously measured using a load cell and an extensometer, respectively. Stress $\boldsymbol{\sigma}$ is then computed by dividing load $\boldsymbol{F}$ by the original cross-sectional area $\boldsymbol{A_0}$ of the specimen, and strain $\boldsymbol{\varepsilon}$ is computed by dividing elongation ($\boldsymbol{L}$-$\boldsymbol{L_0}$) by the original distance $\boldsymbol{L_0}$ (see Equation 1-23 to Equation 1-25) [173]. A typical stress-strain curve obtained by plotting $\boldsymbol{\sigma}$ verse $\boldsymbol{\varepsilon}$ is also shown in Figure 1-7 (c). The magnitude of the slope of the linear line in elastic region equals the magnitude of Young's modulus.

$$\sigma = \frac{F}{A_o} \qquad \textit{Equation 1-23}$$

$$\varepsilon = \frac{L - L_o}{L_o} \qquad \textit{Equation 1-24}$$

$$E = \frac{\sigma}{\varepsilon} \qquad \textit{Equation 1-25}$$

where $\boldsymbol{F}$ is the instantaneous load applied perpendicular to the specimen cross section, $\boldsymbol{A_0}$ is the original cross-sectional area before any load is applied, $\boldsymbol{L_0}$ is the original length before any load is applied, $\boldsymbol{L}$ is the instantaneous length.

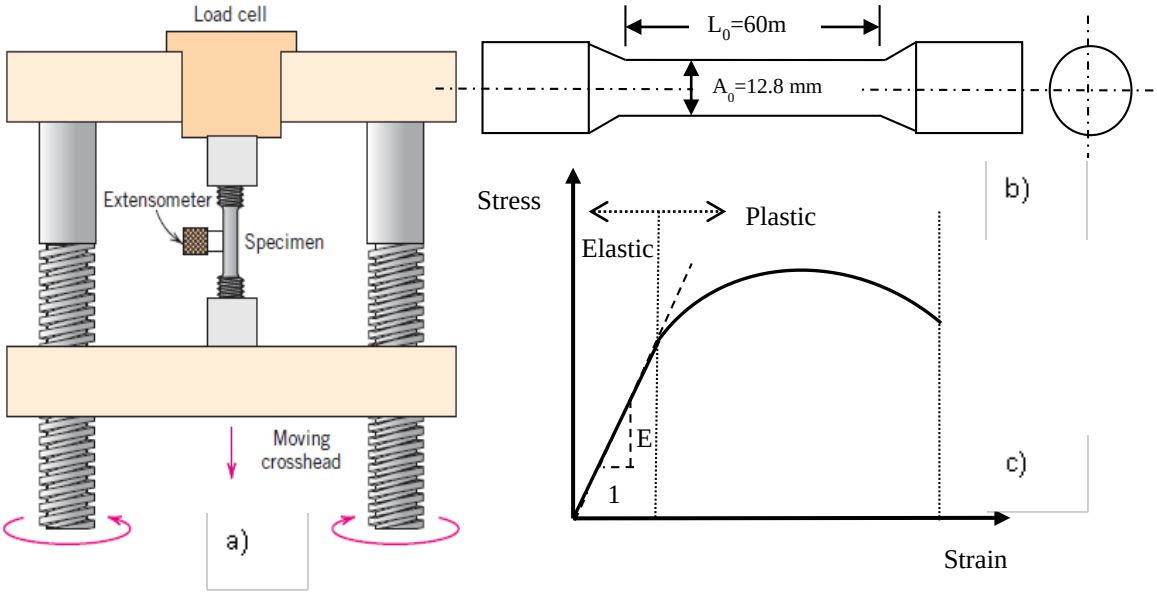


Figure 1-7 a) Schematic representation of the apparatus used to conduct tensile stress–strain tests (adapted from Ref. [172]); b) A typical standard specimen; and c) A typical stress-strain curve.

ii. Torsion test

The shear modulus can be obtained from a torsion test (see Figure 1-8). Point A on the surface of a solid cylindrical shaft of radius r rotates to point B subjected to a torque T_t. The shear stress τ, shear strain γ and shear modulus G are computed by the equations shown below [172].

$$\tau = \frac{T_t r}{J} \qquad \textit{Equation 1-26}$$

$$\gamma = \frac{r\theta_1}{L} \qquad \textit{Equation 1-27}$$

$$G = \frac{T_t L}{J\theta_1} \qquad \textit{Equation 1-28}$$

where θ_1 is the angle of twist, L is the length of the cylindrical solid shaft, T_t is the applied torque, r is the radial distance, J is the polar moment of inertia, for a solid cylinder, $J = \frac{\pi r^4}{2}$.

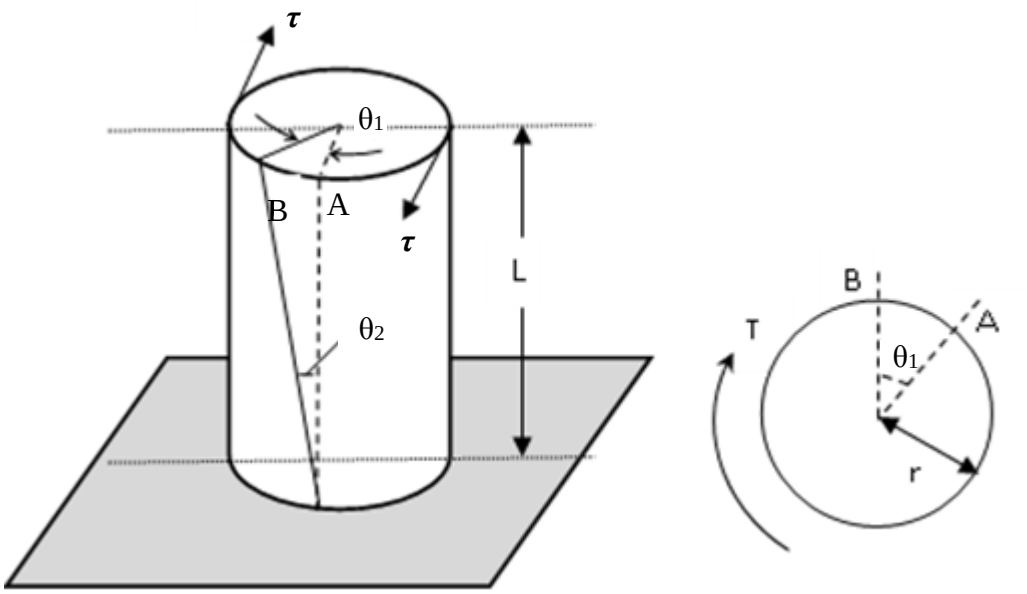


Figure 1-8 Schematic representation of torsional deformation (from Ref. [172]).

iii. Flexure test

Flexural modulus testing is designed to apply small forces but achieve large displacements. It is ideal for brittle materials. There are two test types: 3-point flexure test and 4-point flexure test (see Figure 1-9). The former test is most common for polymers, while the latter is common for wood and composites.

For the displacement of loading points in four-point bending [174]:

$$E = \frac{2\left(F_1 - F_2\right)d_1^2\left(d_1 + 3d_2\right)}{b_t h^3\left(\delta_c - \delta_s\right)} \qquad \textit{Equation 1-29}$$

where:

$\boldsymbol{E}$ = Young's modulus in Pa;

$\boldsymbol{F_1}$ = Lower load level selected from recordings in N;

$\boldsymbol{F_2}$ = Upper load level selected from recordings in N;

$\boldsymbol{d_1}$ = Test jig inner roller to outer roller spacing in four-point bending in m;

$\boldsymbol{d_2}$ = One half of the test jig inner span in four-point bending in m;

$\boldsymbol{b_t}$ = Test-piece width in m;

$\boldsymbol{h}$ = Test-piece thickness in m;

$\boldsymbol{\delta_c}$ = Displacement for the thin test-piece in the jig over load interval $\boldsymbol{F_1}$ to $\boldsymbol{F_2}$, in m;

$\boldsymbol{\delta_s}$ = Displacement for the thick steel bar (replacing the test-piece) in the jig over load interval $\boldsymbol{F_1}$ to $\boldsymbol{F_2}$, in m;

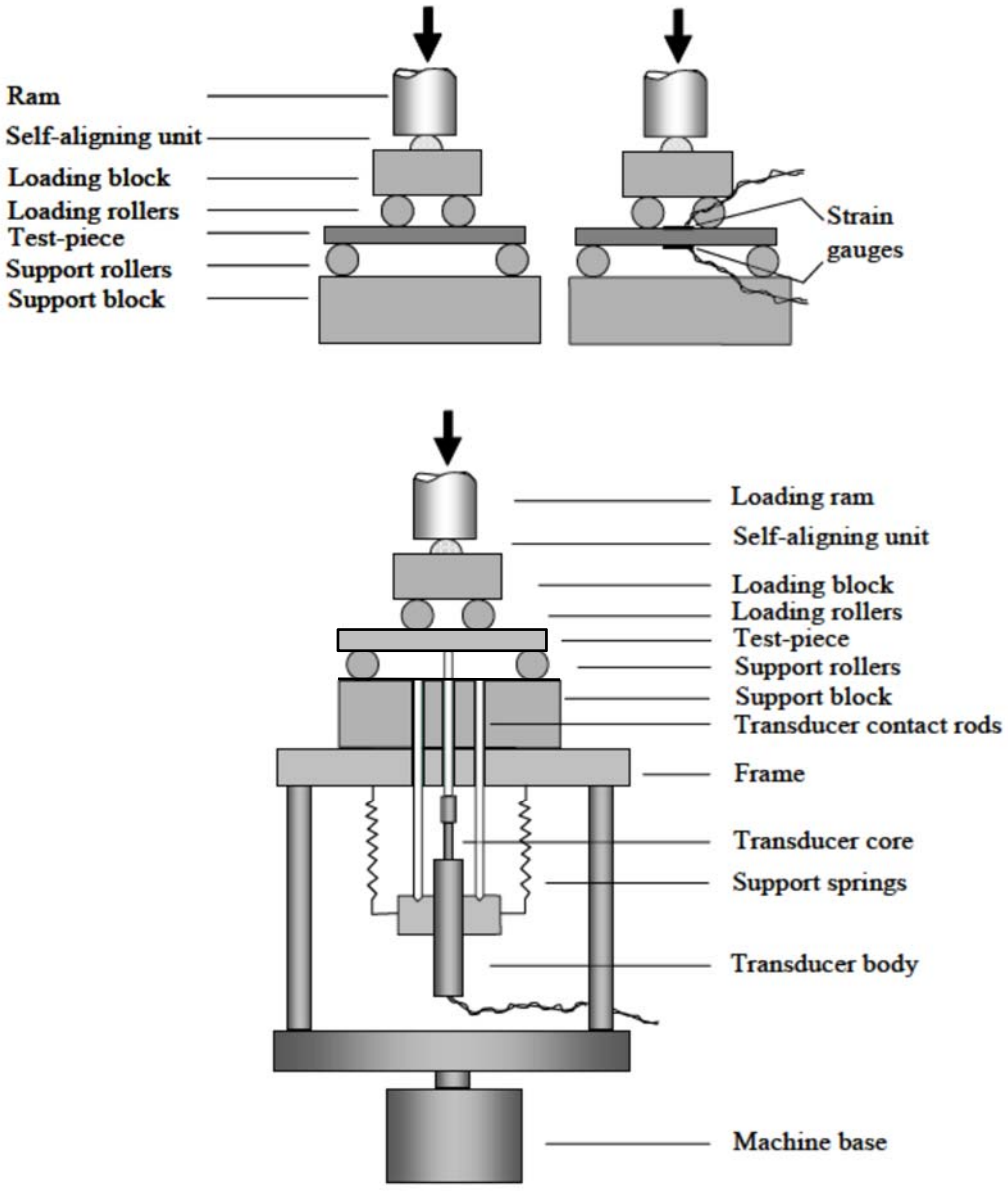


Figure 1-9 Schematic diagrams of (top left) four-point flexure using machine displacement measurement, (top right) strain gauges applied to both sides of the test-piece, and (lower) direct measurement of deflection between fixed points on the test-piece surface using a single linear displacement transducer (from Ref. [174] and modified by the author).

2) Dynamic approaches

In comparison with static approaches, dynamic approaches can very precisely and non-destructively measure elastic moduli of specimens with a greater variety of shapes and dimensions at a wide range of temperatures [164]. Lord and Morrell [174] summarized the relative merits of the dynamic and static approaches for measuring elastic moduli, see Table 1-4. They concluded the theoretical errors in measurement of elastic moduli by dynamic methods are typically of the order of ±1%, in contrast to the much more significant errors generated by static methods at low strains. The advantages against static approaches encourage the developments of dynamic methods, which can be classified into two groups: wave propagation based methods and vibration methods.

Table 1-4 Relative merits of the dynamic and static approaches [174].

	Static approach	**Dynamic approach**
Advantages	▪ "Engineering value" for modulus ▪ Generation of stress-strain curve ▪ Widely available test equipment	▪ Quick, simple, non-destructive ▪ Good inherent accuracy ▪ Uses small specimens ▪ High temperature measurement ▪ Can readily measure shear modulus and Poisson's ratio
Disadvantages	▪ High accuracy strain measurement required ▪ Need averaging extensometry ▪ Specialised test ▪ Larger specimens required ▪ Large interlaboratory scatter ▪ Accurate high temperature measurements are difficult	▪ Sensitive to dimensional tolerances ▪ Methods do not always work well for some materials and composites ▪ Calculations require some knowledge of other material parameters ▪ Equipment not widely available

i. Wave propagation methods

The most commonly used wave propagation method for the measurement of elastic property is the ultrasonic pulse technique. It is non-destructive and measures the speed of wave propagation through the material [175]. The elastic moduli of an isotropic material can be obtained by measuring the speed of sound velocities and material density. The four elastic properties can be determined from Equation 1-30 to Equation 1-33 [176], and all elastic constants can be determined from Equation 1-34 to Equation 1-36 [176]. For materials with preferred orientation, the equations to calculate elastic properties are more complex, requiring more sound speed measurements in specific directions relative to the material symmetry axes.

$$E = \frac{\rho V_S^2 (3V_L^2 - 4V_S^2)}{V_L^2 - V_S^2} \quad \textit{Equation 1-30}$$

$$K = \rho \left(V_L^2 - \frac{4V_S^2}{3} \right) \quad \textit{Equation 1-31}$$

$$G = \rho V_S^2 \quad \textit{Equation 1-32}$$

$$\nu = \frac{V_L^2 - 2V_S^2}{2V_L^2 - 2V_S^2} \quad \textit{Equation 1-33}$$

$$C_{11} = \rho V_L^2 \quad \textit{Equation 1-34}$$

$$C_{12} = \rho (V_L^2 - 2V_S^2) \quad \textit{Equation 1-35}$$

$$C_{44} = \rho V_S^2 \quad \textit{Equation 1-36}$$

where $\boldsymbol{V_L}$ and $\boldsymbol{V_S}$ are the longitudinal and shear velocities.

ii. Vibration methods

Though wave propagation methods were robust and widely performed to determine the elastic constants of single crystals, possible inhomogeneity of test specimen would adversely affect the accuracy of measurements [164]. To deal with this problem, vibration methods have become available. With the knowledge of the size and mass information of the sample, such tests determine elastic properties, which are proportional to the square of material resonances, by making a beam test piece with uniform cross-section vibrate mechanically [174]. According to different vibration

modes, these methods can be reduced to three categories, as illustrated in Figure 1-10 [177].

Figure 1-11 is a schematic graph of a typical thread suspension flexural vibration test apparatus that has the advantage to allow the specimen to oscillate without significant restriction [178]. Electrical signal generated by the oscillator is amplified and transformed into mechanical oscillation through a transducer thus drives specimen at resonance. Another transducer on the other side detects such vibration in the specimen. The vibration transmitted through an amplifier, and is displayed on an oscilloscope. Frequency can be determined by either Lissajous figure analysis or a frequency meter.

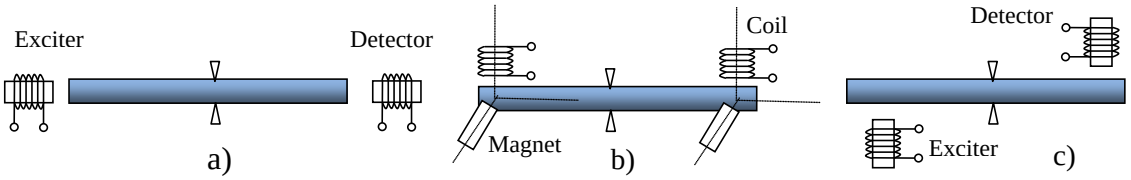


Figure 1-10 Three vibration modes using electromagnetic-acoustic transducers: a) longitudinal, b) torsional, and c) flexural [177].

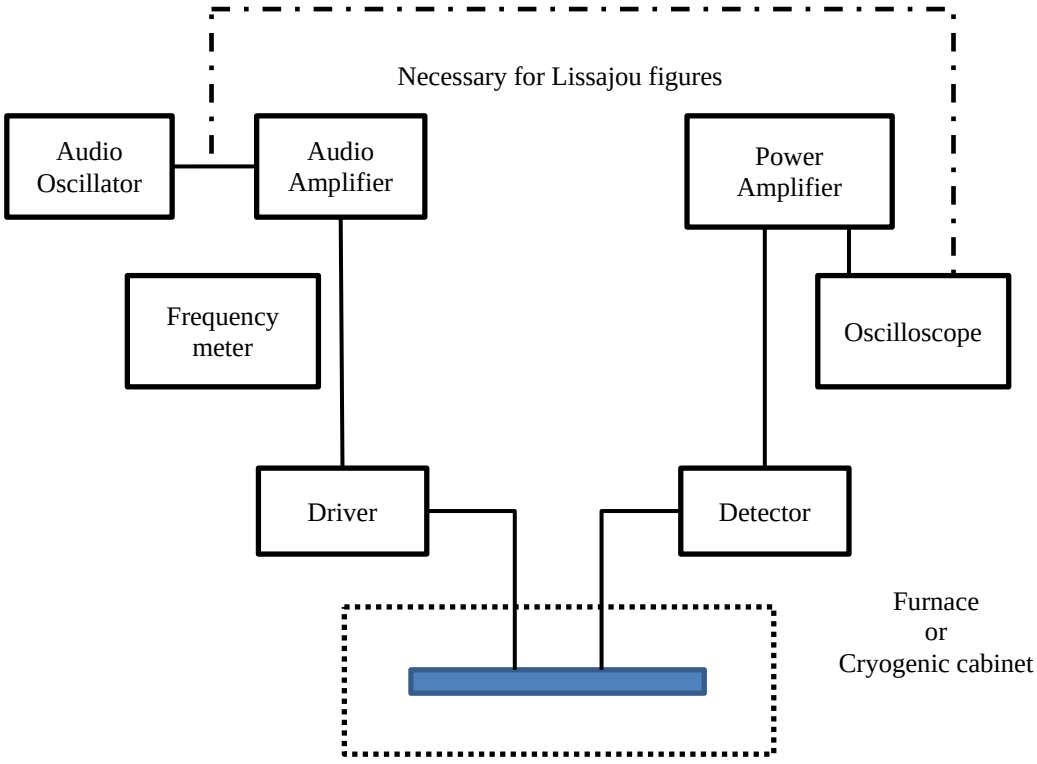


Figure 1-11 Block diagram of a typical thread suspension flexural vibration test apparatus (from Ref. [178]).

1.7 Capturing materials properties correlations using artificial neural networks: an example in hardening of pure metals by high pressure torsion

Capturing correlations from carefully measured experimental data and creating theories are one of the important methods in scientific research. One of the famous examples in history is the discovery of the laws of planetary motion by Johannes Kepler who spent many years in analysing Tycho Brahe's carefully and systematically recorded astronomical data. Capturing correlations has been dramatically accelerated by the application of modern IT technology [79, 123].

As one of the important Data Mining techniques, ANNs, which have been widely used to study diverse phenomena that are not yet accessible to full physical modelling, are capable of capturing linear or nonlinear relationships [126, 179]. The ability of ANNs in exploring relationships between multiple/unknown numbers of variables has been demonstrated by a number of successful applications in materials science [4, 78]. ANNs are especially useful in capturing relationships without knowing whether an explicit physical model exists.

Advantages of ANNs include computationally high parallelism, fault and failure tolerance, learning ability to handle imprecise and fuzzy information, making this modelling very feasible [53]. Combinatorial search using ANNs is a quantitative treatment to handle a large numbers of variables with or without prior knowledge to pick up properties correlations that are perceived to be significant [180, 181]. It has been applied recently to automatically find the correlations between materials properties by searching a large number of data spaces.

In Chapter 5, ANNs have been used to capture correlations between fundamental materials properties and the material’s hardness after high pressure torsion (HPT) processing. A physical model to explain this phenomenon has been published elsewhere [116]. The strength/hardness of metals can be significantly improved by HPT. In a HPT process, a disc sample subjected to compressive force and torsional straining (as shown in Figure 1-12), will result in grain refinement and reach a saturated hardness level [182].

Several factors influencing the hardness of pure metals due to HPT have been investigated, and it has been indicated that the hardness at saturated level correlates to atomic bond parameters, specific heat capacity, specific latent heat of fusion, linear thermal expansion coefficient, activation energy for self-diffusion, melting temperature and shear modulus [183-189]. However, little attention has been given to address the question: whether a limited set of physical properties can explain all measured changes in hardness due to HPT.

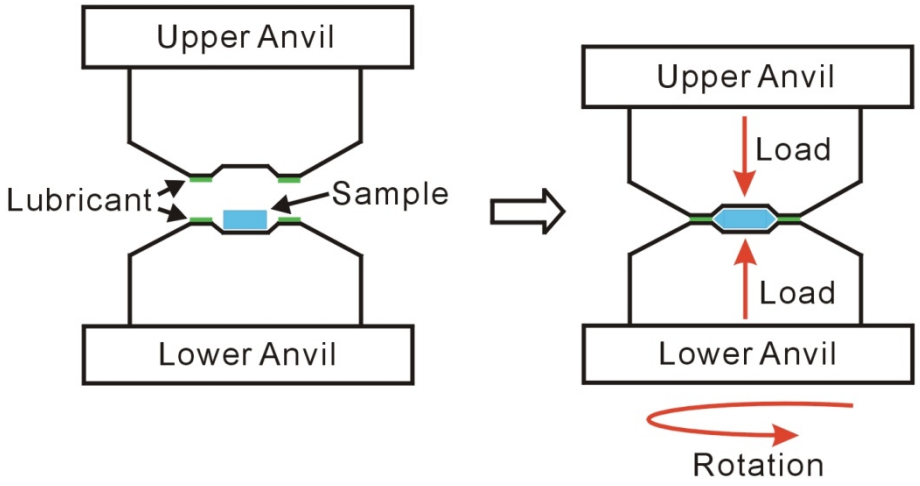


Figure 1-12 The principle of high pressure torsion (from Ref. [182]).

ANNs are generally used for complex system modelling with a large input data set. As a rule of thumb, training subset should be larger than the total number of input variables. It is generally recommended having more than 10 examples per input variable, though the size of data required in real practice is depending on the complexity of the problem and amount of noise in data [190]. Considering scientists would like to benefit from ANNs in a situation where the size of data set has an inherent limitation, it is time to find whether effective ANNs can be established with a small input data. For example, a relatively limited number of pure metals have been processed through HPT. Although the HPT processing has been applied to a wide range of materials [191], for instance, aluminium-based alloys, it is an easy start to create ANNs using results that only relate to pure metals.

In Chapter 5, a systematic method based on artificial neural networks is analysed in detail. ANNs are applied to derive the correlations between the hardness increment of

pure metals and their thermal, electronic, and mechanical properties without any prior knowledge being applied. The aim is to identify the important properties that contribute most to the changes of hardness. In the author's related work [116], a physical model based on strengthening by dislocations and thermally activated recovery was developed to predict the increment of Vickers hardness of pure metals due to HPT. The physical model construction process initially benefited from the parameters identified by ANNs. Although ANNs are established by using only 17 data points (the same 17 metals that were used in the physical model), the results of ANN assessment and the physically based model mutually corroborate each other.

1.8 The discovery of materials properties correlations by artificial neural networks and symbolic regression

Materials property-property correlations have been used increasingly to describe new models, estimate property data, and match materials to designs. Unlike materials structure-property correlations, which are usually obtained from a conceptual view of bonding and structure, materials property-property correlations are often distilled from statistical analysis. They are expected to be presented in a quantitative way, i.e. in the forms of mathematical equation, which allows easy accesses and further explorations [192].

Despite the fact that materials property-property correlations could be complex and highly nonlinear, a wide range of modelling techniques are available for relationship modelling, property prediction and data forecasting. The modelling techniques include artificial neural networks [193], random forests [194], decision trees [195], support vector machines [196], and genetic programming (GP) [197]. Moreover, fixed-form parametric models can also be developed based on expert knowledge. Depending on the level of prior information needed to fully describe the model, Giustolisi and Savic [192] used white-box, grey-box and black-box to represent the situations where most, modest and least information is available in the model constructions process. Obtaining necessary information to construct white-box or grey-box models may not always be possible, thus, the applications of data driven modelling techniques have attracted a great deal of interests from scientists.

One of the emerging powerful methods to automatically find the analytic expressions of correlations is called symbolic regression (SR) via genetic programming. Instead of fitting parameters to an equation of a given form, SR evolutionarily searches for the optimum mathematical expression, including both the numerical parameters and the functional form of relationships between variables, by minimizing the modelling error [28]. The advantage to deal with nonlinear correlations without any prior knowledge makes SR override the traditional statistical models and empirical models that need to predefine equation structures. In contrast with other data driven modelling techniques, which are usually difficult to interpret as functional relationships [198], the solution of symbolic regression can be explicit interpreted into an analytic form of a mathematical correlating equation.

Inspired by the pioneer work conducted by Koza [197], studies reporting the advantages of symbolic regression approach can be found in many disciplines [199]. Precise equations have been obtained by Brezocnik and Gusel [200] to predict radial stress distribution in cold-formed material. McKay *et al.* [201] employed symbolic regression to simulate multiple chemical process systems. Cai *et al.* [202] claimed that the heat transfer correlations extracted from the experimental data had smaller predictive errors than their published counterparts. Laws of geometric and momentum conservation from experimental data captured from various physical systems were re-discovered by Schmidt and Lipson [28]. In hydrology, applying symbolic expression to solve three example problems including rainfall-runoff modelling [203], resistance coefficient in corrugated pipes [204], and end-depth ratio in open channels [205], demonstrate the advantages of the new method as a 'transparent' and structured system identification that outperforms many data-driven methods.

However, one major problem in employing symbolic regression via genetic programming is that the generated symbolic function tends to grow in complexity over time [192]. Several attempts have been made in order to yield more parsimonious model structures including improvements made by Grosman and Lewin [206], least squares optimization method [207], an evolutionary polynomial regression method proposed by Giustolisi and Savic [192], and so on. Nevertheless, Chapter 6 started from another perspective: the ANN method is employed as an additional automatic way to capture

domain knowledge of relevant inputs for SR systems, thus, the analytic function can be significantly reduced.

In Chapter 6, the problem of obtaining mathematical equations (which represent the hidden materials property-correlations) from a large amount of data is investigated. To be specific, the correlating equations are expected to make accurate predictions about the enthalpies of vaporization of compounds. Firstly, ANNs are utilized to test whether any materials property-correlation exists in the dataset prepared from CRC handbook. When a property-correlation is found, the impact of each variable that contributes to the prediction is analysed through three different methods: i) the ANN combinatorial search, ii) the Partial Derivatives method (PaD), and iii) the Connection Weight approach (CW). Variables that have the largest influences on the accuracy of the prediction are identified by doing this. The extra information captured by ANNs is then fed into a standard SR model and an explicit mathematical equation has been obtained to predict the enthalpies of vaporization of 175 compounds.

1.8.1 The enthalpy of vaporization

The enthalpy of vaporization ($\boldsymbol{\Delta H_{vap}}$, given in units of kJ/mol) is the energy required to transform a given quantity of a pure compound from liquid state into gas at the corresponding vapour pressure [208]. It is also known as the latent heat of vaporization, and is usually measured at the normal boiling point where the pressure equals to 101kPa ($\boldsymbol{\Delta H_{vb}}$). The enthalpy of vaporization at the normal boiling point is an important parameter for the design and development of industrial processes at elevated temperatures [209, 210], such as distillation, evaporation, dry, etc. The value of $\boldsymbol{\Delta H_{vap}}$ at one temperature can be used to calculate the heat of vaporization of pure substances at other temperatures by employing the Watson functions indicated as Equation 1-37 [211].

$$\Delta H_{vap2} = \Delta H_{vap1} \left(\frac{1-T_{r2}}{1-T_{r1}}\right)^{n} \qquad \textit{Equation 1-37}$$

where the subscripts 1 and 2 refer to temperatures 1 and 2, $\boldsymbol{T_{ri}}=\boldsymbol{T_i}/\boldsymbol{T_c}$, and a common choice for the constant of n is 0.375 or 0.38.

A few methods for the evaluation of $\boldsymbol{\Delta H_{vb}}$ utilizing other property parameters have been proposed [211]. However, all the estimation methods developed in decades are either modifications to the empirical models developed by the predecessors or white-box

models that were developed merely based on researchers' expert knowledge. Since ANNs can be used to capture correlations between properties. The author is curious to know whether their corresponding analytical forms utilizing the same input variables can also be acquired when the hidden correlations are found. In addition, it is of great interest to develop the correlating equations through SR, and compare such correlating equations with the empirical models proposed in literature. A simple but accurate correlating equation obtained from SR would show a great advantage especially when a new problem is encountered and the underlying mechanism is not wholly understood.

1.8.2 Genetic programming and symbolic regression

Genetic programming (GP) is a highly parallel computational modelling technique that emerged in the late 20th century and increased in popularity since then. Inspired by the evolution theory of natural selection, it simulates the process of genetic transmission of characteristics by performing the operation of reproduction, crossover, and mutation, when searching for a better solution [212]. When use GP to solve a given problem, each potential solution of the problem is referred as an individual, a number of individuals constitute a population (a group of possible solutions). By evaluating the fitness of each generation (i.e. how close the solution fit the input data), a set of feasible candidate solutions from one population is transformed into a newly created population that address user-defined issues with improved fitness.

A flowchart of the 'survival of fittest' process is given in Figure 1-13 [212]. At the very beginning of GP, an initial population of individuals are randomly created as the evolutionary starting point. A fitness value is assigned for each individual according to an error evaluation function. Individuals with higher fitness survive and are reproduced at a higher rate, thus they are more likely to be preserved in the next generation. Meanwhile, the individuals chosen from the current generation are stochastically recombined to form a new population through the operator of crossover and mutation. The new population is then used in the next iteration of the algorithm, until some termination criterions are satisfied: a maximum number of generations have been produced, or a desired fitness level has been reached.

The GP process of solving problems can be viewed as a search for a highly fitted tree-like computer program, where all tree-nodes are selected randomly from the function set

or terminals set, while the root node uses only the function set [197]. A function set consists of operators such as arithmetic operations (+, -, ×, ÷, etc.), mathematical functions (sin, cos, exp, log, etc.), Boolean operations (AND, OR, NOT), conditional operators (If-Then-Else), functions causing iteration (Do-Until), and so on. The terminal set comprises the arguments that the function set can operate on, including constants. An example of such a tree-like program is given in Figure 1-14. The operations of reproduction, crossover, and mutation in a tree-like program are illustrated in Figure 1-15. The major components of GP include the following aspects: (1) the set of terminals, (2) the set of functions, (3) the fitness measurement,(4) the parameters for controlling the run, (5) the criterion for termination and 6) the result designation method. Figure 1-16 is a summary for the basic GP preparatory steps [213].

The special application of GP used in this thesis is called Symbolic Regression. It searches for the optimal model structure (the form of equation) as well as associated parameters simultaneously, and generates a mathematical expression that provides a good solution to the input data points [28]. Unlike the traditional numeric regression problems, no predefined model structure is needed as a priori. In order to distinguish with the 'program-based GP', which has a program output (i.e. code), Elshorbagy *et al.* [214] use 'equation-based GP' to emphasise that an explicit equation can be obtained by SR.

However, to generate parsimonious equations, a problem called 'bloat' needs to be solved. 'Bloat' is used to describe mathematical expressions that have non-functional codes (introns), making the solution harder to understand [199]. 'Bloat' could be caused by an inappropriate choice of input variables. Though it is claimed that SR has the ability to discriminate relevant and irrelevant inputs [215], 'bloat' can still occur due to the existing of redundant input variables which have limit impacts on the output. Having the redundant input variables removed in the beginning would make the generated mathematical equations easier to be interpreted.

Moreover, mathematical expressions generated by SR are differing in size, shapes, and complexity, due to the computational resources, time, and expenses control. Identifying variables that have little or no predictive power and thus removing them before the implementation of SR, would offer substantially advantages. Nevertheless, input variables for SR modelling are generally chosen by users based on human intuition,

experiences and experiment observations [200]. It is desirable to have an additional automatic ways to determine the appropriate set of input variables from a possible large variable-pool for SR modelling [216].

1.8.3 ANNs to identify the contributions of input variables

ANNs are one of the most widely used data-driven techniques that have the ability to deal with complex information. As demonstrated by the applications in Chapter 3 ~ Chapter 5, ANNs show great predictive power in capturing the linear or non-linear relationships between input variables without having to assume the relation function in advance. However, interpreting the symbolic meaning behind ANNs (in regard to the weight matrix and network internal behaviour) into a comprehensible form of knowledge has been a long standing issue [217].

Though ANNs are often regarded as 'black-boxes' that the relationships are encoded as weight vectors [124, 218], a number of methods studying variable contributions in ANNs have been developed in recent years and received growing interests [219-222]. The most notable advances were made by Gevrey *et al.* [221] and Olden *et al.* [218] respectively, who agreed that ANNs could adequately identify redundant and noisy variables. Nevertheless, Gevrey *et al.* [221] compared different variable-importance evaluation methods based on an empirical dataset, and found the Partial Derivatives method, which calculated the partial derivatives of the output according to the input variables, outperforming other variable contribution analysis methods. But a comparison study to access the relative importance of the input factors provided by Olden *et al.* [218] based on a simulated dataset, indicates the Connection Weight approach that uses weight vectors of ANNs, is the optimal method to fully rank the importance of input variables.

In Chapter 6, both the PaD method and the CW approach are deployed to study the contributions of input variables. A combinatorial search is also conducted to analyse the predictive performance of input variable combinations. The most important input variables are identified, and are subsequently used to develop the correlating equations through SR. Thus, explicit equations that exclude redundancy (uninformative candidate input variables) are generated.

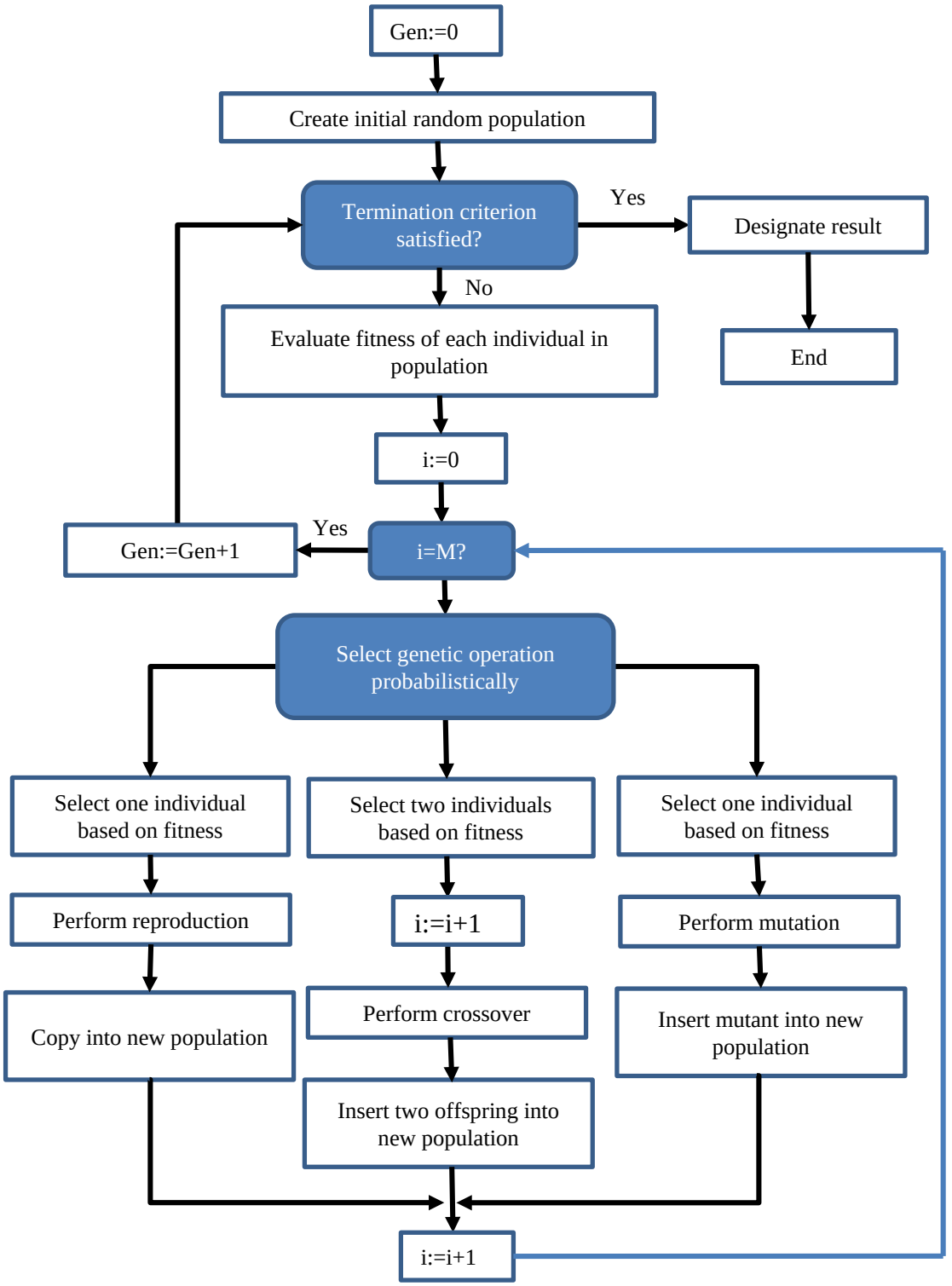


Figure 1-13 Flowchart of the genetic programming paradigm (from Ref. [212]). Index i refers to an individual in the population of size M, and GEN is the number of the current generation.

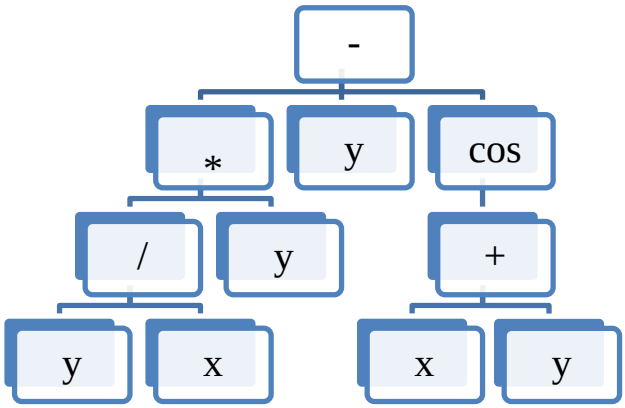
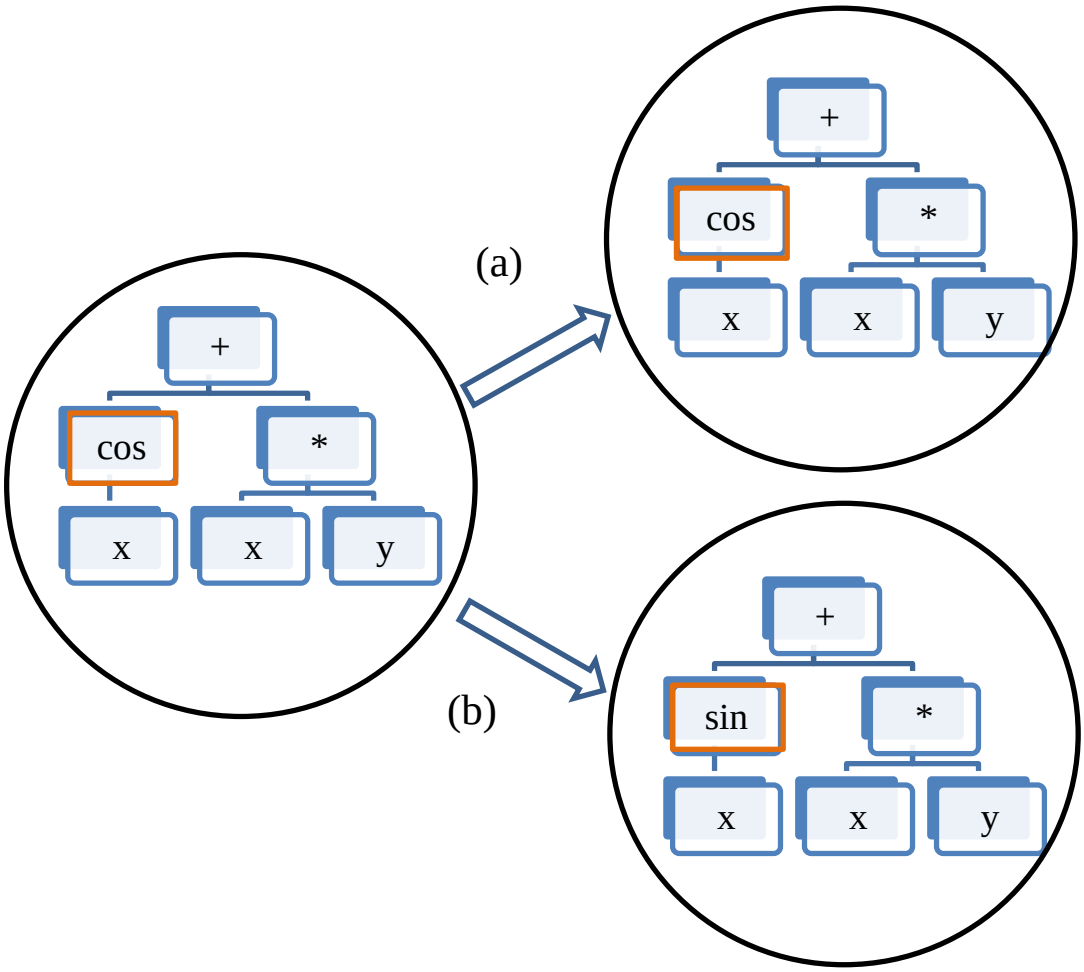


*Figure 1-14 A tree-like hierarchically structured computer program corresponding to the equation: (y/x)*y-y-cos(x+y).*



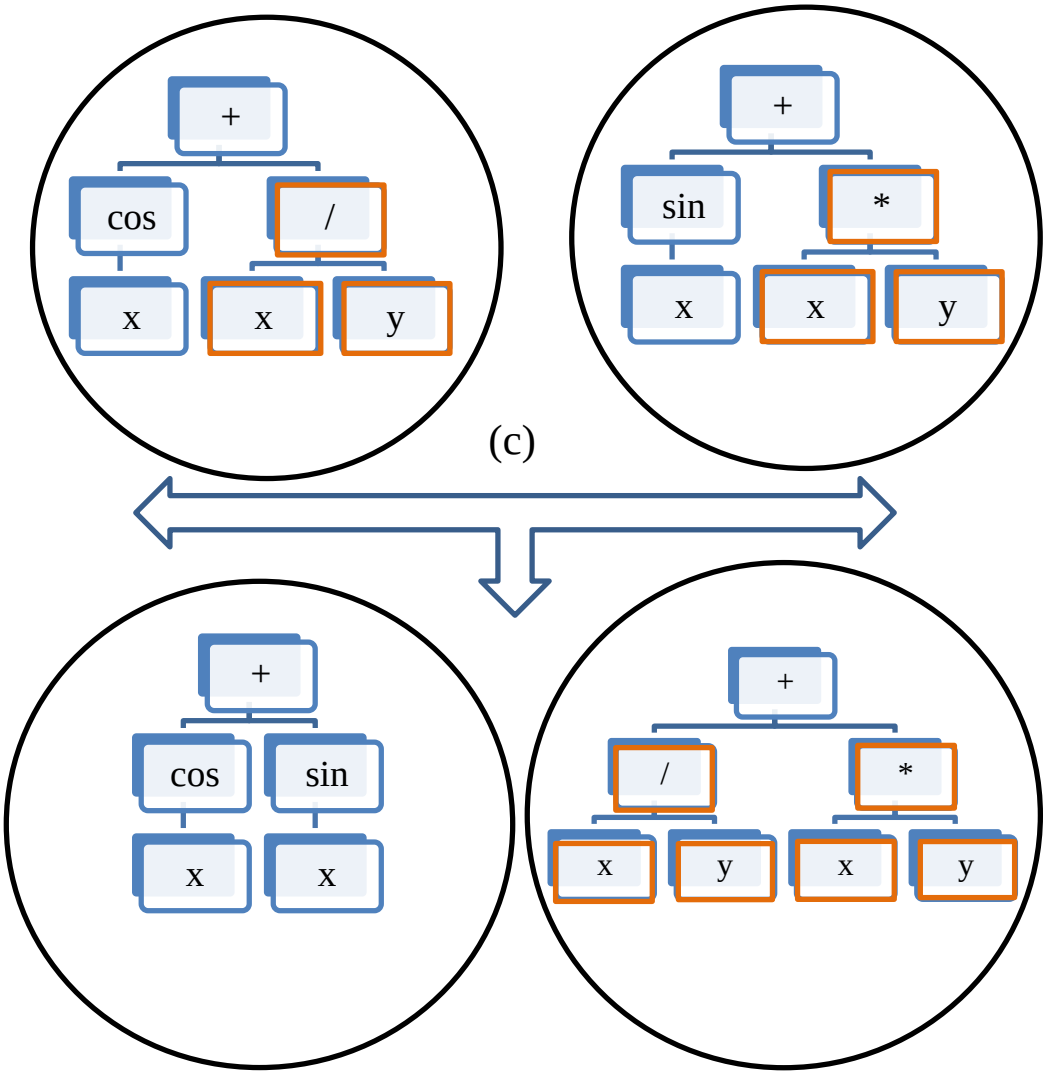


Figure 1-15 Illustration of GP operations: a) reproduction, b) mutation, and c) crossover.

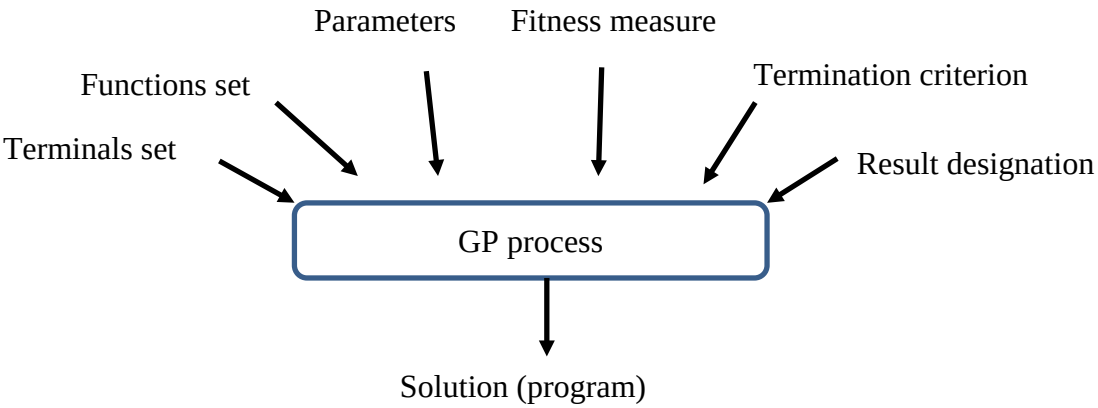


Figure 1-16 Major preparatory steps of the basic GP process (from Ref. [213]).

2. ANN methodology and configuration

With better understanding of biological brain function and the faster development of scientific technology, modern neural network research has bifurcated into two fields: one is concerned with creating effective network architectures and algorithms, and the other is concerned with the applications of ANNs [21]. The work presented in the thesis belongs to the second field that a well-accepted standard BPANN model is employed in the application of materials science to explore property-property correlations.

According to Masters [223] and Heaton [224] that problems that require two hidden layers are rarely encountered, as well as the promising initial results obtained when only one hidden layer was used, a three-layer feed forward backpropagation network containing an input layer, a hidden layer and an output layer has been chosen to be investigated for capturing the hidden property-property correlations. The number of neurons in the input layer equals the vectors/dimensions of the input data, the number of neurons in the output layer is constrained by the number of output data. In the four ANNs applications, the output of networks is defined to be one property. Hence, there is only one neuron in the output layer. Determining the optimum number of neurons in hidden layer ($\boldsymbol{Z_o}$) by adequate ANN training in a reasonable/acceptable amount of time is a long-standing puzzle since it depends on multiple factors [179]:

- the size of input and output data;
- the number of training cases,;
- the quality of the data sets;
- the complexity of problems intended to be solved;
- the architecture of ANNs;
- the type of activation function and transfer function;
- the training algorithm;
- regularization.

Based on the author's experience, the choice of $\boldsymbol{Z_o}$ is generally considered to obey the following two rules:

i. $\boldsymbol{Z}$ should be between the size of the input layer and the size of the output layer (i.e. 1) in order to avoid over-fitting or under-fitting.
ii. The total number of connections in ANNs to be fitted should be smaller than the number of the cases available for ANNs training in order to avoid over-fitting.

The larger number calculated from the above rules is denoted as $\boldsymbol{H}$. A loop trial-and-error program suggested by Malinov and Sha [225] is then applied to determine the best number of neurons within the range of [1, $\boldsymbol{H}$] that achieves the best predictive performance.

The predictive performance is assessed not only based on the observation cases, but also on the evaluation cases that were not used in the network training process. ANNs are trained by feeding a balanced set of training examples that effectively represent global signals into the networks. With enough information process capacity (neurons), training errors are reduced and general features hidden in the complicated data set are captured. However, if the training examples are not sufficient, which will cause a problem called under-fitting, or there are too many neurons to learn trivial features in the training examples that turn out to be irrelevant with the general population, which in turn results in over-fitting, the ANNs obtained will not be able to generalize well on a new dataset [90]. In order to make sure that the obtained ANNs have a good generalisation ability, the performance of ANNs using ‘unseen’ data, which are already collected but not been used in the training processes, is also evaluated and used as one of the criteria for the network performance evaluation.

Apart from determination of the number of neurons in the hidden layer, the loop programme, as illustrated in Figure 2-1, is employed to randomly divide the total data set into two sub-sets, i.e. observation cases and evaluation cases, at a ratio of 4:1 for $\boldsymbol{Q}$ times. Because the amount of training time will increase dramatically with the increase in the number of training cases and hidden neurons, it would be time consuming to set $\boldsymbol{Q}$ to be a very large value for all problems. It is the author's experience that for regression problems, if the size of the observation cases is smaller than 500, $\boldsymbol{Q}$ setting to 30 is large enough to obtain satisfactory ANN models in most cases. Thus, in the

experiments described in Chapter 4 ~ Chapter 6, $\boldsymbol{Q}$ is set to 30 (though some ANN models may not converge). While in the experiment described in Chapter 3, where all ANN models need to be successfully constructed for an effective correlation-group analysis, the minimum value of $\boldsymbol{Q}$ is set to 30, and the upper limit of $\boldsymbol{Q}$ equals to the times when ANNs finally converge.

The entire programs are run on the Matlab 2010a platform [226] where initial weights and bias of the network are automatically generated. All input data are normalized to a uniform range of [-1, 1] using Equation 2-1, which is a standard procedure to improve the speed of learning process. The 'trainbr' training function, that can update weight and bias values according to Bayesian regularization, is used together with the mean square error performance function. It minimizes a combination of squared errors and weights to determine the best combination of squared errors and weights [227]. 'Tansig' (hyperbolic tangent sigmoid transfer function) is used as transfer function for the hidden layers. The transfer function for the output layer is selected to be 'Purelin' (linear transfer function), which is popularly used especially in solving regression problems.

$$y_i = \frac{2(x_i - x_{min})}{x_{max} - x_{min}} - 1 \qquad \textit{Equation 2-1}$$

where $\boldsymbol{x_i}$ is the $\boldsymbol{i}^{\text{th}}$ input before normalization, $\boldsymbol{x_{min}}$ is the minimum value of the inputs, $\boldsymbol{x_{max}}$ is the maximum value of the inputs, and $\boldsymbol{y_i}$ is the $\boldsymbol{i}^{\text{th}}$ input after normalization.

The network is trained for 800 epochs to meet a training goal of 1×10^{-8}. During the training, the weights and biases are constantly updated to map the outputs with the inputs until: 1) the maximum number of epochs (800) is reached, or, 2) the minimum performance value is obtained, i.e. mean squared error (MSE, see Equation 2-2) equals to 1×10^{-8}.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \qquad \textit{Equation 2-2}$$

where $\boldsymbol{a_i}$ is the network predicted value, $\boldsymbol{t_i}$ is the target output, $\boldsymbol{e_i}$ is the difference between the network predicted value and the target output, and $\boldsymbol{N}$ is the number of training cases available [227].

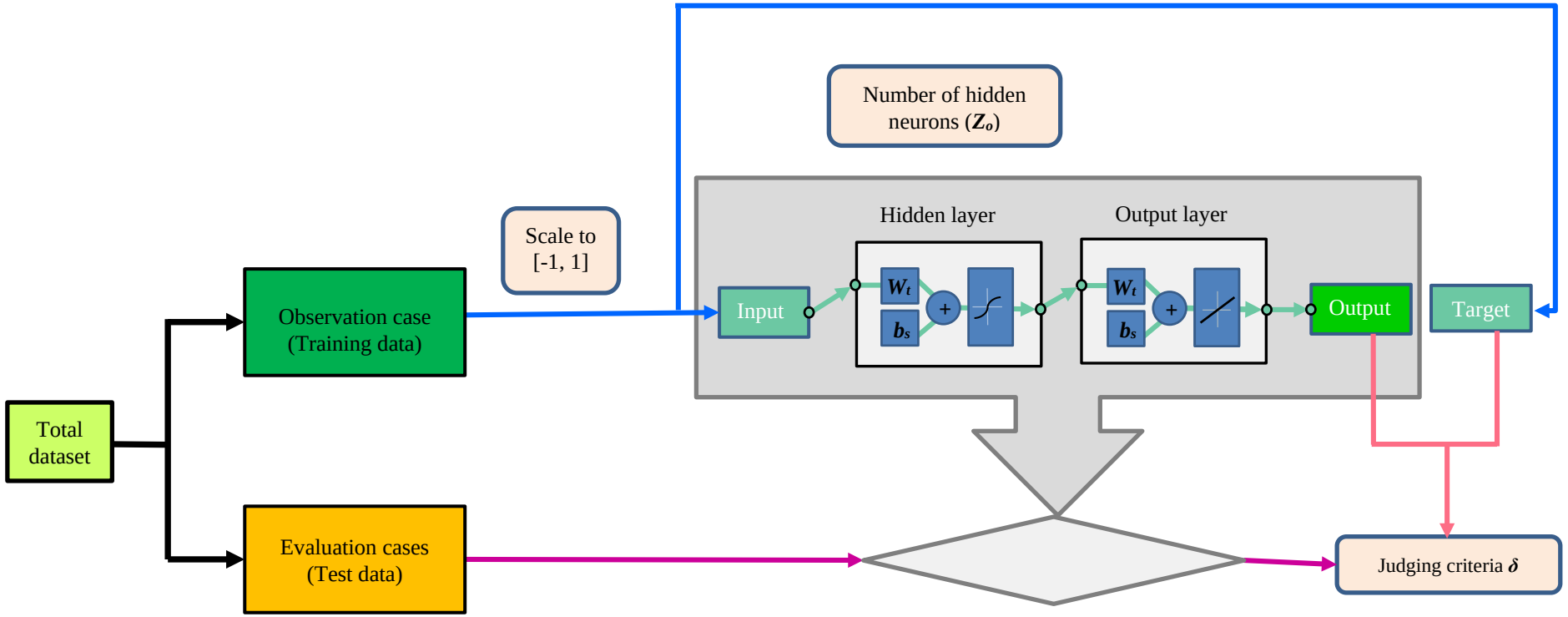


*Figure 2-1 A flowchart showing how the desired Neural network with one hidden layer was achieved. 'Tansig' is the transfer function used in the hidden layer and 'Purelin' for the output layer. '$\boldsymbol{W_t}$' is the weight, while '$\boldsymbol{b_s}$' stands for bias. The total available data are randomly partitioned as observation cases and evaluation cases at a ratio of 4:1. This process is repeated for **Q** times. For each time, the number of hidden neurons is decided in a loop program to achieve the smallest $\boldsymbol{\delta}$. The designate ANN is the one that has the smallest $\boldsymbol{\delta}$ in all the **Q** times.*

2.1 Inputs and outputs of ANNs

The inputs and outputs of ANNs for each application are discussed in the experimental section of each chapter. To summarise:

In Chapter 3, a database of 37 metallic elements for the construction of ANNs is prepared from CES EduPack [228]. The inputs and outputs of ANNs are chosen from 24 properties, which include: 1) atomic radius, 2) atomic weight, 3) boiling point, 4) bulk modulus, 5) cohesive energy, 6) debye temperature, 7) electrical resistivity, 8) heat of fusion, 9) heat of vaporization, 10) lattice parameter ***a***, 11) magnetic susceptibility, 12) melting temperature, 13) molar volume, 14) neutron absorption cross section (0.025 eV), 15) neutron scattering cross section (0.025 eV), 16) Poisson's ratio, 17) shear modulus, 18) specific heat capacity, 19) surface energy, 20) T-dependence of resistivity, 21) thermal conductivity, 22) thermal expansion coefficient, 23) work function, and 24) Young's modulus.

In Chapter 4, the values of Young's modulus, shear modulus, bulk modulus and Poisson's ratio for the 68 pure metals are collected from 12 different sources. One property out of the four properties is selected as the output, and any other two properties are used as the inputs.

In Chapter 5, the database of 17 pure metals is collected from several journal articles. The target output is the absolute increment of hardness after HPT. The input properties are chosen from 13 properties. The 13 properties are: 1) atomic number,2) binding energy per nucleon, 3) cohesive energy, 4) density, 5) heat of fusion, 6) lattice parameter ***a***, 7) melting temperature, 8) molar volume, 9) shear modulus, 10) specific heat capacity, 11) thermal expansion coefficient, 12) work function, and 13) Burgers vector.

In Chapter 6, the database of 175 organic and inorganic compounds is collected from CRC Handbook [229]. Enthalpy of vaporization at boiling point is the target output of ANNs and the SR models. The input properties are chosen from five properties: 1) normal boiling point, 2) critical point, 3) critical pressure, 4) dipole moment, and 5) molecular weight.

2.2 Early stopping

A method called early stopping can improve network generalization when dataset size is relatively large [230]. Generally, data are randomly divided into three subsets at a ratio of 8:1:1, which are the training, validation and testing data respectively. The training data are used to train the network by updating the network weights and biases. The errors from validation data and test data are monitored to ensure the network generalises well during training:

- if the validation error increases over 5 iterations, the training will be terminated;
- if the iteration number of minimum error reached in the test set is significantly different from that of the validation set, a new partition of data may be needed.

2.3 Bayesian regularization

Another important technique called Bayesian regularization can improve network generalization when dataset size is relatively small. The objective of the ANN training process is to minimize the objective function. Initially, the objective function is equivalent to the sum of squared errors between the network predicted value and the target output ($\boldsymbol{E_D}$), while Bayesian regularization adds an additional term to modify the objective function from Equation 2-3 to Equation 2-4 [231].

The original objective function $\boldsymbol{F_o}$:

$$F_o = E_D = \sum_{i=1}^{N}(e_i)^2 = \sum_{i=1}^{N}(t_i - a_i)^2 \qquad \textit{Equation 2-3}$$

The modified objective function $\boldsymbol{F_m}$:

$$F_m = \beta E_D + \alpha E_w \qquad \textit{Equation 2-4}$$

where $\boldsymbol{E_w}$ is the sum of squares of the network weights, and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are objective function parameters that automatically update in each iteration until reach convergence.

In Bayesian regularization framework, it is assumed that the true underlying function has a degree of smoothness, and the network response will be smooth if the size of network weights is constrained [231]. Matlab employs a Gauss-Newton approximation to progressively refine the regularization parameters (i.e. $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$).

It has been demonstrated that the Bayesian regularization produces networks which have excellent generalization capabilities [232]. Because Bayesian regularization tends to provide better generalization performance than Early stopping method especially when the size of the dataset is relatively small [227], only Bayesian regularization is employed in the experiments described in Chapter 3 ~ Chapter 6 to improve ANNs generalization ability. Therefore, training data and test data are used instead of observation cases and evaluation cases to describe ANN results for the sake of convenience.

2.4 General evaluation criteria

In order to evaluate the overall performance of neural networks based on the total data set including both the observation cases and the evaluation cases, two methods can be used. The first one is $\boldsymbol{\delta}$ (see Equation 2-5 and Equation 2-6), which is a comprehensive evaluation of $\boldsymbol{S_l}$ and $\boldsymbol{R_l}$ for both the training data and test data. The smaller the $\boldsymbol{\delta}$, the better ANNs perform, and the stronger the correlations. Another way is to simply compare the difference between the experimental data and the prediction values. In Chapter 3, the two evaluation criteria are combined to access correlation-group performances according to a quadratic superposition rule. In Chapter 4 ~ Chapter 6, when the difference of $\boldsymbol{\delta}$ is relative small between different ANNs, the second evaluation criterion, i.e. the average error, is more favoured in describing the ANN performances.

$$\begin{aligned}\delta &= |\varphi^2_{observation} - \varphi^2_{evaluation}| \\ &= |\varphi_{observation} + \varphi_{evaluation}| \times |\varphi_{observation} - \varphi_{evaluation}|\end{aligned} \qquad \textit{Equation 2-5}$$

$$\varphi = |S_l - 1| + (1 - R_l) \qquad \textit{Equation 2-6}$$

where $\boldsymbol{S_l}$ is the slope of the linear regression line for both the observation cases and the evaluation cases, and $\boldsymbol{R_l}$ is the regression coefficient.

The criterion $\boldsymbol{\delta}$ is defined by Equation 2-5 and Equation 2-6. This criterion is first introduced in Ref. [10]. The smaller the $|\varphi_{\text{observation}} + \varphi_{evaluation}|$, the lower error of the ANNs from either the observation cases that used to train the network and the evaluation cases that used to test the generalization ability of the network. The smaller the $|\varphi_{\text{observation}} - \varphi_{evaluation}|$, the smaller difference between the accuracy of

observation cases and evaluation cases, which is expect to be as close to zero as possible. This is another indicator of the network generalization ability.

The performance of an ANN increases with the decreases of $\boldsymbol{\delta}$, and the desired ANN is the one with the smallest $\boldsymbol{\delta}$. In this thesis, the best ANN out of $\boldsymbol{Q}$ independent trials that has the smallest value of $\boldsymbol{\delta}$ is chosen. An example is given in Figure 2-2 (see more discussion about the figure in Section 5.2.3). It shows ANNs predicting the hardness increment due to HPT from shear modulus and melting temperature. For an accurate prediction, the solid red and dot blue line should coincide, and both the slope $\boldsymbol{S_I}$ and correlation coefficient $\boldsymbol{R_I}$ should approach 1.

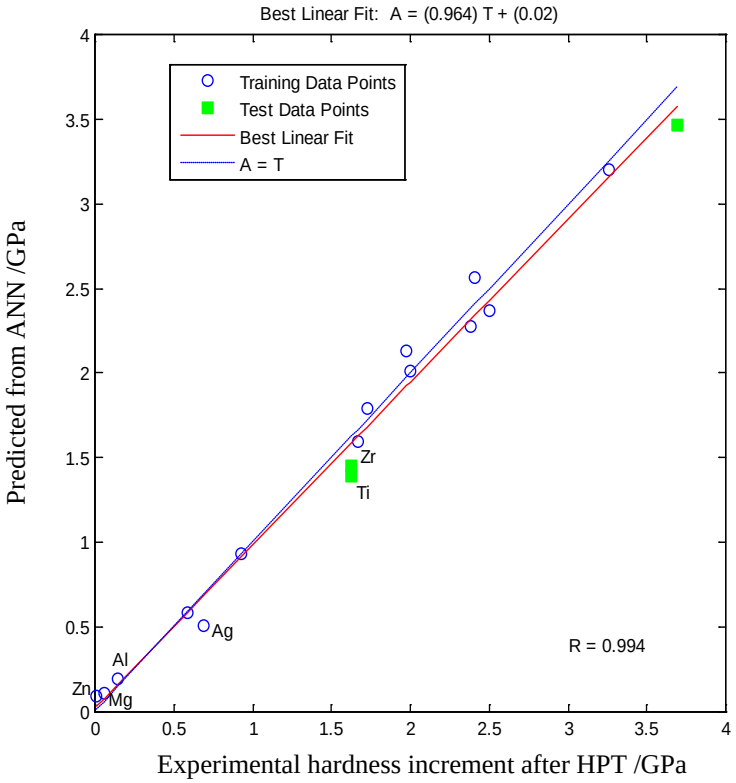


Figure 2-2 Result of ANNs in predicting the hardness increment due to HPT from shear modulus and melting temperature for 17 elements with $\boldsymbol{\delta}$ *equals to 0.04 and 12.67% error.*

3. Capturing property correlations through a combinatorial ANN search

3.1 Introduction

Property correlations are useful in understanding the properties, generating new information, calculating missing property values, testing the internal consistency of experimental data, and the construction of theoretical models. However, capturing property correlations generally requires substantial human effort and prior expert knowledge. The difficulty of correlations exploration is exacerbated with the increase of complexity that exists in the correlations as well as the ever-enlarging databases.

In the following section, an ANN combinatorial search is applied to investigate the binary, ternary and quaternary order correlations of 24 properties in 37 metallic elements. Instead of creating composition 'libraries', property 'libraries' are prepared and analysed. The ANN combinatorial search acts as an automatically exhaustive search method that evaluates all of the possible combinations of property correlations. To effectively identify meaningful and nontrivial property correlations and correlation-groups, an evaluation criteria that combines the parameters of model performance (ψ), ANNs generalization ability (δ) and correlation error (E_c) is proposed. Then, a diverse set of correlations that was top-ranked by this new evaluation criterion is discussed.

3.2 Experiment

3.2.1 Data collection

All data are collected from CES [228]. Though errors can be found in the database, the error rate is considered to fall below 5%. Hence, the data quality of CES will not heavily bias the ANN modelling. Compared to organic or inorganic compounds, elements in the Periodic Table have most completed and reliable property data: there are total 35 properties recorded for the elements. Because not every element has a full 35 property record, a trade-off is made between the number of elements and the number of available property records. Finally, 37 elements, which are in solid state under 295±5 K

(except Hg), with a full record of 24 different properties, constitute the entire ANNs input dataset as listed in Appendix I.

The 24 properties chosen according to the author's interests and according to their availability in the database of CES are listed in Table 3-1. As can be seen, this study extends the earlier work [233] by searching for a much larger domain of property correlations, and introducing new evaluation criteria to evaluate the importance of observed property correlations (which is discussed in Section 3.2.3).

3.2.2 Input variables and output variables

Prior to the application of the ANN analysis, no data transformation was performed besides the procedures described in Chapter 2. Though it is acknowledged that taking logarithmic pre-treatment may result in models with better performances, the standard procedures (without cleaning, integration, etc.) worked sufficiently well for the purpose of this study. Because this experiment aims to find the strongest property correlations through the ANN combinatorial search, there is no need to remove highly correlated properties in the inputs.

In order to perform the ANN combinatorial search, the 24 properties listed in Table 3-1 were used in turns as the output variable of ANNs (i.e. the property to be predicted). In total, the ANNs searched three different orders of property correlations, namely, binary, ternary, and quaternary order property correlations. Each binary, ternary, and quaternary order of property correlation was constituted by one input variable, two input variables, and three input variables, respectively. Once the output variable was determined, input variable 'libraries' representing all possible combination were selected from the rest 23 properties. Since ANNs can be highly parallel distributed utilizing multiple computer cores, the advantages of employing ANNs to process the huge amount of data in the search for knowledge are evident.

Table 3-1 The 24 properties used in the ANN combinatorial search.

Property symbol	Property name	Unit
a	Lattice parameter, a	nm
A_r	Atomic weight	dimensionless
C_p	Specific heat capacity	J/kg·K
E	Young's modulus at 300 K	GPa
E_{coh}	Cohesive energy	kJ/mol
G	Shear modulus at 300 K	GPa
H_{fus}	Heat of fusion	kJ/mol
K	Bulk modulus at 300 K	GPa
r	Atomic radius	nm
T_b	Boiling point	°C
T_m	Melting point	°C
V_m	Molar volume	m^3/kmol
W	Work function	eV
T_r	T dependence of resistivity	Ω/°C
α_L	Thermal expansion coefficient at 300 K	μstrain/K
ΔH_{vb}	Heat of vaporization at the normal boiling point	kJ/mol
θ_D	Debye temperature	°C
λ	Thermal conductivity at 300 K	W/m·K
ν	Poisson's ratio	dimensionless
ρ_e	Electrical resistivity at 300 K	μohm·cm
σ_A	Neutron absorption cross section (0.025 eV)	barns
σ_S	Neutron scattering cross section (0.025 eV)	barns
Υ	Surface energy(liquid)	J/m^2
χ_m	Magnetic susceptibility	dimensionless

3.2.3 Evaluation criteria

One problem with the ANN combinatorial search is the extensive amount of results that need to be analysed. In this work, nearly 5000 ANN models have been created. A way to assess the correlation importance (i.e. how strong the correlation is) of those models

is in a deep need. Two criteria were applied to identify meaningful and nontrivial property correlations in this study:

1. Generalization ability $\boldsymbol{\delta}$. The generalization ability of an ANN model is an important factor that shows how well the correlation can be applied to new data, which is not used in the model construction. As discussed in Chapter 2, $\boldsymbol{\delta}$ can be assessed by analysing the model performance on the training data and the test data (see the discussion about Equation 2-5 and Equation 2-6 on page 56).

$$\delta = \left|\varphi^2_{training} - \varphi^2_{test}\right| \qquad \textit{Equation 2-5}$$

$$\varphi = |S_l - 1| + (1 - R_l) \qquad \textit{Equation 2-6}$$

2. Correlation error ($\boldsymbol{E_c}$). The validity of a property-correlation is reflected by its prediction accuracy ($\boldsymbol{A}$), which indicates how accurate the correlation is when it is applied to the given data. Because $\boldsymbol{E_c}$ equals (1-$\boldsymbol{A}$) when $\boldsymbol{A}$ is larger than zero, the higher accuracy means the lower prediction error. Thus, correlations that have the lowest prediction errors are perceived to show the best performances. On the other hand, the correlation accuracy could be heavily impaired by errors, and generally ~5% original data are subject to contamination [234, 235], the correlation error presented in this thesis always excluded the 5% data that have the largest errors.

Property correlations can be evaluated on single correlation basis or on a correlation-group basis. The concept of a correlation-group is illustrated in Figure 3-1. The ternary order correlation-group is constituted by three different correlations that utilizing the same three properties (X, Y and Z). For each correlation, one of the three properties is delegated as the output variable, and the other two properties are chosen to be the input variables. Similarly, the binary order correlation-group is constituted by two correlations that utilizing one of the two properties as the output variable, and the other property as the input variable. In addition, the quaternary order correlation-group is constituted by four correlations that utilizing one out of the four properties as the output variable, and the other three properties as the input variables.

It is acknowledged that root mean square of errors (RMSE) and coefficient of determination ($\boldsymbol{R_l}$) are the two modelling evaluation methods that widely adopted in the

literature [205]. Because the magnitude of average prediction error partly depends on the magnitude of experimental data, it makes no sense to compare the magnitude of error of different correlation models in predicting different properties. As for $\boldsymbol{R_l}$, it has already been taken into the account by the generalization criteria $\boldsymbol{\delta}$. Thus, it is not suitable to use the two parameters (RMSE and $\boldsymbol{R_l}$) as the overall evaluation criteria in this experiment.

The two criteria ($\boldsymbol{\delta}$ and $\boldsymbol{E_c}$) mentioned above are considered to be equally important and are assessed simultaneous in three different ways, where the completed picture of model performance is given by $\boldsymbol{\psi}$:

1) A linear superposition rule: $\psi = \delta + E_c$
2) A quadratic superposition rule: $\psi^2 = \delta^2 + {E_c}^2$
3) The rank: correlation-models are sorted based on $\boldsymbol{\delta}$ and $\boldsymbol{E_c}$ respectively. Therefore each correlation-model will have two ranking indexes $\boldsymbol{I(\delta)}$ and $\boldsymbol{I(E_c)}$. The final index of $\boldsymbol{\psi}$ is judged by $[\boldsymbol{I(\delta)} + \boldsymbol{I(E_c)}]$.

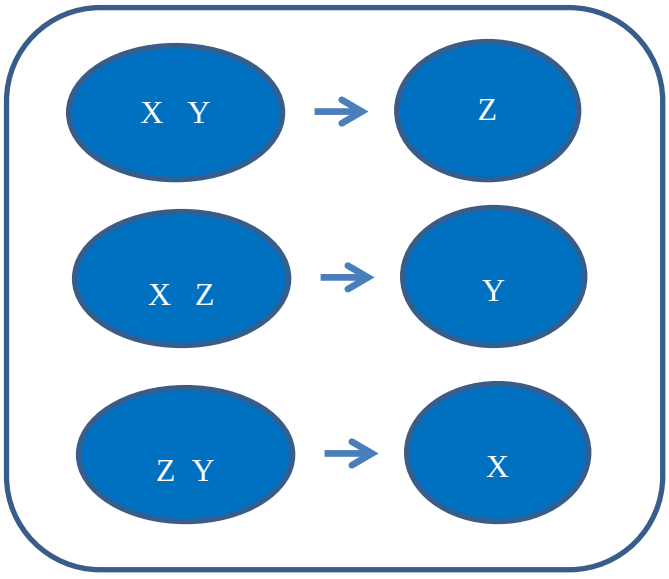


Figure 3-1 An example of a ternary order correlation-group that is constituted by three different correlations utilizing the property X, Y and Z. The three properties have been used as the output variable in turns, when the other two properties are used as the input variables.

As for a correlation-group constituted by $\boldsymbol{N}$ ($\boldsymbol{N}$=2, 3 or 4) correlations, the group performance ($\boldsymbol{\psi_N}$) is evaluated according to Equation 3-1, which is also assumed to follow a quadratic superposition rule. A good correlation-group means the involved properties have common response to each other.

$$\psi_N = \sqrt{\bar{\psi}^2 + sd(\psi)^2} \qquad \textit{Equation 3-1}$$

where $\bar{\boldsymbol{\psi}}$ and $\boldsymbol{sd}(\boldsymbol{\psi})$ is the average value and standard deviation of $\boldsymbol{\psi}$ for all the correlation-models ($\boldsymbol{\psi_1}, \boldsymbol{\psi_2}$ ... $\boldsymbol{\psi_n}$).

3.3 Results and discussion

In total, 552 binary order property correlations, 6072 ternary order property correlations and 42504 quaternary order property correlations have been created and investigated through the performing of ANN combinatorial search. For the correlations prioritized as the top 1% of each order, all the three assessments (proposed in Section 3.2.3) in evaluating individual property relationship reach very similar conclusions. Appendix II and Appendix VII tabulates the top 50 property correlations and the top 25 property correlation-groups of each order based on the quadratic superposition ANNs evaluation criteria. To enable potential usage for interested readers, the entire results are available on the author's webpage at www.researchgate.net.

ANN models with small correlation errors indicate the captured relationships have strong connections to nature laws. A large number of strong correlations are obtained through the combinatorial search, among which, some correlations have not been previously attended to, while some correlations have already been proposed in literature (especially for binary order correlations). It is beyond the framework of this study to discuss every strong correlation discovered by ANNs. In order to show the typical types of correlations that the ANN combinatorial search can find, two binary order correlation-groups and three ternary order correlation-groups are analysed in detail. In addition, the quaternary order property correlation between the work function and the confounding effect generated by surface energy, thermal conductivity, and lattice parameter $\boldsymbol{a}$, is also discussed.

3.3.1 Top binary order correlations

Appendix II and Appendix III list the top binary order correlations obtained by the ANN combinatorial search, including:

a) the relation between cohesive energy of solid and heat of vaporization;
b) the relation between atomic radius and molar volume (which is evident);
c) the relation between boiling point and heat of vaporization (which will be further discussed in Chapter 6);
d) the relation between shear modulus and Young's modulus (the ratio of shear modulus to Young's modulus generally equals to 3/8 [154]);
e) the relation between atomic weight and specific heat capacity.

To the author's best knowledge, the first correlation is hardly acknowledged in literature (though the correlation between cohesive energy of liquid metal and heat of vaporization is well known [236]), thus it is discussed emphatically in this section. While the relation between atomic weight and specific heat capacity, which is also known as the Dulong-Petit Law, is discussed in brief. The correlations captured by ANNs indicate that the ANN combinatorial search is a way of illuminating the facts, regardless of whether the facts have been discovered or not.

1) Cohesive energy and heat of vaporization at the normal boiling point

By definition, the heat of vaporization at normal boiling point (ΔH_{vb}) is the energy per mol required to transform a substance from liquid into a gaseous state under the standard pressure (101kPa) at its normal boiling point [208]. The average discrepancy of ΔH_{vb} between CES data [228], which is used in the ANN combinatorial search, and the data recently corrected by Zhang *et al.* [237], is about 4.2%. A more detailed discussion about ΔH_{vb} can be found in Chapter 6. Among the several thermal properties such as melting point, heat of fusion, thermal expansion coefficient, and Debye temperature, the ANN combinatorial search indicates that ΔH_{vb} is most strongly correlated with boiling point (T_b). There are some empirical equations available to make a fast estimation of ΔH_{vb} from T_b, for example, the equation proposed by Kistiakowsky [208].

The cohesive energy (E_{coh}) is the energy per mole must be added to completely break the bonding of atoms of a solid at 0 K at 1 atm, separating the condensed materials into

isolated neutral free atoms, which have the same electronic configurations, at infinity (the gaseous state) [238, 239]. $\boldsymbol{E_{coh}}$ is different from lattice energy as the latter is referring to ionic solid crystals [238].

The value of $\boldsymbol{E_{coh}}$ offers an insight into the bonding strength between atoms in the solid materials [239]. As can be seen from the values of $\boldsymbol{E_{coh}}$ for metals that are listed in Appendix I, the alkali metal crystals have the smallest values of the cohesive energy (~100 kJ/mol), because they are relatively weakly bonded. The transition element metals are strongly bounded, as they have the largest values (400 ~ 900 kJ/mol). While the alkaline earth metals and the poor metals, such as Al, Zn, Cd, In, Sb, Hg, Tl, Pb, and Bi, have intermediate values of $\boldsymbol{E_{coh}}$ (100 ~ 200 kJ/mol).

It should be noted that $\boldsymbol{E_{coh}}$ mentioned in this work is for bulk materials rather than the corresponding nanoparticles. Since atoms on the surface of nanoparticles are incompletely bonded [238], $\boldsymbol{E_{coh}}$ of nanoparticles is much lower than that of the corresponding bulk materials [240]. The particle size dependence also holds for melting point and surface tension [241, 242]. Without any further notice, all the properties mentioned here are for bulk materials.

There is no direct measurement to obtain the absolute value of $\boldsymbol{E_{coh}}$ [243]. Experimental results mentioned in literature can usually be traced back to the values compiled by Kittel [238] and Brewer [244]. In fact, Kittel claimed his data was supplied by Brewer. The average difference of $\boldsymbol{E_{coh}}$ between CES data, which were used in the ANN combinatorial search, and the data provided by Kittel is less than 1%. For pure crystals, Farid and Godby [245] summarised the three experimental ways to obtain $\boldsymbol{E_{coh}}$: i) by measuring the heat of sublimation at various temperatures, and extrapolating the sublimation curve in the pressure-temperature plane to zero Kelvin, ii) by determining of the heat of reaction based on the measurements of the equilibrium constant, and ii) by calculating from thermodynamic tables at crystal phase ($\boldsymbol{cr}$) and gas phase ($\boldsymbol{g}$) through Equation 3-2 [245].

$$\Delta_f H^o_{0\,K} = \Delta_f H^o_{298.15\,K} + (H^o_{0\,K} - H^o_{298.15\,K})_g - (H^o_{0\,K} - H^o_{298.15\,K})_{cr} \qquad \textit{Equation 3-2}$$

where $\boldsymbol{\Delta_f H^o_T}$ is the standard-state enthalpy of formation of the gas from the crystal at temperature $\boldsymbol{T}$, and $\boldsymbol{H^o_T}$ is the standard-state enthalpy at temperature $\boldsymbol{T}$.

There are a number of studies in the correlation of cohesive energy with other materials properties. Because E_{coh} provides a measure of the intermolecular forces of metals, it is found to increase linearly with the melting point (T_m) [239] and the bulk modulus (B, the inverse of compressibility) [246]. An empirical inverse relationship between cohesive energy and thermal expansion coefficient (α_L) has been observed by Tsuru *et al.* [239] for metals and ceramics, and by Gangopadhyay *et al.* [247] for liquid transition metal alloys.

Unfortunately, Tsuru *et al.* [239] gave no data on how accurate the relationships were. Nevertheless, the average error of E_{coh} between the experimental data [228] and the predicted values by ANNs from T_m or α_L in the present work is in the range of ±13% and ±20%, respectively. The correlation coefficient R_l between bulk modulus and cohesive energy is 0.876 in ANN models, in contrast to 0.817 in the empirical model proposed by Tamura [246]. In another words, the ANN combinatorial search has also evaluated those correlations proposed in literature and determined they are less attractive based on the criteria described in Section 3.2.3.

However, the correlation between the cohesive energy of the solid and the heat of vaporization at the normal boiling point, obtained by the ANN combinatorial search, is somewhat surprisingly strong: the ANN prediction accuracy is about 97%. The cohesive energy of the solid equals the heat of sublimation at absolute zero temperature [245], which is the sum of heat of fusion, the heat to raise material to vaporization temperature, and the heat of vaporization [248]. The heat of vaporization is equivalent to the cohesive energy of the liquid [247], the magnitude of which should be smaller than the cohesive energy of the solid. Hence it is quite interesting to find the cohesive energy of the solid can be accurately predicted by simply using the heat of vaporization.

As shown in Figure 3-2 (a), the ANN predicted values ΔH_{vb} from E_{coh} are plotted as the ordinate against ΔH_{vb} experimental values as abscissa for the 37 metallic elements, and in Figure 3-2 (b), the ANN predicted values E_{coh} from ΔH_{vb} (Y axis) are plotted against E_{coh} experimental values (X axis). The deviation from the straight line is <10% for most of the solids, except for Ba. Nevertheless, ΔH_{vb} and E_{coh} are highly correlated as the ANN predictions agree very well with the experimental data with an average correlation error less than 2.5%, i.e. the average prediction accuracy is higher than 97.5%.

Because the relationship between ΔH_{vb} and E_{coh} is a binary order correlation, it is possible to draw a 2-D plot of ΔH_{vb} against E_{coh} using experimental values (Figure 3-3). As can be seen, E_{coh} increases with the increases of ΔH_{vb}, except for two elements: Zr and Sb. The distribution of experimental data points seems to follow a straight line. A linear regression fit gives $E_{coh} = 1.0476\Delta H_{vb} + 11.731$ with a high value of correlation factor R_l (0.998) and low average error (±3.29%), where both properties are in kJ mol^{-1}. Like the correlations found by ANNs, it also indicates Ba as an outlier with a deviation of 10.34%. Unfortunately, no explanation is available to justify the abnormal behaviour of Ba.

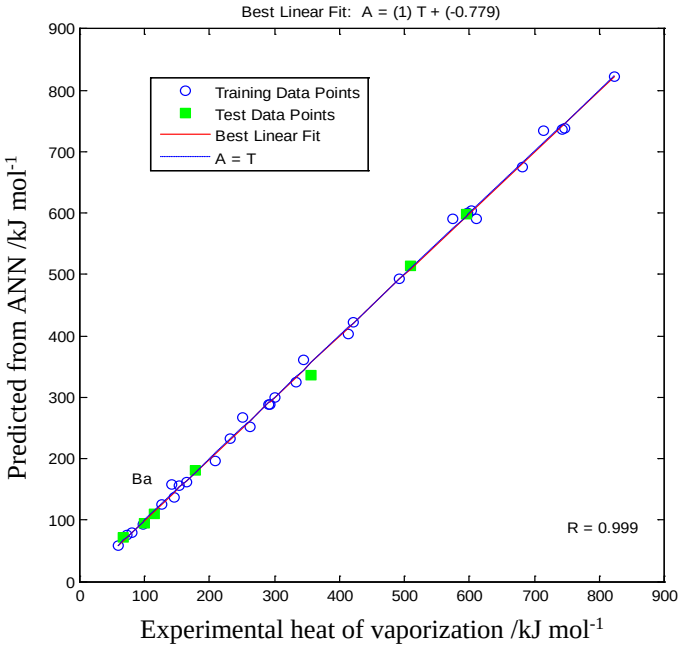
As previously stated, the cohesive energy of the solid is equal to the sum of heat of fusion and the heat of vaporization, thus it is easily to understand the increasing E_{coh} with the rise of ΔH_{vb}. It is reasonable to assume that for each element, its E_{coh} is larger than its ΔH_{vb}. However, E_{coh} is a bit smaller than ΔH_{vb} for Zr (E_{coh}=610 kJ/mol, ΔH_{vb} =611 kJ/mol) and Sb (E_{coh}=261 kJ/mol, ΔH_{vb} =262 kJ/mol) in CES [228], thus the reliability of the two data-pairs is questioned. To address this question, both data (E_{coh} and ΔH_{vb}) are carefully checked. Results show very small variances of E_{coh} for Zr (1.1%) and Sb (1.5%) in CES database as compared to the values tabulated by Kittel [238]. But large disagreements of ΔH_{vb} are found for Zr and Sb, and the detailed investigation is shown below.

If the value of E_{coh} data for Zr is reliable, and the correlations between E_{coh} and ΔH_{vb} are accurately captured by the ANN combinatorial search, the value of ΔH_{vb} for Zr can be obtained from the established ANN model using E_{coh} as the input. The predicted ΔH_{vb} for Zr is 590.6 kJ/mol, and it agrees well with the data given by Zhang *et al.* [237] (591 kJ/mol), and the data compiled by Hultgren *et al.* [249], (139112 cal/g-atom, which is corresponding to 582 kJ/mol).

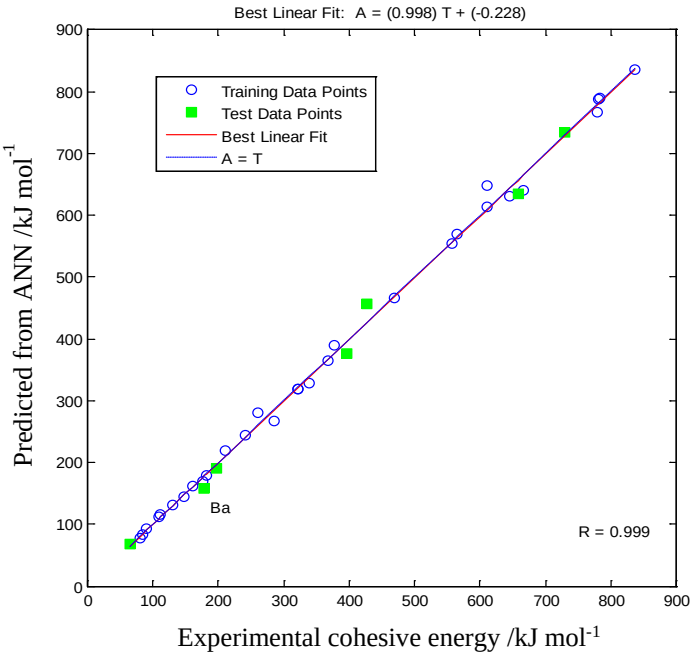
It is more complicated for Sb. If the value of E_{coh} for Sb is reliable, the predicted ΔH_{vb} for Sb is 251.1 kJ/mol, which differs by more than 10% from the value given by Zhang *et al.* [237] (193 kJ/mol). In fact, Zhang *et al.* [7, 237] had to choose either 195 kJ/mol [250], 193 kJ/mol [112] or 67.9 kJ/mol [251] from the limited sources they checked, and suggested a smaller value (176 kJ/mol) might be more appropriate. The value ~193 kJ/mol seems to be first published in Lange's Handbook of Chemistry 10th edition [252], which gave 46.63 cal/g-atom (195 kJ/mol), and changed to 46.23 kcal/mol (193 kJ/mol)

in its later editions [112]. While the value of 67.9 kJ/mol can be traced back to the data compiled by Sinke [253], where it is claimed that for Sb "the heat of vaporization of 1 gram atomic weight at 1910 K to the equilibrium vapour is 16230 cal" (which corresponds to 67.9 kJ/mol). So it is highly possible that ΔH_{vb} is wrongly recorded in Lange's Handbook of Chemistry (should be 16.23 kcal/mol rather than 46.23 kcal/mol for atomic Sb).

At lower temperatures, the substance almost entirely exists in the molecular form of Sb_4. The proportion of Sb_2 and Sb in the gas increases when the temperature is approaching to the boiling point (1860 K), and the dominant species are believed to be Sb_4 and Sb_2 at 1860 K [249]. If Sb_3 is used to represent the state of gas in combination of Sb_4 and Sb_2, the corresponding ΔH_{vb} accounted for the molecular form Sb_3 would be 203.7 kJ/mol. But there is more Sb_2 than Sb_4 in the gas when the temperature is equal to the boiling point [249], and the existing of Sb should not be neglected. So the real ΔH_{vb} accounted for the molecular form of Sb_x ($X \in [1,4]$) should be smaller than 203.7 kJ/mol, but larger than 67.9 kJ/mol. The deviation of ΔH_{vb} predicted by ANNs for Sb (251.1 kJ/mol) in contrast to the real value is possibly due to using the less accurate value of E_{coh} for Sb. Indeed, Brewer [244] used the thermodynamic data compiled by Hultgren *et al.* [249] to calculate E_{coh}, but it is noted that the vapour pressure data below 900 K were evaluated as Sb_4 (neglecting Sb_2 and Sb) [249]. In this case, the corresponding ΔH_{vb} accounted for the molecular form Sb_4 would be 271.6 kJ/mol, which is very close to the ANN prediction in this experiment.



(a)



(b)

Figure 3-2 The correlation between the cohesive energy of the solid (E_{coh}) and the heat of vaporization at normal boiling point (ΔH_{vb}) obtained by the ANN combinatorial search for 37 metallic elements: a) Predict ΔH_{vb} from E_{coh}, both experimental (X axis) and predicted data (Y axis) are for ΔH_{vb}, and the correlation error (E_c) is 2.2%; b) Predict E_{coh} from ΔH_{vb}, both experimental (X axis) and predicted data (Y axis) are for E_{coh}, and the correlation error is 2.5%.

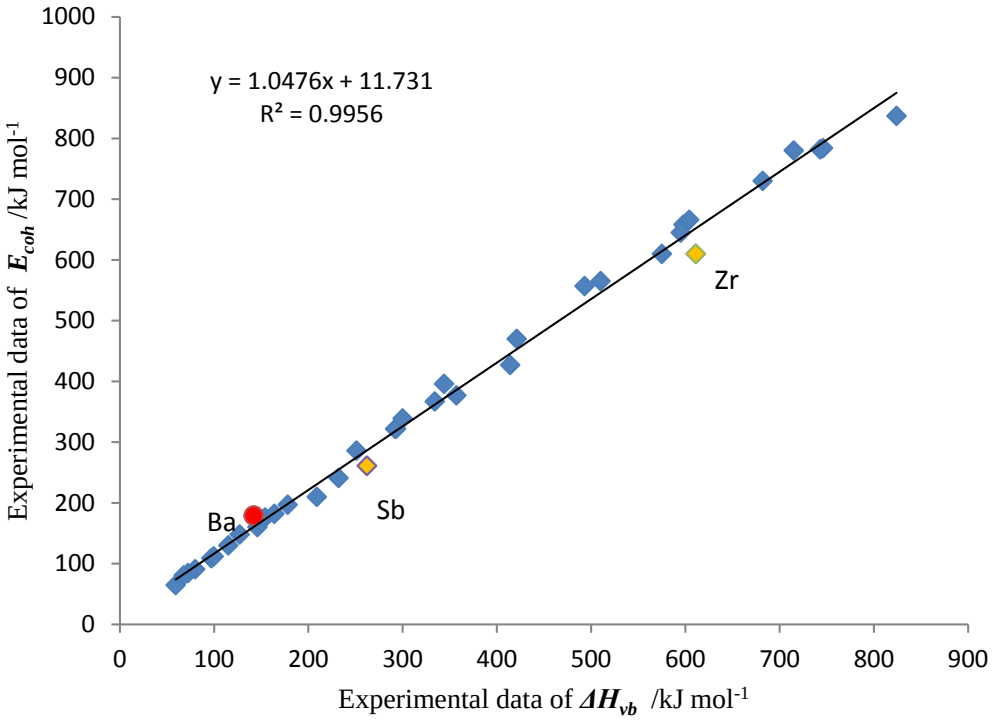


Figure 3-3 Linear correlation between the cohesive energy of the solid (E_{coh}) and the heat of vaporization at normal boiling point (ΔH_{vb}).

2) Atomic weight and specific heat capacity

The atomic weight (A_r) is the ratio of the average mass per atom of the element to 1/12 of the mass of neutral atom of carbon-12, e.g. $A_r(^{12}C) = 12$ [254]. A recommended table of atomic-weight values is published annually by International Union of Pure and Applied Chemistry (IUPAC). The average difference between the most recent IUPAC data [254] and the data recorded in CES [228] is less than 0.1%. The specific heat capacity (C_p) used here is the energy required to heat 1 gram of a material by 1 K at constant pressure (atmospheric pressure) [228]. For solids, C_p is very close to C_v, i.e. the specific heat capacity measured at constant volume. C_p data in CES are compared with the values recorded in CRC Handbook [229], and the average difference is 1.03%. Hence, the input data accuracy should have a very small impact on the correlations captured by ANNs.

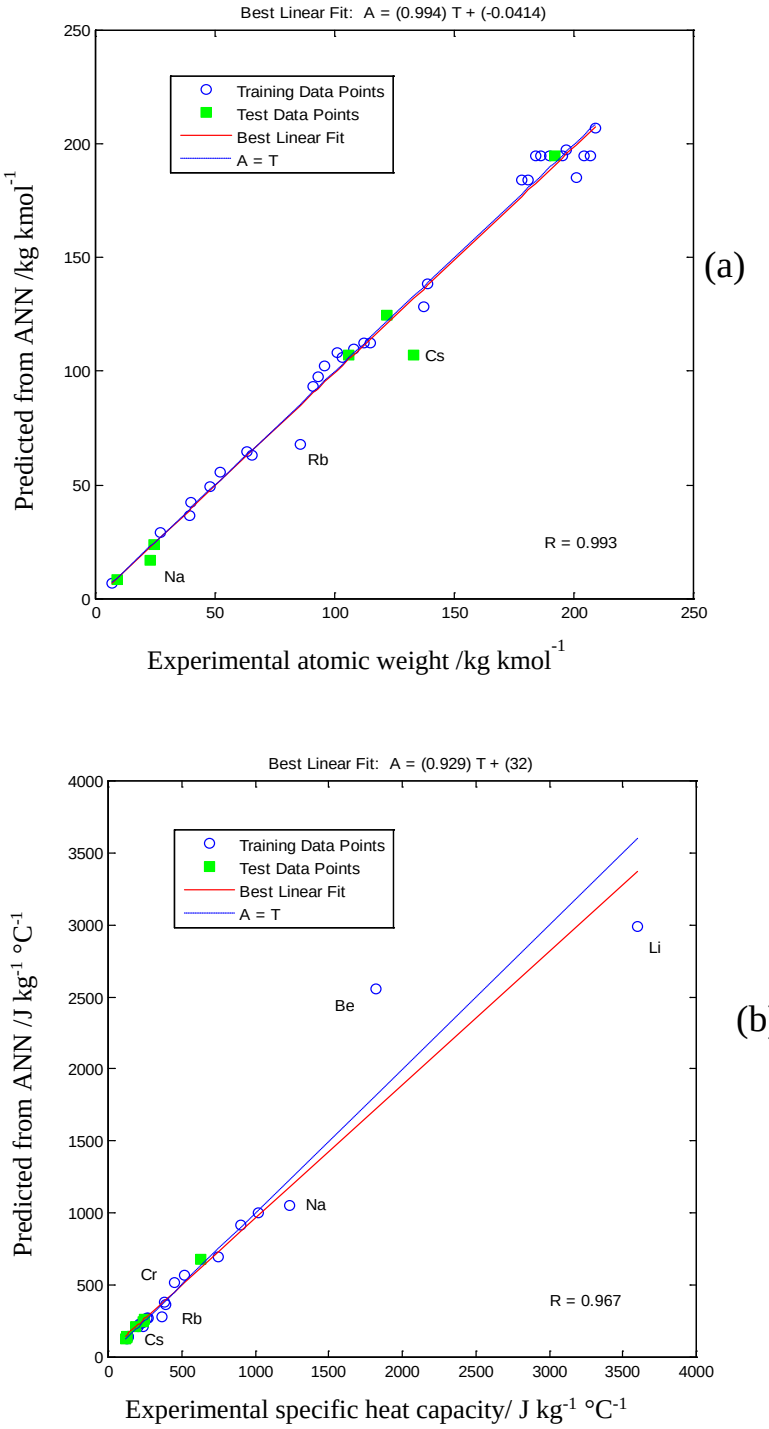


Figure 3-4 The correlation between the atomic weight ($\mathbf{A_r}$) and specific heat capacity ($\mathbf{C_p}$) obtained by the ANN combinatorial search for 37 elements: a) Predict $\mathbf{A_r}$ from $\mathbf{C_p}$, both experimental (X axis) and predicted data (Y axis) are for $\mathbf{A_r}$, and the correlation error is 4.3%; b) Predict $\mathbf{C_p}$ from $\mathbf{A_r}$, both experimental (X axis) and predicted data (Y axis) are for $\mathbf{C_p}$, and the correlation error is 4.7%.

As shown in Figure 3-4, the ANN combinatorial search indicates atomic weight and specific heat capacity are strongly correlated, as the coefficients of determination $\boldsymbol{R_l}$ for both predictions are higher than 0.96, with the average correlation errors less than 5%. Among the 37 elements, beryllium and lithium have the largest prediction errors, possibly because they have the highest values in $\boldsymbol{C_p}$ and lowest values in $\boldsymbol{A_r}$, which leave them on the prediction domain edge of the ANN, therefore the predictions are less accurate.

The relationship between atomic weight and specific heat capacity can be interpreted by the correlation between molar mass and specific heat capacity. For the 37 elements used in the ANN combinatorial search, their molar masses are equal to their atomic weights multiplied by the molar mass constant (i.e. 1 g mol^{-1}), assuming that the ideal crystalline solid is built of monatomic unit cells. Such relation might be enunciated as the Dulong-Petit Law that "the atoms of all simple bodies have exactly the same capacity for heat" [255]. In modern terms, the law states the molar heat capacity ($\boldsymbol{C_m}$) of solid elements is a constant [256]:

$$C_m = MC_p = 3R = 6\ cal\ mol^{-1}K^{-1} = 25\ J\ mol^{-1}K^{-1} \qquad \textit{Equation 3-3}$$

where

$\boldsymbol{C_m}$ — molar specific heat capacity, J mol^{-1} K^{-1};

$\boldsymbol{M}$ — molar mass; for a solid metallic element, it equals to $\boldsymbol{A_r}$, g mol^{-1};

$\boldsymbol{C_p}$ — mass specific heat capacity, J g^{-1} K^{-1};

$\boldsymbol{R}$ — universal gas constant, which equals to 8.314 J mol^{-1} K^{-1}.

The relationship between atomic weight and specific heat capacity is attributed to lattice vibrations. A crystal has three modes of vibration per atom (one longitudinal mode and two transverse modes), each corresponding to a quadratic kinetic energy term and a quadratic potential energy term [256-258]. Similar to the ANN predictions, a significant deviation (~50%) is also observed for beryllium under the Dulong-Petit Law.

Besides the Dulong-Petit Law, Schwarz [259] proposed an empirical non-linear function (Equation 3-4) to correlate molar specific heat capacity (at 25˚C) with atomic weight for each element family in the Periodic Table:

$$C_p = \frac{1}{C_1 \times A_r + C_2} \qquad \textit{Equation 3-4}$$

where $\boldsymbol{C_p}$ is in the unit of cal g^{-1} K^{-1}, and C_1 and C_2 are constants for different element families in the Periodic Table, e.g. for alkali-metal family, C_1 =0.1319, C_2=0.27.

It is worth to point out that the correlation captured by ANNs is constrained by the characteristics of input data. In this work, $\boldsymbol{C_p}$ data are considered to be temperature-independent values obtained at room temperature, thus the correlation will not reflect the empirical behaviour that specific heat capacity of solids decrease exponentially at low temperatures, i.e. $\boldsymbol{C_p}(T) \propto \boldsymbol{T}^3$, when $\mathbf{T} \rightarrow 0$.

3.3.2 Top ternary order correlations

By analysing the top ternary order property correlations, three types of ternary order correlations were obtained by the ANN combinatorial search, if X and Y represent the two explanatory properties (input variables) and Z is the response property (output variable):

i. Z can be accurately predicted by either X or Y, but X and Y produce a small confounding effect to improve the prediction accuracy. For example, the ternary order correlation between cohesive energy, boiling point and heat of vaporization;
ii. Z is correlated to X or Y to a certain degree, and the accuracy to predict Z is significantly improved due to the confounding effect of X and Y. For example, the ternary order correlation between heat of vaporization, surface energy and molar volume;
iii. Z is poorly correlated to the single variable X or Y, but Z can be accurately predicted due to the confounding effect of X and Y. For example, the ternary order correlation between shear modulus, bulk modulus and Poisson's ratio.

1) Cohesive energy, boiling point and heat of vaporization

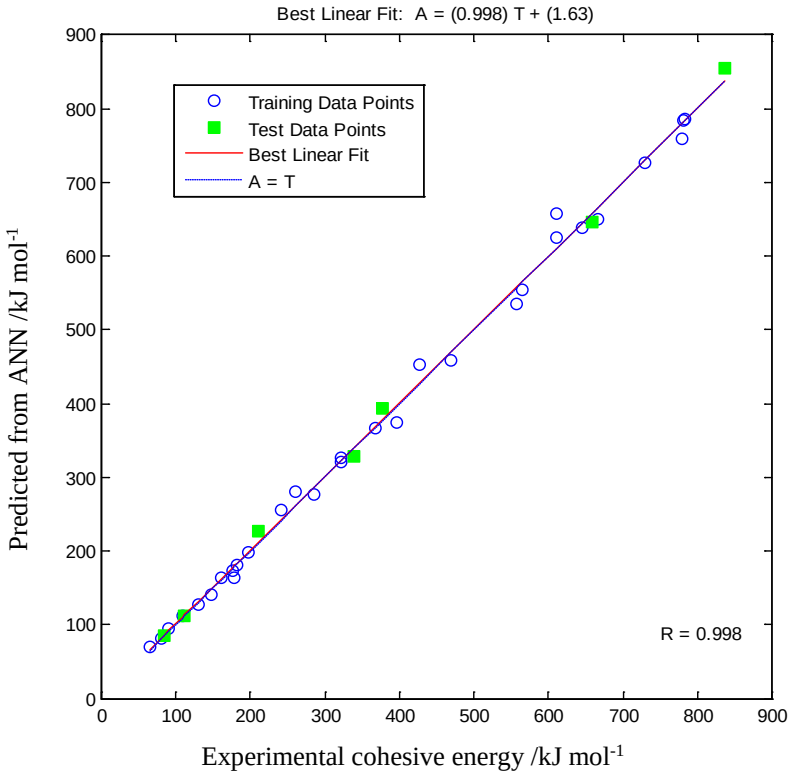
As discussed in Section 3.3.1, there is a good relationship between cohesive energy and heat of vaporization, and between heat of vaporization and boiling point. In fact, the ANN combinatorial search also indicates cohesive energy is highly correlated with boiling point (see Appendix II and Appendix III). Hence, it is not surprising to find that

cohesive energy, heat of vaporization and boiling point constitutes a good ternary order correlation-group, i.e. each property can be predicted by the combination of other two properties at a reasonable accuracy.

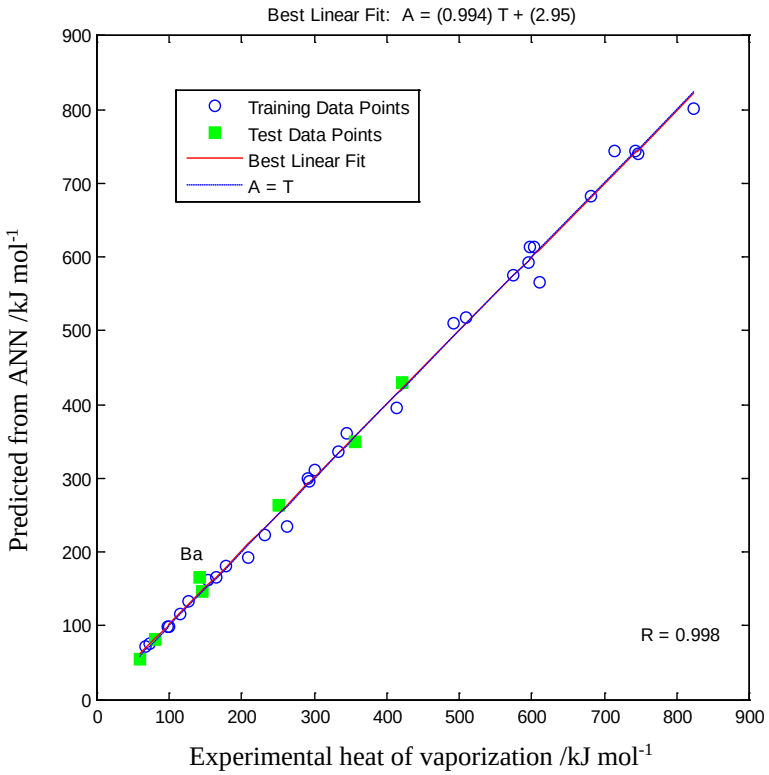
The ternary order correlation-group is illustrated by Figure 3-5. It shows the average accuracy to predict E_{coh} from the other two properties is ~97%, which is similar to the average accuracy to predict ΔH_{vb}, but a bit higher than the average accuracy to predict T_b (~93%). As shown in Table 3-2, the influence of T_b seems to be very small in the prediction of E_{coh} or ΔH_{vb}, because the ANN evaluation parameters are almost the same when the prediction is made with or without T_b. Table 3-2 also indicates that E_{coh} and ΔH_{vb} are equally important in predicting T_b.

Table 3-2 The binary and ternary order correlations between cohesive energy, boiling point and heat of vaporization.

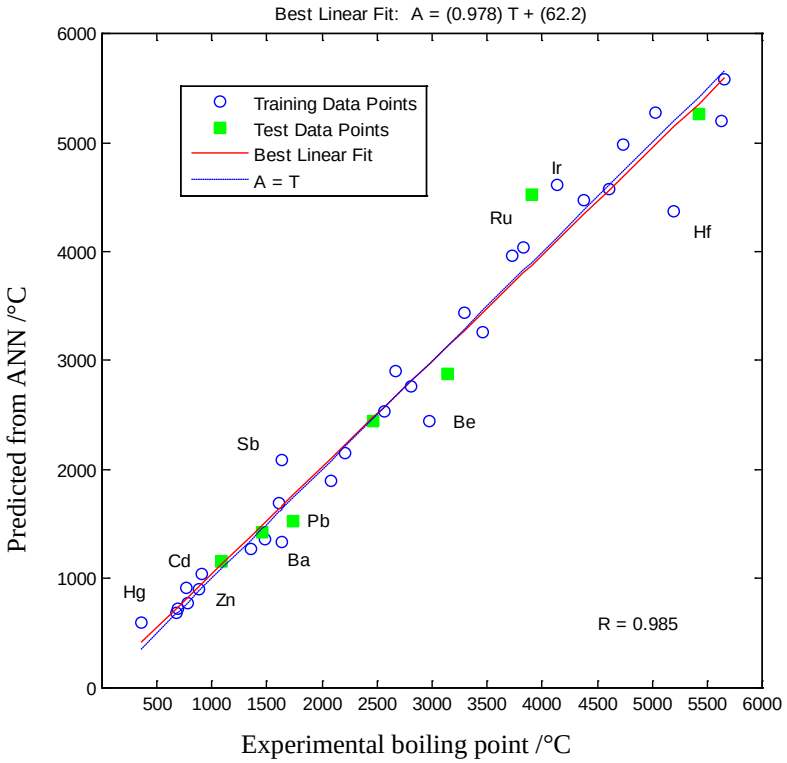
Conditions		**ANNs evaluation parameters**		
Predicted property	**Input property**	ψ	δ	E_c
ΔH_{vb}	E_{coh}	0	0.001	2.2%
	T_b	0.006	0.024	7.5%
	E_{coh} and T_b	0.001	0.008	2.8%
E_{coh}	ΔH_{vb}	0.001	0.003	2.5%
	T_b	0.005	0.024	6.9%
	ΔH_{vb} and T_b	0.001	0.004	2.7%
T_b	E_{coh}	0.005	0.026	6.9%
	ΔH_{vb}	0.005	0.024	6.9%
	E_{coh} and ΔH_{vb}	0.006	0.037	6.9%


Best Linear Fit: A = (0.998) T + (1.63)
Training Data Points
Test Data Points
Best Linear Fit
A = T
R = 0.998
Predicted from ANN /kJ mol-1
Experimental cohesive energy /kJ mol-1
0
100
200
300
400
500
600
700
800
900

(a)


Best Linear Fit: A = (0.994) T + (2.95)
Training Data Points
Test Data Points
Best Linear Fit
A = T
Ba
R = 0.998
Predicted from ANN /kJ mol-1
Experimental heat of vaporization /kJ mol-1
0
100
200
300
400
500
600
700
800
900

(b)



(c)

Figure 3-5 The correlation between cohesive energy, heat of vaporization, and boiling point, which is obtained through the ANN combinatorial search: a) Predict E_{coh} *from* ΔH_{vb} *and* T_b*, both experimental (X axis) and predicted data (Y axis) are for* E_{coh}*, and the correlation error is 2.7%; b) Predict* ΔH_{vb} *from* E_{coh} *and* T_b*, both experimental (X axis) and predicted data (Y axis) are for* ΔH_{vb} *and the correlation error is 2.8%; c) Predict* T_b *from* E_{coh} *and* ΔH_{vb}*, both experimental (X axis) and predicted data (Y axis) are for* T_b *and the correlation error is 6.9%.*

2) Heat of vaporization, surface energy and molar volume

Surface energy Υ is an important property to characterize materials processing behaviours, such as welding and sintering [260]. It quantifies the energy that is necessary to create new surfaces by unit area [261], and is also defined as the excess energy at the surface of material compared to the bulk [262]. It is reported that the surface energy of liquid metals is a temperature dependent property that decreases with increasing temperature [263]. At a relative small temperature interval (100 K), the surface energy is perceived to be a constant [264].

The surface energy data used in the ANN combinatorial search is for liquid metals, where the surface energy (J/m^2) is equivalent to the surface tension (N/m). Surface tension can be measured by the oscillating drop method using levitation under microgravity (see Equation 3-5) [265]. The total CES surface energy data [228] has an average error of ~1.6% comparing to the experimental results recorded in Ref. [260].

$$\gamma = \frac{3}{8}\pi f^2 m \qquad \textit{Equation 3-5}$$

where $\boldsymbol{f}$ is the oscillation frequency and $\boldsymbol{m}$ is the mass of the droplet.

Table 3-3 summarises all the six binary order correlations and the three ternary order correlations between heat of vaporization, surface energy and molar volume. As can be seen, the correlations between any of the two properties are relatively weak, with the correlation errors ranging from 14.3% to 52.1% (the average error of the six binary order correlations is 29.5%). However, the correlation-group constituted by the three ternary order correlations, has a substantially lower average correlation error of 11.7%, indicating there are stronger correlations owing to the confounding effects of any two properties.

The ANN results are in line with the empirical relationship between heat of vaporization ($\boldsymbol{\Delta H_{vb}}$), surface energy ($\boldsymbol{\Upsilon}$) and molar volume ($\boldsymbol{V_m}$), showing in Equation 3-6 [266]. A more complicated equation linking heat of vaporization, surface energy, molar volume, and a self-defined structure factor, was derived by Overbury *et al.* [264].

$$\boldsymbol{\gamma} = \frac{\Delta \boldsymbol{E}}{C\boldsymbol{V_m}^{2/3}} = \frac{\Delta \boldsymbol{H_{vb}} - P\Delta V}{C\boldsymbol{V_m}^{2/3}} \approx \frac{\Delta \boldsymbol{H_{vb}} - \boldsymbol{RT}}{C\boldsymbol{V_m}^{2/3}} \qquad \textit{Equation 3-6}$$

where:

$\boldsymbol{\Upsilon}$ = surface tension in dyn/cm;

$\boldsymbol{V_m}$ = molar volume in cm^3;

$\boldsymbol{\Delta E}$ = internal energy in cal;

$\boldsymbol{R}$ = universal gas constant (1.986 cal K^{-1} mol^{-1});

$\boldsymbol{T}$ = absolute temperature in K;

C = a constant (~14).

Because heat of vaporization and cohesive energy (of solids at 0 K) is highly correlated (as discussed in Section 3.3.1), it is not surprising that ANNs also find a correlation between cohesive energy, molar volume and surface energy (see Appendix V). In fact, this correlation was first deployed by Stephan in 1886 (see Equation 3-7), to transition metals [260], as both properties reflect the energy required to break the bonds of atoms. It is worth to mention that the result of the ANN combinatorial search indicates such correlation is applicable not only for transition metals but also for some other metals.

$$\gamma = \frac{C_t E_{coh}}{{V_m}^{2/3}} \qquad \textit{Equation 3-7}$$

where C_t is a constant that is rarely theoretically determined.

Table 3-3 The binary and ternary order correlations between heat of vaporization, surface energy and molar volume.

Conditions		**ANNs evaluation parameters**		
Predicted property	**Input property**	ψ	δ	E_c
ΔH_{vb}	γ	0.074	0.186	19.8%
	V_m	0.997	0.852	52.1%
	γ and V_m	0.018	0.051	12.4%
γ	ΔH_{vb}	0.081	0.16	23.6%
	V_m	0.249	0.422	26.6%
	V_m and ΔH_{vb}	0.014	0.017	11.6%
V_m	γ	0.027	0.081	14.3%
	ΔH_{vb}	0.888	0.863	37.9%
	γ and ΔH_{vb}	0.014	0.039	11.1%

3) Shear modulus, bulk modulus, and Poisson's ratio

Shear modulus G, bulk modulus K, and Poisson's ratio ν are fundamental mechanical properties that are used to describe materials' elastic behaviour. The definitions, the common measurement methods, and the internal correlations of the three properties are

discussed in Chapter 1 together with Young's modulus. Here a brief summary is provided.

When a material is subjected to elastic deformation: the shear modulus is defined as the ratio of the shear stress to the shear strain on the loading plane; the bulk modulus quantifies a material's resistance to change in volume under a pressure; and Poisson's ratio is the negative ratio of the lateral (transverse) strain to the axial strain in the loading direction [149]. The three elastic properties reflect the strength of the interatomic forces bonding the material, and can be measured by static tests or dynamic tests [164]. For anisotropic materials (e.g. single crystals), the values of the three properties depend on the directions in measurements, and the average data can be derived from the single crystal elastic constants through different approximation methods, such as the Voigt-Reuss-Hill average method [156].

To investigate the confounding effect of shear modulus, bulk modulus, and Poisson's ratio, Table 3-4 tabulates the three ternary order correlations as well as the six binary order ANN results. In total, there are nine pairs. Comparing the binary and ternary order correlations, it is found that all the six binary order correlations have very high values of total model performance ($\boldsymbol{\psi}$), ANNs generalization ability ($\boldsymbol{\delta}$) and correlation error $\boldsymbol{(E_c)}$, thus they are perceived to be weak correlations, i.e. one property is poorly related to the other property. In particular, it is beyond the ability of ANNs to predict the values of bulk modulus or shear modulus merely based on the data of Poisson's ratio, as the prediction errors are as large as ~200%. This is partly due to the fact that for crystalline metals, the magnitudes of $\boldsymbol{G}$ and $\boldsymbol{K}$ have significant variances, but their $\boldsymbol{\nu}$ normally lies in the range of 1/4 to 1/3 [153].

However, the prediction ability of ANNs is significantly improved by adding one more property. Comparing the six binary order correlations and the three ternary order correlations, the evaluation parameters, i.e. model performance ($\boldsymbol{\psi}$) and ANNs generalization ability ($\boldsymbol{\delta}$) and correlation error ($\boldsymbol{E_c}$) are substantially reduced from an average of 2.55, 0.63 and 99% to an average of 0.02, 0.03 and 9.2%, respectively. This means that good ternary order relationships exist between the three elastic properties. Indeed, such relationships can be explained by Equation 1-21, where each property can be expressed in terms of any other two properties:

$$G = \frac{3K(1-2v)}{2(1+v)}$$

$$K = \frac{2G(1+v)}{3(1-2v)}$$

$$v = \frac{3K-2G}{2(3K+G)}$$

The ANN evaluation parameters of the ternary order correlations between shear modulus, bulk modulus and Poisson's ratio are small enough to rank them as a top correlation group. However, as the validated correlating equations exist (Equation 1-21), the relative large errors in predicting ***G*** (18.0% error) and ***K*** (7.5% error) are very unusual, especially considering that two data with the largest prediction errors have been excluded in the calculation of the ANN correlation error. Even using Equation 1-21, an average error of 9.1% for ***G***, 8.42% for ***K*** and 4.56% for ***v*** are obtained (two data with the largest error have also been excluded in the calculation).

Table 3-4 The binary and ternary order correlations between shear modulus, bulk modulus and Poisson's ratio.

Conditions		**ANNs evaluation parameters**		
Predicted property	**Input property**	ψ	δ	E_c
K	***v***	6.72	0.94	241.5%
	G	0.18	0.11	40.4%
	G and ***v***	0.01	0.01	7.5%
G	***K***	0.34	0.14	56.9%
	v	6.02	0.75	233.5%
	K and ***v***	0.03	0.01	18.0%
v	***K***	1.75	1.32	11.7%
	G	0.30	0.54	9.8%
	G and ***K***	0.01	0.08	2.1%

In this case, doubts are cast on the values of the three elastic properties (plus Young's modulus) recorded in CES. It is suspected that the large errors are caused by using incorrect data to train ANNs. As shown in Chapter 4, the assumption of 5% incorrect data underestimates the overall errors in the values of elastic properties. Thus,

correlations found within the ANN combinatorial search involving any elastic property will not be further discussed in this chapter due to the lack of confidence in data.

3.3.3 Top quaternary order property correlations

Because the ANN combinatorial search tests all possible property combinations, the strong correlations obtained may have redundant variables, especially for property correlations of higher order. As shown in Appendix VI, almost all the top 50 quaternary order property correlations can be simplified to be the correlation between atomic weight and specific heat capacity, where the other two properties play minor roles in the ANN models. By comparing the quaternary order property correlations with the binary and ternary order property correlations, quaternary order property correlation-groups are identified into four categories.

For a quaternary order property correlation-group constituted by (X_1, X_2, X_3 and Z), where X_i represents the possible explanatory property (input variable) and Z is the response property (output variable):

i. Z is only correlated to X_1, while X_2 and X_3 have little influence on Z. For example, the quaternary order correlation-group contains cohesive energy, heat of vaporization, atomic weight, and specific heat capacity. When Z is taken to be the cohesive energy, it can be explained by the heat of vaporization as discussed in Section 3.3.1, where the atomic weight and the specific heat capacity play no role in predicting the cohesive energy. When Z is taken to be the specific heat capacity, it can be well predicted by the atomic weight, where the cohesive energy and the heat of vaporization have very small impacts in the prediction.
ii. Z is related to X_1, and is correlated to X_2 and X_3 due to their confounding effect. For example, the quaternary order correlation-group constituted by heat of vaporization, cohesive energy, surface energy and molar volume. It can be considered as a binary order correlation between heat of vaporization and cohesive energy, added to a ternary order correlation between heat of vaporization, surface energy, and molar volume.
iii. Z is correlated to any two independent properties, which means a ternary relationship exists between any three of the four properties. For example, the

quaternary order correlation-group constituted by Young's modulus, shear modulus, bulk modulus, and Poisson's ratio.

iv. X_1, X_2 and X_3 all contribute to the correlation with Z, among which some variables may have larger impacts than the others. Although such quaternary order property correlation-group is rare in the present work, some interesting quaternary order property correlations are found that belong to this type. One typical quaternary order property correlation that is using surface energy, thermal conductivity, and lattice parameter ***a*** to predict work function, is discussed with more details.

1) Surface energy, thermal conductivity, lattice parameter *a* and work function

The thermal conductivity ($\boldsymbol{\lambda}$) measures the rate at which heat will flow through a solid per unit time across a surface with area ***As***, and it can be calculated by Equation 3-8 [267]:

$$\boldsymbol{\lambda} \equiv \frac{\boldsymbol{Q_h}}{\boldsymbol{A_s}\left(\frac{\boldsymbol{dT}}{\boldsymbol{dx}}\right)} \qquad \textit{Equation 3-8}$$

where:

$\boldsymbol{Q_h}$ = heat flux in W;

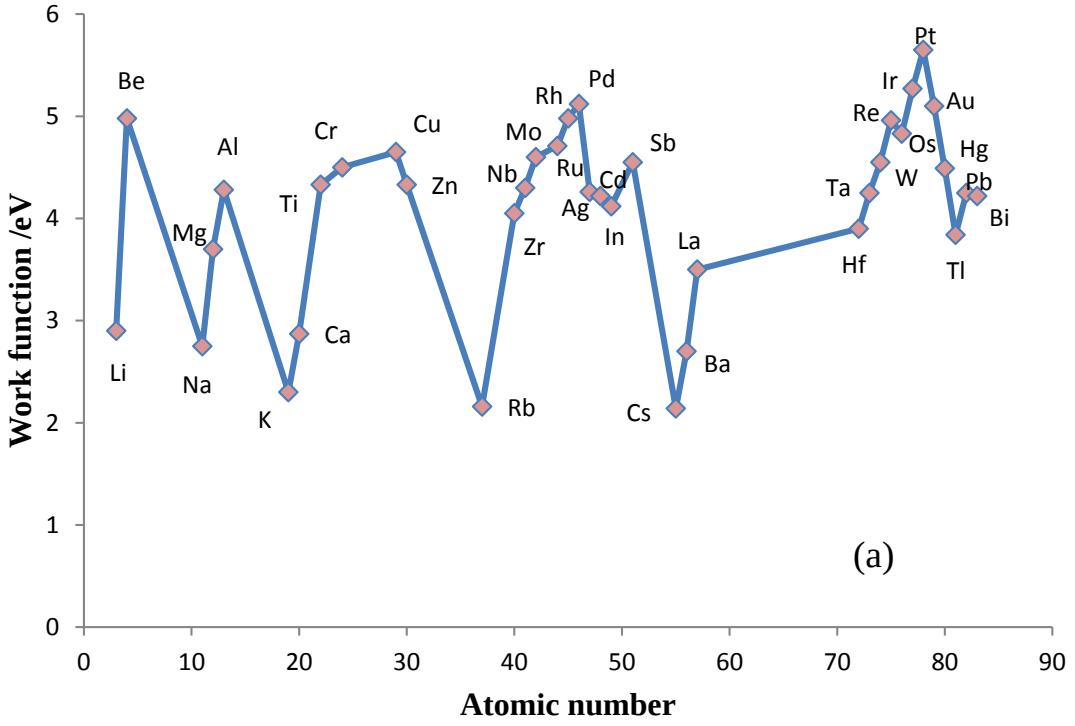
$\boldsymbol{T}$ = materials surface temperature in K;

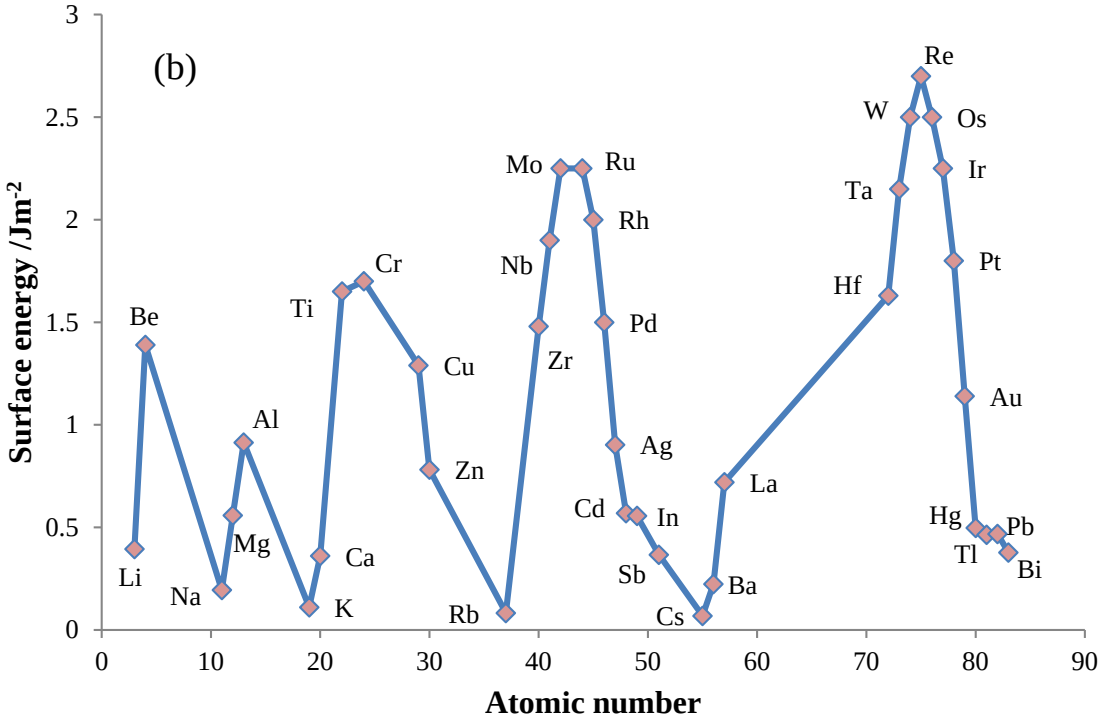
$\boldsymbol{x}$ = distance between the two temperatures in m.

The work function ($\boldsymbol{W}$) of a metal is the energy (usually measured in electron volts) required to extract one electron from highest filled level in the Fermi distribution of a solid to be at rest in vacuum at 0 K [268]. The data of work function recorded in CES [228] are the same as the data recently tabulated by Michaelson [269], excepted a very small variation for Mg (3.7 eV instead of 3.66 eV).

The value of work function can be obtained by the absolute methods (thermionic, photoelectric, and field emission) or the relative method (contact potential difference) [270]. Measurements of work function are sensitive to the anisotropy of sample [271] as well as its surface condition, such as surface impurities (i.e. oxides and gases) [268],

and the surface morphology (i.e. roughness) [272]. Ibragimov and Korolkov [273] noticed a slight temperature dependence of work functions for metals.





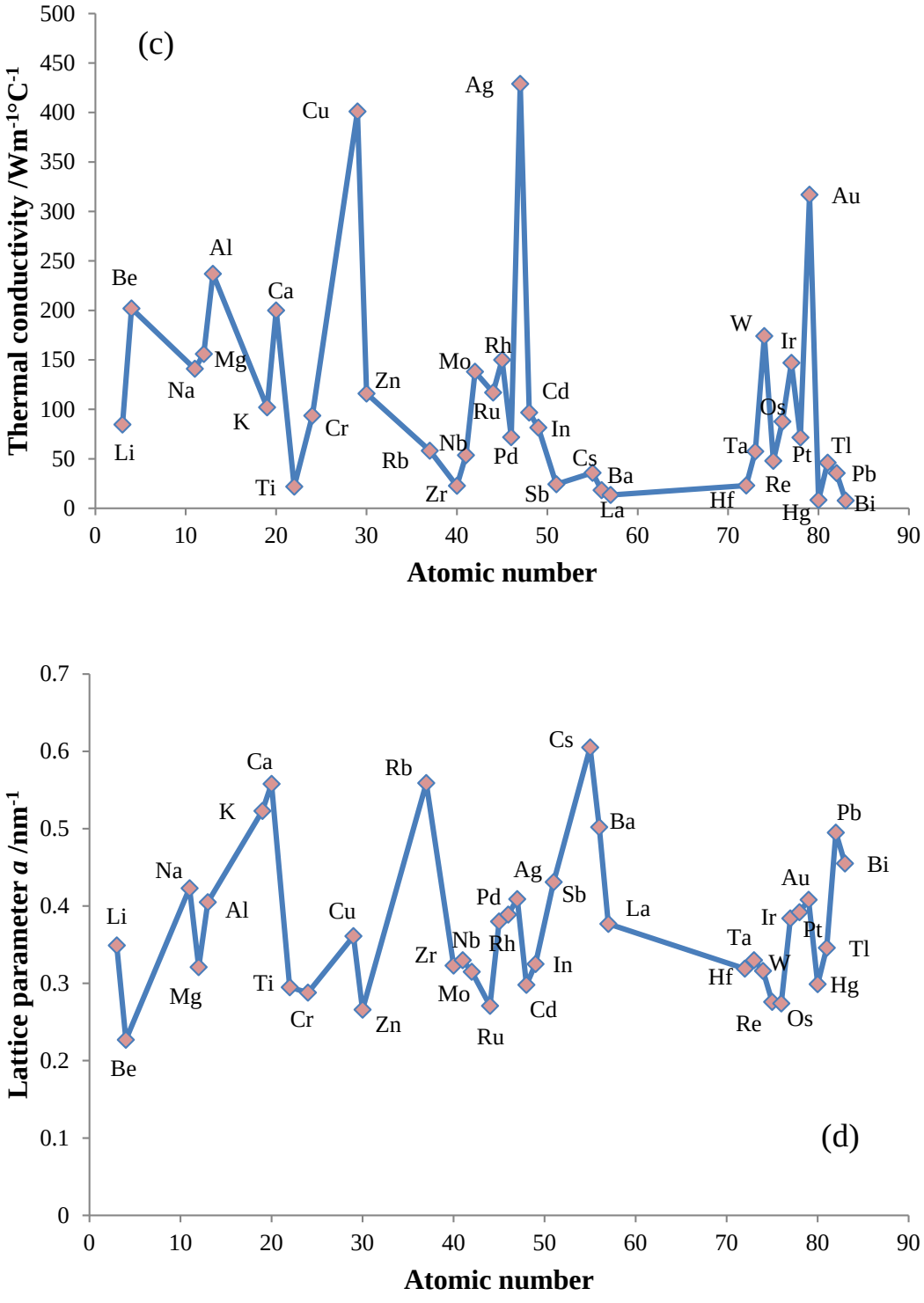


Figure 3-6 Plots of the atomic number of a metal versus its: a) work function, b) surface energy, c) thermal conductivity, and d) lattice parameter ***a****.*

The work function of a metal is found to be a periodic function of its atomic number, and is correlated with its electronegativity [274] and its first ionization potential [268]. Empirical relations between work function and surface energy [275], melting temperature [276], sublimation entropies [277], and Young's modulus [278] have also been reported for metals in the Periodic Table. However, the proposed correlations are of binary order. Good binary order correlations can be relatively easily identified through a regression analysis or by plotting properties against each other. Ternary order correlations sometimes can be recognized by plotting their 3D graphs but are difficult to be obtained through traditional regression analysis unless their analytical forms are pre-defined, not to mention quaternary order property correlations. However, this is not a problem for the ANN method.

The ANN combinatorial search found a good quaternary order property relationship between work function, surface energy, thermal conductivity and lattice parameter ***a*** for the 37 elements. As shown in Figure 3-6, the data of work function, surface energy and thermal conductivity exhibit similar periodic trends when plotting against atomic number, while an roughly inverse trend is observed for the data of lattice parameter ***a***.

The relative importance of each property in the quaternary relationship with work function is explored by comparing the binary, ternary, and quaternary order correlations listed in Table 3-5. The surface energy is identified to be the variable that contributes most in the prediction of the work function, and followed by the lattice parameter ***a***, and finally by the thermal conductivity. The work function and the surface energy of metals are expected to correlate with each other as they characterize the surface properties of a given material. In addition, both the work function and the thermal conductivity refer to electrons, thus the relation between the two properties seemingly origin in nature. However, the physical meaning behind the correlation between lattice parameter ***a*** and work function needs a further investigation. It could be the confounding effect of the three properties that substantially reduces the total correlation error to 2.7%.

*Table 3-5 The binary, ternary and quaternary order correlations between work function, surface energy, thermal conductivity and lattice parameter **a** for the 37 elements.*

Conditions		**ANNs evaluation parameters**		
Predicted property	**Input property**	ψ	δ	E_c
	Υ	0.11	0.32	6.6%
	a	0.15	0.38	8.6%
	λ	3.15	1.77	15.8%
W	Υ and a	0.02	0.12	4.5%
	Υ and λ	0.14	0.36	7.5%
	a and λ	0.64	0.80	10.7%
	Υ, a and λ	0.00	0.04	2.7%

3.4 Conclusion

The ANNs combinatorial approach is devised as an exhaustive search method to evaluate all possible property correlations hidden in the database. Instead of creating composition 'libraries', property 'libraries' based on the 24 properties for 37 elements are prepared and analysed accordingly. A new evaluation criterion (ψ) that combines the parameters of ANN generalization ability (δ) and correlation error (E_c), is introduced to justify the importance of the observed property correlations. For property correlation-group, such evaluation criterion can be applied according to a quadratic superposition rule.

In addition, this work summarises the types of strong correlations that are identified by the ANN combinatorial search. Though it is impossible to discuss all the explored correlations, two typical examples of binary order property correlation-groups, three ternary order property correlation-groups, and a quaternary order property correlation, are discussed respectively. They are:

1) Binary order correlation groups:
    a) cohesive energy and heat of vaporization at the normal boiling point;
    b) atomic weight and specific heat capacity.
2) Ternary order correlations groups:
    a) cohesive energy, boiling point and heat of vaporization;
    b) heat of vaporization, surface energy and molar volume;
    c) shear modulus, bulk modulus, and Poisson's ratio.
3) Quaternary order property correlation: using surface energy, thermal conductivity and lattice parameter ***a*** to predict work function,

Among those correlations, some have been discovered in history as the results of deductive reasoning from known principles or serendipity, such as the Dulong-Petit Law, while some correlations are first introduced, for example, the relation between cohesive energy of solid and heat of vaporization at the normal boiling point. Though empirical correlations obtained by ANNs do not provide deep insight into the physical nature, they could enrich the understanding of the fundamental material properties. Furthermore, the potential of applying the ANN combinatorial search to capture correlations hidden in a database is very promising. The captured correlations can be employed to predict unknown data, test theoretical models or hypotheses, and to check the internal consistency of experimental data. It is assumed that the present work will stimulate further activity in this field for discovering new physical laws. In the meantime, the poor data quality noticed in the elastic properties database motives the author to conduct a careful data evaluation in the next chapter.

4. Verification of the elastic properties of the elements through ANNs

4.1 Introduction

The Young's modulus (***E***), shear modulus (***G***), bulk modulus (***K***) and Poisson's ratio (***v***) of pure metals are the important and fundamental mechanical properties. Their data quality plays a vital role in defining the accuracy of many materials-based activities [113, 117]. However, an inspection of elastic moduli data in the 68 metallic elements will show large discrepancies exist in the prestigious handbooks and databases. Because the discrepancy is substantial for many elements, and more importantly, to inspire confidence in users, the author feels it is necessary to verify the inconsistent data.

A data-correction method utilizing binary order correlations through the ANN approach has been proposed recently [7]. In this chapter, to minimize the using of prior knowledge, the ANN method is extended to capture the ternary order correlations between elastic properties, and utilize such correlations to correct suspect data. The construction of ANN models is based on the premise that most handbooks and databases have recorded correct values for the most commonly known pure metals. To ensure the data accuracy, comparisons are made between the values obtained by the ANN method, the results generated by the correlating equations (i.e. Equation 1-21), and the available experimental values. To explain data disparity, both experimental and theoretical factors are discussed.

4.2 Experiment

The error in a database normally lies between 1% and 5% [234, 235]. However, examining the elastic property data recorded in literature for 68 pure metallic elements, more than 80% data are found to have variances larger than 10%. Such finding serves the purpose to conduct an integrated quality research of elastic property data to distinguish errors from reasonable variances.

4.2.1 Data discrepancy in handbooks and databases

As tabulated in Appendix VIII ~ Appendix XI, the values of Young's moduli, shear moduli, bulk moduli and Poisson's ratio at 295±5 K of the 68 elementary metals, are collected from seven different books published from 1960s to 2010s, one journal article, one recently released electronic databases, and three Internet sources with very high search rankings. The total 12 sources are referred as 'the source pool' in the following sections. The largest variation in 13 different editions of CRC Handbook of Chemistry and Physics, or in 3 editions of Tables of Physical and Chemical Constants, or in 2 editions of Handbook of Mechanical Engineering, or in the database of CES released in from 2008 to 2011, is scarcely perceptible. Therefore, the value from the last edition of each source is used. Conversion to SI unit needs to be done before the evaluation. Elements are flagged based on a 10% variation between the minimum and maximum values recorded in the source pool. If (Max-Min)/Min is larger than 10%, the property value of such element is recognized as an exceptional value and need to be checked. A critical analysis is made later by tracking back to the original literature to locate the reasons of disparity.

4.2.2 Data pre-treatment

1) Annotation removed

Young's modulus, shear modulus, bulk modulus and Poisson's ratio are the mechanical properties to describe materials in the real world. The differences in experimental conditions would influence property values. In the source pool, some data are labelled as estimate values, approximate values, calculated values or values read from the graph. To enable data comparisons, such annotations are removed. In addition, only the values at ambient pressure and room temperature (295±5 K) are considered unless otherwise specified.

No literature/database in the source pool records specimens' purity data except for one handbook [279], in which, for example, the Young's modulus of ruthenium is recorded as "447 GPa (at 296 K, annealed, purity 99.8%)", and the lowest purity listed the handbook

is 99.1% for scandium. Because most sources do not specify the measurement conditions, only the principal numeric value, i.e. '447', is retained in order to compare with the data from other sources.

2) SI unit conversion

Because different units of elastic properties are used in the source pool, data are normalized into the SI unit 'GPa'. During the SI unit conversion, the maximum number of decimal digits is designated to be three for Poisson's ratio and two for other elastic properties. The unit conversions are described as below:

Table 4-1 Unit conversion.

Non-SI Unit	SI Unit /GPa
1 dyn/cm^2	1×10^{-10}
1 psi	6.894×10^{-6}
1 kbar	0.1
1 kgf/cm^2	9.807×10^{-5}

3) Data distribution information

For each element, the maximum value (Vmax), the minimum value (Vmin), the mean value (Vmean), the median value (Vmedian), the variance (Vvar) and the value that has the highest occurrence (Vmode) in the source pool are calculated and listed in Appendix VIII ~ Appendix XI. Data distribution information is analysed for each element in Section 4.3.3.

4.2.3 ANN methodology

The principal of the ANN methodology to detect and correct errors is based on the method proposed by Zhang *et al.* [7], who evaluated the boiling point and the heat of vaporization of elements. In contrast to the utilizing of binary order correlations, this

work is extended to employ the ternary order correlations among four elastic properties, so it would be suitable for wider applications.

The first step is to identify the inconsistent data values. Elements are judged based on the basis of a 10% variation between the minimum and maximum values in the source pool. If (Max-Min)/Min (denoted as 'Dminmaxp' in Appendix VIII ~ Appendix XI) is larger than 10%, the data are recognized as inconsistent and need to be verified or judged. If (Max-Min)/Min is smaller than or equal to 10%, the element is recognized as having consistent data with reasonable uncertainty, and the median of all available data in the source pool will be used.

Table 4-2 Systematic methodology for error corrections, where 1 represents an element with consistent data and 0 indicates values are inconsistent. Consistent data are used to construct ANN models that represent property correlations; such models are then used to evaluate inconsistent data (from Ref. [7]).

Categories	A	B	C	Methodology
I	1	1	1	Use consistent data to construct ANNs
II	1	1	0	Apply ANNs to inconsistent data
	1	0	1	
	0	1	1	
III	1	0	0	
	0	1	0	
	0	0	1	
IV	0	0	0	

According to Zhang *et al.* [7], the ternary order correlation constituting by the property A, B and C can be classified into four categories, as shown in Table 4-2, where 1 represents the property records have consistent data and 0 indicates the records are inconsistent. Category I has consistent records of all three properties, which will be used to train three

ANNs (i.e. ANN1, the prediction of C using A and B, ANN2, the prediction of B using A and C, and ANN3, the prediction of A using B and C). The other records are classified into Categories II-IV according to the number of inconsistent properties.

For elements in Category II:

The inconsistent properties can be predicted directly from the other two properties. For example, if the property C has inconsistent records, the value of C can be predicted from the values of A and B by using ANN1. Assume the predicted value $C_{(p)}$ is obtained from ANN1, the predicted value $C_{(p)}$ is compared with all the values of that property recorded in the source pool, and is replaced with the closest original value $C_{(o)}$.

For elements in Category III:

Two mutual correlated ANNs bound the reasonable values. For example, if the property B and C have inconsistent records, while the records in property A are consistent, for ***m*** different records of property B $\{B_{(o)j}, j = 1,2,\ldots m\}$ and ***n*** different records of property C $\{C_{(o)k}, k = 1,2,\ldots n\}$, the ANN2 will predict ***n*** records of property B $\{B_{(p)k}, k = 1,2,\ldots n\}$ from property A and ***n*** property C. The fractional prediction errors of property B are $Y = \{y_{jk}\}$, where $y_{jk} = \left|1 - \frac{B_{(p)k}}{B_{(o)j}}\right|, j = 1,2,\ldots m, k = 1,2,\ldots n.$ Similarly, ANN3 will predict ***m*** records of property C $\{C_{(p)j}, j = 1,2,\ldots m\}$ from property A and ***m*** property B. The fractional prediction errors of property C are $Z = \{z_{jk}\}$, where $z_{jk} = \left|1 - \frac{C_{(p)j}}{C_{(o)k}}\right|, j = 1,2,\ldots m, k = 1,2,\ldots n$. The correct combination is the one that has the minimum value of e_{jk}, where $e_{jk} = \sqrt{(y_{ik} + z_{jk})^2 + (y_{jk} - z_{jk})^2}$.

For elements in Category IV:

Three mutual correlated ANNs bound the reasonable values. For ***l*** different records of property A $\{A_{(o)i}, i = 1,2,\ldots l\}$, ***m*** different records of property B $\{B_{(o)j}, j = 1,2,\ldots m\}$ and ***n*** different records of property C $\{C_{(o)k}, k = 1,2,\ldots n\}$, the ANN1 will predict $m \times n$ records

of property A $\{A_{(p)jk}, j=1,2,\ldots m, k=1,2,\ldots n\}$ from ***m*** property B and ***n*** property C. The fractional prediction errors of property A are $X=\{x_{ijk}\}$, where $x_{ijk}=\left|1-\frac{A_{(p)jk}}{A_{(o)i}}\right|$ $i=1,2,\ldots,l,\ j=1,2,\ldots m, k=1,2,\ldots n$.

Similarly, ANN2 will predict $l\times n$ records of property B $\{B_{(p)ik}, i=1,2,\ldots l, k=1,2,\ldots n\}$ from ***l*** property A and ***n*** property C. The fractional prediction errors of property B are $Y=\{y_{ijk}\}$, where $y_{ijk}=\left|1-\frac{B_{(p)ik}}{B_{(o)j}}\right|, i=1,2,\ldots,l,\ j=1,2,\ldots m, k=1,2,\ldots n.$ Again, ANN3 will predict $l\times m$ records of property $\{C_{(p)ij}, i=1,2,\ldots l,\ j=1,2,\ldots m\}$ from ***l*** property A and ***m*** property B. The fractional prediction errors of property C are $Z=\{z_{ijk}\}$, where $z_{ijk}=\left|1-\frac{C_{(p)ij}}{C_{(o)k}}\right|, i=1,2,\ldots,l,\ j=1,2,\ldots m, k=1,2,\ldots n$.

The correct combination is the one that has the minimum value of e_{ijk}, where $e_{ijk}=\sqrt{\left(x_{ijk}+y_{ijk}+z_{ijk}\right)^2+\left(x_{ijk}-y_{ijk}\right)^2+\left(x_{ijk}-z_{ijk}\right)^2+\left(y_{ijk}-z_{ijk}\right)^2}$.

The elastic properties of the consistent elements in Category I are listed in Table 4-3. Ideally, there would be 12 ANNs constructed using the consistent data from Categories 1, which would be applied to verify the inconsistent data in Category II ~ Category IV. For isotropic materials, the correlations between the four elastic properties (***E***, ***G***, ***K*** and ***v***) are ternary order correlations, i.e. one property is correlated to any other two properties, meaning three properties out of four will suffice to construct ANNs. Therefore, there are several correlations that can be utilized to predict the property that has inconsistent records, which may results in different predictions.

For example, if ANN1 predicts ***E*** using ***G*** and ***K***, ANN2 predicts ***E*** using ***G*** and ***v***, ANN3 predicts ***E*** using ***K*** and ***v***, and the property ***E*** has inconsistent records, ***E*** can be predicted from ANN1, ANN2 or ANN3 utilizing any two properties that have consistent values. As a result, three predicted values $\boldsymbol{E_{(p)1}}$, $\boldsymbol{E_{(p)2}}$ and $\boldsymbol{E_{(p)3}}$ will be obtained from each ANN model. The predicted values $\boldsymbol{E_{(p)1}}$, $\boldsymbol{E_{(p)2}}$ and $\boldsymbol{E_{(p)3}}$ are compared with all the values of the

property $\boldsymbol{E}$ recorded in the source pool, and is replaced with the closest original value $\boldsymbol{E_{(o)1}}$, $\boldsymbol{E_{(o)2}}$ and $\boldsymbol{E_{(o)3}}$.

However, in a few cases, ANN1, ANN2 and ANN3 may not yield the same results. If the magnitudes of $\boldsymbol{E_{(o)1}}$, $\boldsymbol{E_{(o)2}}$ and $\boldsymbol{E_{(o)3}}$ are very close to each other, i.e. the (Max-Min)/Min of ($\boldsymbol{E_{(o)1}}$, $\boldsymbol{E_{(o)2}}$ and $\boldsymbol{E_{(o)3}}$) is smaller than 10% (the same standard to judge data consistency), the median value of $\boldsymbol{E_{(o)1}}$, $\boldsymbol{E_{(o)2}}$ and $\boldsymbol{E_{(o)3}}$ is selected. However, if $\boldsymbol{E_{(o)1}}$, $\boldsymbol{E_{(o)2}}$ and $\boldsymbol{E_{(o)3}}$ differ from each other, i.e. (Max-Min)/Min is larger than 10%, an methodology is needed to solve the discrepancy. Such methodology is described by Table 4-6 in the section 'Employ ANNs to verify data').

1) ANN constructions

The number of consistent data in the source pool is very small. In the 68 metallic elements, only five elements (Co, Dy, Fe, Ta and Tb) have consistent data, i.e. the variances of the four elastic properties are all smaller than 10%. Five elements (Ac, Fr, Np, Pa, Ra) have only one data available in the source pool (collected from CES [228]), and Pm only has estimate values.

To construct reliable ANNs, the input dataset needs to be large enough to cover the whole problem domain. Because a dataset that contains only five input-output data pairs (i.e. values of the five consistent elements) is too small to build a reliable ANN model, more elements' data need to be added. On the other hand, since problem domains are not always clear in a general situation, it is better to have the fewest subjective decisions on which element should be added.

It is reasonable to assume that most handbooks and databases have recorded correct values for the most commonly known pure metals. Thus, the values of elastic property that has the highest occurrence in the source pool of twenty-two well studied and widely used pure metal (Ag, Al, Au, Ba, Bi, Ca, Cr, Cu, Mg, Mn, Mo, Nb, Ni, Pb, Pd, Pt, Sn, Sr, Ti, V, W and Zn), are pre-chosen and defined as 'consistent' without any pre-treatment. The values of the twenty-two pure metals together with the five consistent data (Co, Dy, Fe, Np, Ta and Tb), are used to build ANN models (shown in Table 4-3). It should be mentioned that the elastic property data of the twenty-two elements might be incorrect.

However, ANNs are capable of giving fairly unbiased simulations without compromising to a small portion of tampered data. A further study as described in Section 4.3.1 is also performed to validate the elastic property data of the twenty-two elements.

*Table 4-3 Dataset used to build the ANNs includes the consistent values of 5 elements (Co, Dy, Fe, Np, Ta and Tb) and the most common values of 22 elements (Ag, Al, Au, Ba, Bi, Ca, Cr, Cu, Mg, Mn, Mo, Nb, Ni, Pb, Pd, Pt, Sn, Sr, Ti, V, W and Zn). **E** is Young's modulus, **G** is shear modulus, **K** is bulk modulus and **v** is Poisson's ratio.*

Atomic number	Symbol	*E*	*K*	*G*	*v*
47	Ag	82.7	103.6	30.3	0.37
13	Al	70	76	26	0.345
79	Au	78.5	171	26	0.42
56	Ba	12.8	9.6	4.86	0.28
83	Bi	32	31	12	0.33
20	Ca	19.6	17	7.4	0.31
27	Co	209	181.5	82	0.31
24	Cr	279	160.2	115	0.21
29	Cu	129.8	137.8	48.3	0.343
66	Dy	61.4	40.5	24.7	0.247
26	Fe	211	170	82	0.28
12	Mg	44.7	35.6	17.3	0.29
25	Mn	191	120	79.5	0.24
42	Mo	324.8	261.2	125.6	0.31
41	Nb	104.9	170.3	37.5	0.397
28	Ni	199.5	177.3	76	0.31
82	Pb	16	45.8	5.59	0.44
46	Pd	121	187	43.6	0.39
78	Pt	170	276	60.9	0.39
50	Sn	49.9	58.2	18.4	0.33
38	Sr	15.7	12	6.03	0.28
73	Ta	185.7	196.3	69.2	0.34
65	Tb	55.7	38.7	22.1	0.261

Atomic number	Symbol	*E*	*K*	*G*	*v*
22	Ti	116	108.4	45.6	0.32
23	V	127.6	158	46.7	0.365
74	W	411	311	160.6	0.28
30	Zn	104.5	69.4	41.9	0.249

2) ANNs simulations

The four elastic properties in the ANN training and testing dataset are for 27 metallic elements with atomic number ranging from 12 to 83, ***E*** ranging from 12.8 to 411 GPa, ***K*** ranging from 9.6 to 311 GPa, ***G*** ranging from 4.86 to 160.6 GPa, and ***v*** ranging from 0.21 to 0.44. Except the ANN model of predicting ***v*** from ***E*** and ***G*** could not be established (ANN model failed to converge after 30 independent trials), the ANNs of predicting one elastic property from any other two properties are all successfully constructed, see Figure 4-1 ~ Figure 4-4. Their correlation coefficients are larger than 0.99, except for the ANN of predicting ***v*** from ***E*** and ***K***, whose correlation coefficients equals to 0.987. The average error between the ANN predictions and values listed in Table 4-3 is far below 10%, except for the ANN model of predicting ***K*** from ***E*** and ***G***, whose average error reaches as high as 13%, thus is relatively less reliable, and should be used with great caution. The summary of ANN simulations of the four elastic properties correlations is listed in Table 4-4.

*Table 4-4 A summary of ANN simulations of the elastic properties correlations that are constituted by **E**, **G**, **K** and **v**.*

Property group	ANNs correlation	Average error percentage
E, K, v	Predict ***v*** from ***E*** and ***K***	2.34%
	Predict ***K*** from ***E*** and ***v***	3.08%
	Predict ***E*** from ***K*** and ***v***	3.54%

Property group	ANNs correlation	Average error percentage
G, K, v	Predict ***v*** from ***K*** and ***G***	1.99%
	Predict ***G*** from ***K*** and ***v***	5.87%
	Predict ***K*** from ***G*** and ***v***	3.98%
G, E, v	Predict ***v*** from ***E*** and ***G****	—
	Predict ***G*** from ***E*** and ***v***	6.11%
	Predict ***E*** from ***G*** and ***v***	1.47%
G, E, K	Predict ***K*** from ***E*** and ***G****	13.15%
	Predict ***E*** from ***K*** and ***G***	1.10%
	Predict ***G*** from ***E*** and ***K***	1.43%

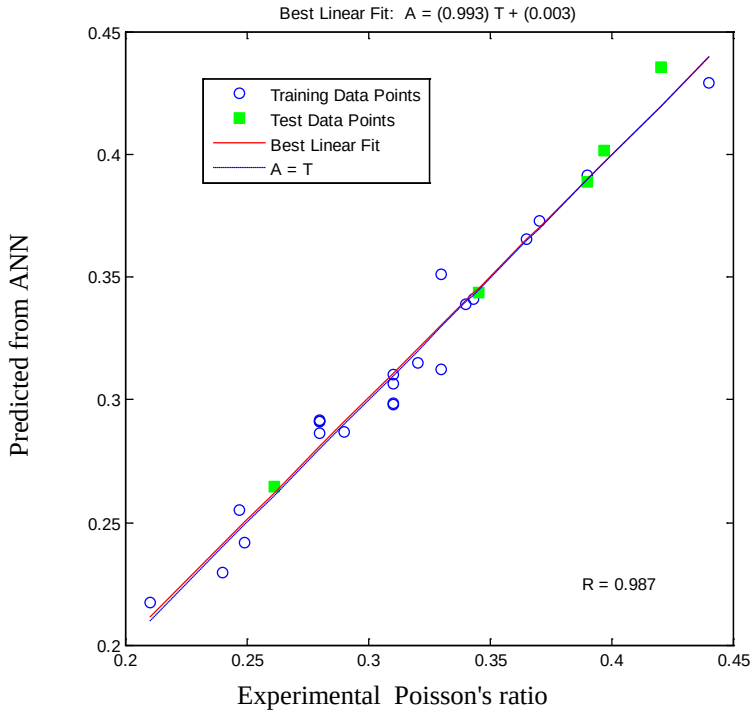
* The ANN model failed to converge in 30 independent trials, thus, it is inapplicable for further usage.

** ANNs should be used with great caution due to large prediction error.

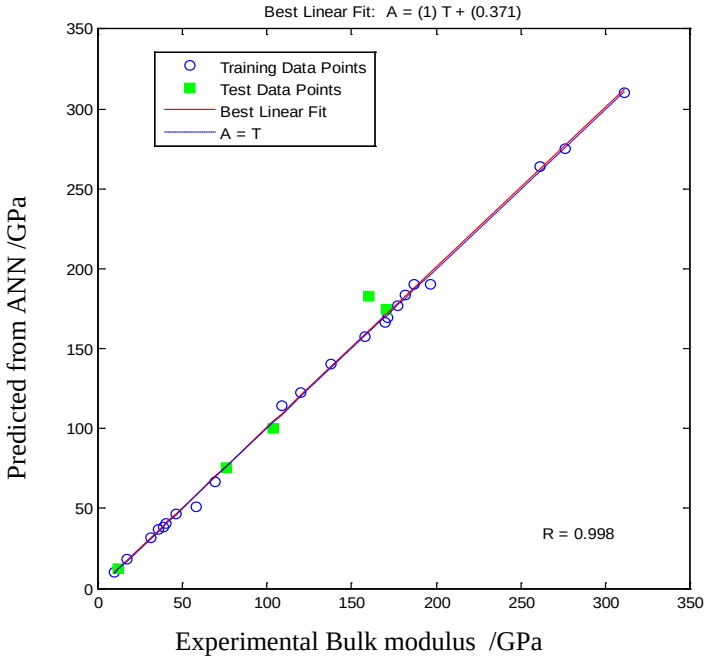
Theoretically, a ANN model with one hidden layer (with sigmoid transfer function) and enough neurons is capable of representing versatile functional relationships between the inputs and the outputs [227]. However, the ANN model predicting ***v*** from ***E*** and ***G*** failed to converge after 30 independent trials. A possible reason is that the initial weights and bias in training ANNs could be set inappropriately that cause the ANN to oscillate between relatively poor solutions. Different training functions, more hidden layers and more training may help to yield a good result.

However, the author has no intention to use all possible means in constructing a valid ANN model to predict ***v*** from ***E*** and ***G*** or obtaining a more accurate ANN model to predict ***K*** from ***E*** and ***G***. First, in the practice of employing ANNs to find property correlations in materials science, there is no way to know whether a strong hypothetical relationship exists without any prior knowledge, knowledge that is seldom available in a general situation. The ANN results summarized in Table 4-4 only indicate there are ten strong correlations and one relatively weak correlation. Realizing that ***v*** can be predicted from ***E*** and ***G*** and ***K*** can be accurately obtained from ***E*** and ***G*** requires prior knowledge. Secondly, it is also interesting to see if the ANN data policing methodology still holds

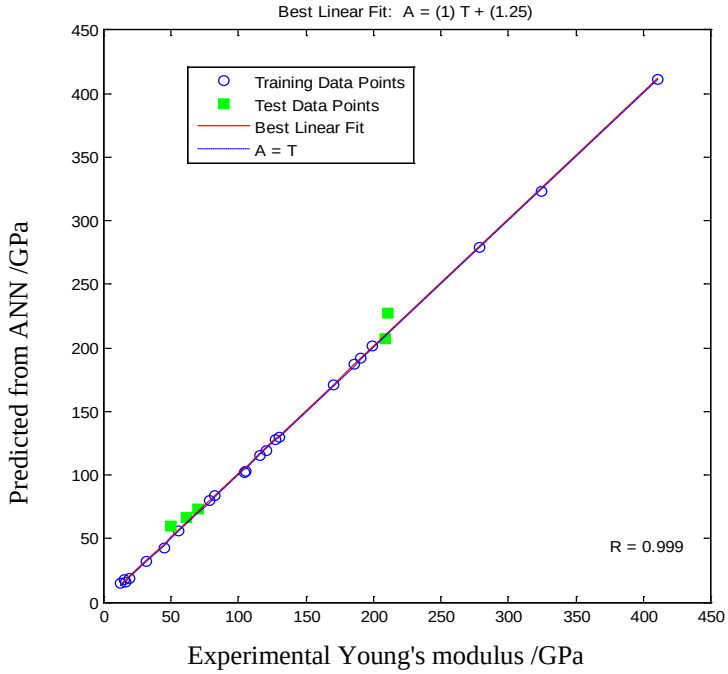
when ANNs have not successfully identified a hidden correlation. In this case, the 11 ANNs listed in Table 4-4 are utilized in the next step.



a)

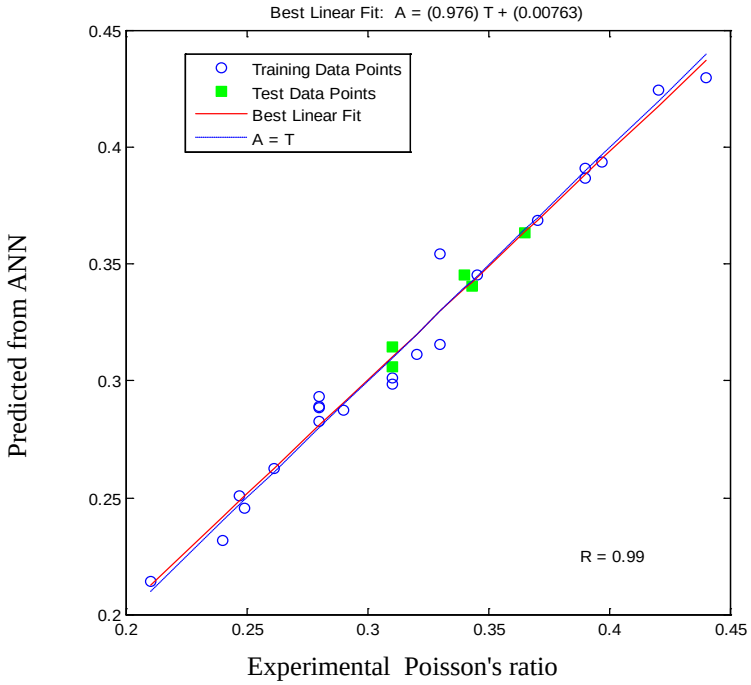


b)

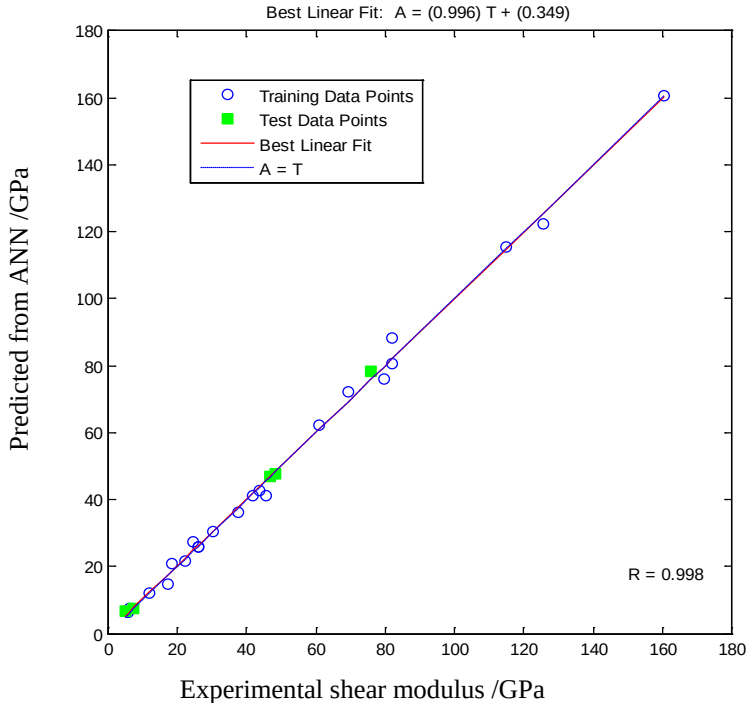


c)

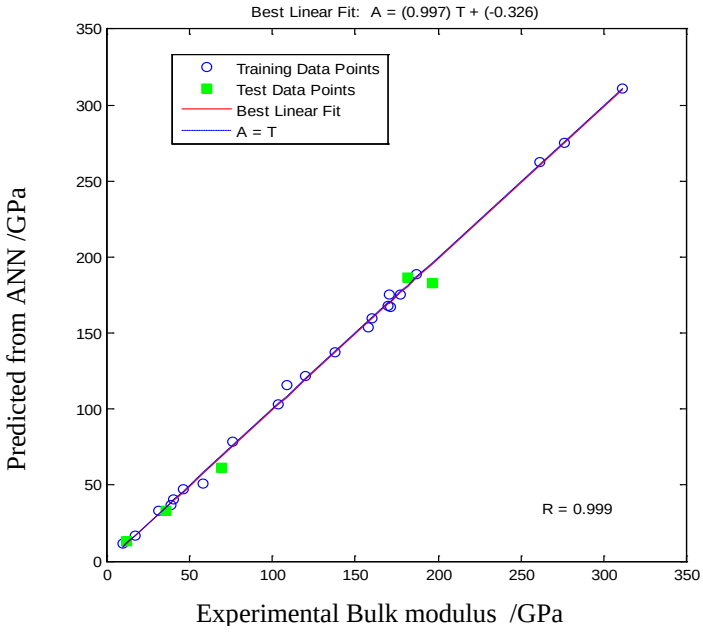
*Figure 4-1 ANN correlations of property group (**E**, **K** and **ν**): a) Predict **ν** from **E** and **K**, b) Predict **K** from **E** and **ν**, c) Predict **E** from **K** and **ν**.*


Best Linear Fit: A = (0.976) T + (0.00763)
Training Data Points
Test Data Points
Best Linear Fit
A = T
R = 0.99
Predicted from ANN
Experimental Poisson's ratio
0.2
0.25
0.3
0.35
0.4
0.45

a)

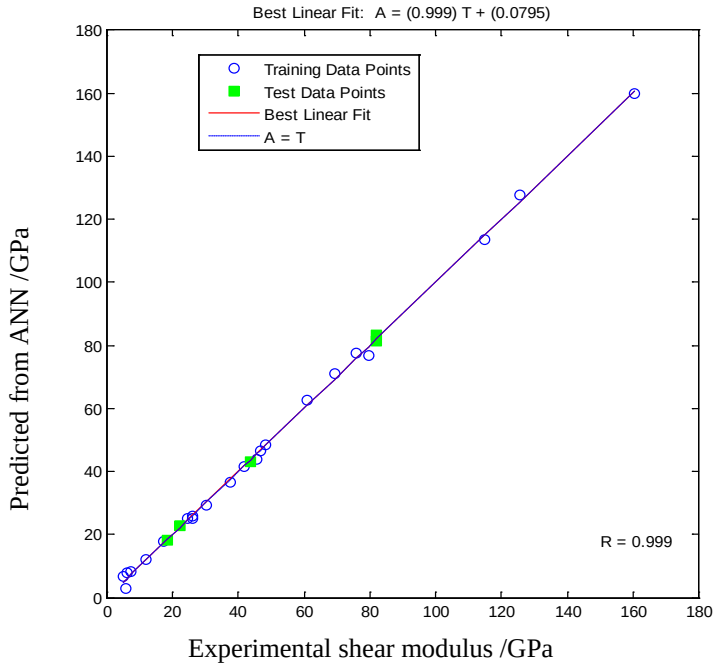

Best Linear Fit: A = (0.996) T + (0.349)
Training Data Points
Test Data Points
Best Linear Fit
A = T
R = 0.998
Predicted from ANN /GPa
Experimental shear modulus /GPa
0
20
40
60
80
100
120
140
160
180

b)

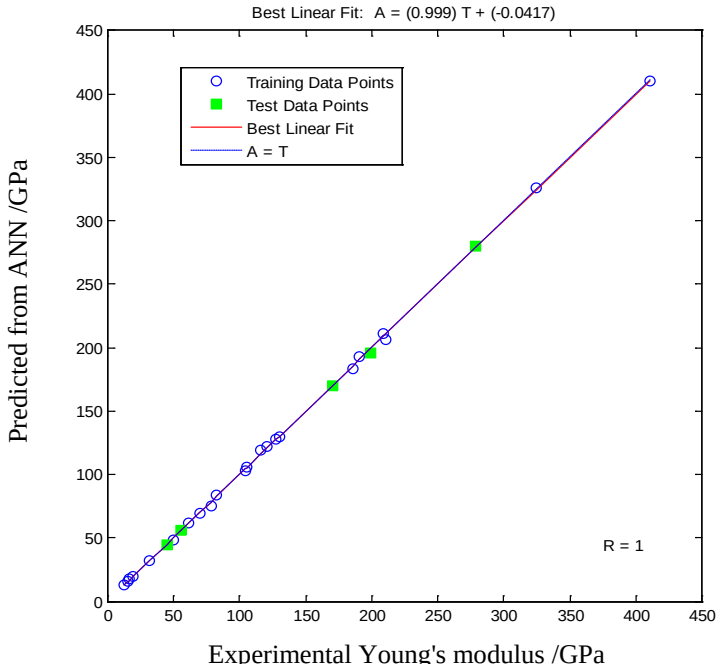


c)

*Figure 4-2 ANN correlations of property group (**G**, **K** and **v**): a) Predict **v** from **K** and **G**, b) Predict **G** from **K** and **v**, c) Predict **K** from **G** and **v**.*

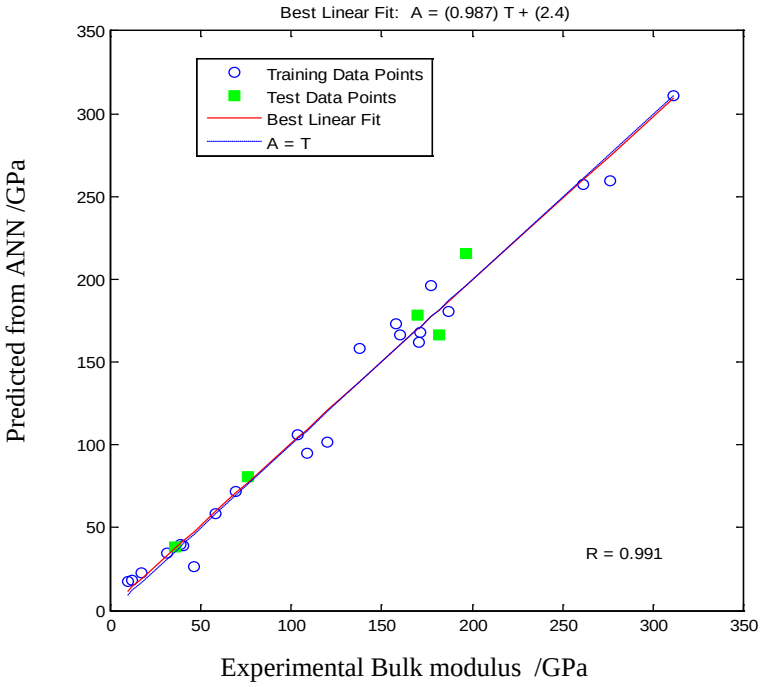


a)

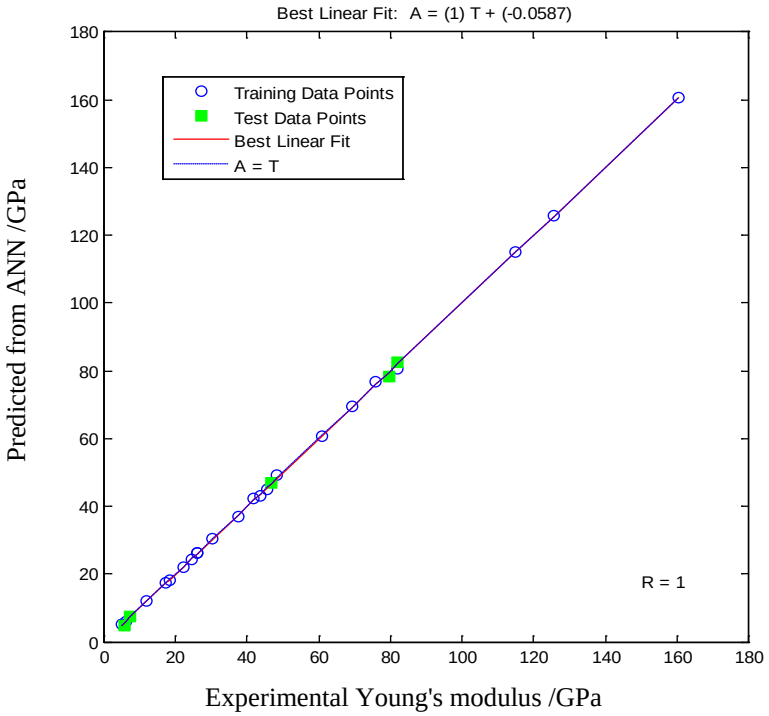


b)

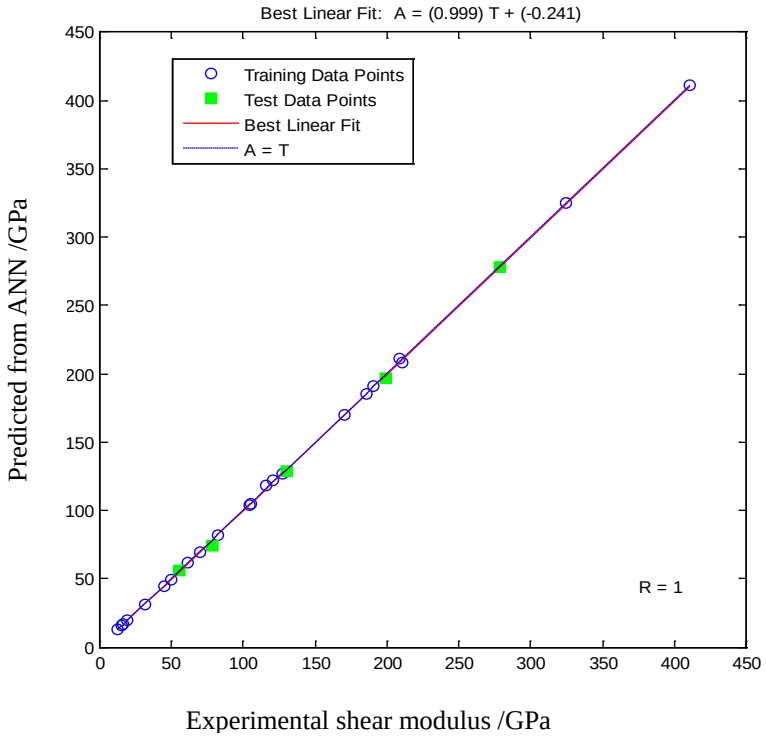
Figure 4-3 *ANN correlations of property group (**G**, **E** and **ν**): a) Predict **G** from **E** and **ν**, b) Predict **E** from **G** and **ν**.*



a)



b)



c)

*Figure 4-4 ANN correlations of property group (**G**, **E** and **K**): a) Predict **K** from **E** and **G**, b) Predict **E** from **K** and **G**, c) Predict **G** from **E** and **K**.*

3) Employ ANNs to verify data

The eleven ANNs of elastic property correlations constructed from 27 data are then used to predict inconsistent data as demonstrated in Table 4-5. Similarly, inconsistent data can also be calculated through the correlating equations that listed in Table 1-2. The data verified by ANNs are listed in Table 4-9, and are compared with the data calculated from the correlating equations.

Table 4-5 Correlation for error checking in the source pool.

Categories	Inconsistent property	Total elements	Element Symbol	Useful correlation
1	-	22	Ag, Al, Au, Ba, Bi, Ca, Cr, Cu, Mg, Mn, Mo, Nb, Ni, Pb, Pd, Pt, Sn, Sr, Ti, V, W, Zn	—
		5	Co, Dy, Fe, Ta, Tb	
		6*	Ac, Fr, Np, Pa, Ra, Pm	
2	***G***	2	Gd, Nd	Predict ***G*** from ***K*** and ***v***
				Predict ***G*** from ***E*** and ***v***
				Predict ***G*** from ***E*** and ***K***
	E	—	—	—
	K	2	Er, Ir	Predict ***K*** from ***E*** and ***v***
				Predict ***K*** from ***G*** and ***v***
				Predict ***K*** from ***E*** and ***G*****
	v	—	—	—
3	***E***, ***G***	3	Li, Pr, Tc	Predict ***E*** from ***v*** and ***K***
				Predict ***G*** from ***K*** and ***v***
				Predict ***E*** from ***G*** and ***v***
				Predict ***G*** from ***E*** and ***v***
				Predict ***E*** from ***K*** and ***G***
				Predict ***G*** from ***E*** and ***K***
	K, ***E***	1	Rh	Predict ***K*** from ***G*** and ***v***
				Predict ***E*** from ***G*** and ***v***

Categories	Inconsistent property	Total elements	Element Symbol	Useful correlation
				Predict ***E*** from ***K*** and ***ν***
				Predict ***K*** from ***E*** and ***ν***
				Predict ***K*** from ***E*** and ***G*****
				Predict ***E*** from ***K*** and ***G***
	K, ***G***	—	—	—
	K, ***ν***	8	Cs, Os, Re, Ru, Sc, Tm, Y	Predict ***K*** from ***E*** and ***G*****
				Predict ***K*** from ***E*** and ***ν***
				Predict ***ν*** from ***E*** and ***K***
				Predict ***K*** from ***G*** and ***ν***
				Predict ***ν*** from ***K*** and ***G***
	E, ***ν***	1	Th	Predict ***E*** from ***K*** and ***G***
				Predict ***ν*** from ***K*** and ***G***
				Predict ***E*** from ***K*** and ***ν***
				Predict ***ν*** from ***E*** and ***K***
	ν, ***G***	—	—	—
4	***ν***, ***K***, ***E***	2	Ho, Zr	Predict ***K*** from ***E*** and ***G*****
				Predict ***E*** from ***K*** and ***G***
				Predict ***K*** from ***G*** and ***ν***
				Predict ***ν*** from ***K*** and ***G***
				Predict ***K*** from ***E*** and ***ν***
				Predict ***ν*** from ***E*** and ***K***
				Predict ***E*** from ***K*** and ***ν***
	ν, ***K***, ***G***	—	—	—
	ν, ***E***, ***G***	3	Hf, Na, Pu	Predict ***ν*** from ***K*** and ***G***
				Predict ***G*** from ***K*** and ***ν***
				Predict ***ν*** from ***E*** and ***K***
				Predict ***E*** from ***K*** and ***ν***
				Predict ***G*** from ***E*** and ***K***
				Predict ***E*** from ***K*** and ***G***

Categories	Inconsistent property	Total elements	Element Symbol	Useful correlation
	K, ***E***, ***G***	4	Ce, In, K, Tl	Predict ***E*** from ***K*** and ***v***
				Predict ***K*** from ***E*** and ***v***
				Predict ***K*** from ***G*** and ***v***
				Predict ***G*** from ***K*** and ***v***
				Predict ***E*** from ***G*** and ***v***
				Predict ***G*** from ***E*** and ***v***
				Predict ***G*** from ***E*** and ***K***
				Predict ***E*** from ***K*** and ***G***
				Predict ***K*** from ***E*** and ***G*****
5	***K***, ***G***, ***E***, ***v***	10	Be, Cd, Eu, Ga, La, Lu, Rb, Sm, U, Yb	Predict ***E*** from ***K*** and ***v***
				Predict ***K*** from ***E*** and ***v***
				Predict ***v*** from ***E*** and ***K***
				Predict ***G*** from ***E*** and ***K***
				Predict ***E*** from ***K*** and ***G***
				Predict ***K*** from ***E*** and ***G*****
				Predict ***G*** from ***K*** and ***v***
				Predict ***K*** from ***G*** and ***v***
				Predict ***v*** from ***K*** and ***G***

* The elements are omitted from the evaluation for lack of reliable data.

** The ANN model should be used with great caution due to the large error prediction.

Due to the complexity of the data recorded in the source pool, in a few cases, elastic properties predicted utilizing different correlations established by ANNs or by the correlating equations may not yield consistent results. Take rhodium (Rh) as an example, Young's modulus and bulk modulus calculated using known shear modulus (149.45 GPa) and Poisson's ratio (0.26) are 377 GPa and 262 GPa, respectively. The corresponding closest data-pair recorded in the source pool is 379 GPa and 270.3 GPa, which are consistent with the best data pair (380 GPa and 276 GPa) determined with shear modulus and Equation 1-21-2 and Equation 1-21-5, but is in contradiction with the best data pair

(318.5 GPa and 221.2 GPa) selected solely according to Poisson's ratio and Equation 1-21-8 and Equation 1-21-11. Methods 1, 2 and 3 (as listed in Table 4-6) can resolve such discrepancy:

Table 4-6 The systematic methodology to resolve the discrepancy among different ANN predictions. 1 indicates property records are consistent in the source pool, and 0 indicates property records are inconsistent in the source pool.

Categories		A	B	C	D	Methodology
1		1	1	1	1	
2	2.1	1	1	1	0	Method 1
	2.2	1	1	0	1	
	2.3	1	0	1	1	
	2.4	0	1	1	1	
3	3.1	1	1	0	0	Method 2
	3.2	1	0	1	0	
	3.3	1	0	0	1	
	3.4	0	1	1	0	
	3.5	0	1	0	1	
	3.6	0	0	1	1	
4	4.1	1	0	0	0	Method 3
	4.2	0	1	0	0	
	4.3	0	0	1	0	
	4.4	0	0	0	1	
5		0	0	0	0	Method 1, 2 or 3

Method 1: Take Category 2.1 as an example. With a unique record of property A, B and C, ***l*** different sets of property D $\{D_i, i = 1,\ldots l\}$ will be obtained from different correlations, i.e. using the correlation of A and B to predict D, using the correlation of A and C to predict D, and using the correlation of B and C to predict D. With the correlation of using A and D to predict B and the correlation of using C and D to predict B, property

B $\{B_{(AD)i}, B_{(CD)i}, i = 1,\ldots l\}$ will be obtained from each set of property (A and D_i), and from each set of property (C and D_i), respectively. The fractional errors of property B is $X = \{x_{(AD)i}, x_{(CD)i}\}$, where $x_{(AD)i} = \left|1 - \frac{B_{(AD)i}}{B}\right|$, $x_{(CD)i} = \left|1 - \frac{B_{(CD)i}}{B}\right|, i = 1,2,\ldots l$. Similarly, a set of property A $\{A_{(BD)i}, A_{(CD)i}\}$ with corresponding fractional errors $Y = \{y_{(BD)i}, y_{(CD)i}\}$, and a set of property C $\{C_{(AD)i}, C_{(BD)i}\}$ with corresponding fractional errors $Z = \{z_{(AD)i}, z_{(BD)i}\}$ will be obtained. The best set to be chosen has the minimum value of e_i, where

$$e_i = \sqrt{\left(x_{(AD)i} - x_{(CD)i}\right)^2 + \left(y_{(BD)i} - y_{(CD)i}\right)^2 + \left(z_{(AD)i} - z_{(BD)i}\right)^2 + \left(x_{(AD)i} + x_{(CD)i} + y_{(BD)i} + y_{(CD)i} + z_{(AD)i} + z_{(BD)i}\right)^2}$$

Method 2: Take Category 3.1 as an example. Within a unique record of property A and property B, ***l*** different sets of property C and property D $\{C_i, D_i, i = 1,2,\ldots l\}$ may be obtained from different correlations. Again, using the ***l*** different sets of (property C and property D), ***l*** different sets of (property A and property B) $\{A_i, B_i, i = 1,2,\ldots l\}$ may be obtained, with corresponding fractional errors of property A, i.e. $x_{(CD)i} = \left|1 - \frac{A_i}{A}\right|$, $i = 1,2,\ldots l$ and property B, i.e. $y_{(CD)i} = \left|1 - \frac{B_i}{B}\right|, i = 1,2,\ldots l$. The best set to be chosen has the minimum value of e_i, where

$$e_i = \sqrt{\left(x_{(CD)i} - y_{(CD)i}\right)^2 + \left(x_{(CD)i} + y_{(CD)i}\right)^2}.$$

Method 3: Take Category 4.1 as an example. Within a unique record of property A, ***l*** different sets of property B, property C and property D $\{B_{i,}, C_i, D_i, i = 1,2,\ldots l\}$ may be obtained from different correlations. Again, using those ***l*** different sets, ***l*** different sets of property A $\{A_{(BC)i}, A_{(BD)i}, A_{(CD)i}, i = 1,2,\ldots l\}$ from property set (B and C), (B and D), and (C and D) may be obtained. Their corresponding fractional errors are $x_{(BC)i}$, $x_{(BD)i}$ and

$x_{(CD)i}$, where $x_{(BC)i} = \left|1 - \frac{A_{(BC)i}}{A}\right|$, $x_{(BD)i} = \left|1 - \frac{A_{(BD)i}}{A}\right|$, $x_{(CD)i} = \left|1 - \frac{A_{(CD)i}}{A}\right|$, $i = 1,2,\ldots l$. The best set to be chosen has the minimum value of e_i, where

$$e_i = \sqrt{\left(x_{(BC)i} - x_{(BD)i}\right)^2 + \left(x_{(BC)i} - x_{(CD)i}\right)^2 + \left(x_{(BD)i} - x_{(CD)i}\right)^2 + \left(x_{(BC)i} + x_{(BD)i} + x_{(CD)i}\right)^2}$$.

For Category 5, because at least one elastic property predicted by employing different correlations established by either ANNs or the correlating equations listed in Table 1-2 has a variance smaller than 10%, such property data can be treated as consistent. Methods 1, 2 and 3 described above can then be applied to yield the best predictions of the other properties.

4.3 Result and discussion

4.3.1 Validation of ANNs

Two questions may arise in employing ANNs to capture the correlations hidden in elastic property data. The first one is whether any experimental evidence supports the hypothesis that the elastic property data of the 27 elements that were used to build ANNs are correct. The second one is to what degree the elastic property correlations simulated by ANNs from a relatively small dataset are analogous to the known relationships represented by the correlating equations listed in Table 1-2. The two questions are addressed as below.

1) Valid inputs for ANN constructions

Because ANNs are data-driven modelling approaches, the relationships captured by ANNs highly depend on the input data. In order to validate the elastic properties of 27 elements which are used to train ANNs, the pre-defined consistent data (i.e. the most common values in the source pool), are compared with the elastic properties calculated from the single crystal elastic constants through the VRH averaging method, except for Mn and Ca (see Table 4-7 and Table 4-8). To the author's knowledge, no elastic constant measurement on the single crystal of manganese or calcium is available. The experimental values used for the two elements were tabulated by Koster [140].

A comparison of Table 4-7 and Table 4-8 indicates the pre-defined consistent elastic properties of 27 elements are in excellent agreement with experimental values, except for barium and lead, whose Young's modulus and shear modulus calculated from single crystal elastic constants are more than 10% larger than the pre-defined values. The only literature providing the single crystal elastic constants for barium is poorly determined, as neither C_{11} nor C_{12} is directly obtained from the experiment [280]. Though it is not possible to assign a reason for the large discrepancy observed for lead, the elastic properties calculated from the single crystal elastic constants (provided by Ref. [281]) are probably more reliable than the values that first appeared in Ref. [140], which had little experimental information.

Nevertheless, the discrepancy of elastic property relating to barium and lead and the uncertainty of elastic property associated with manganese or calcium for lacking single crystal data, should not affect the intrinsic relationships established by the ANNs. These data still obey the correlating equations that are listed in Table 1-2, which the ANNs are supposed to capture

Table 4-7 Comparison of elastic properties of 27 pre-defined consistent elements between the values used to train ANNs (denoted as Anp.), and the values calculated from single crystal elastic constants (denoted as Exp.). See text for a discussion for Mn and Ca.

Symbol	Young's modulus /GPa		Shear modulus /GPa		Bulk modulus /GPa		Poisson's ratio	
	Exp.	Anp.	Exp.	Anp.	Exp.	Anp.	Exp.	Anp.
Ag	80.99	82.7	29.56	30.3	103.78	103.6	0.37	0.37
Al	70.04	70	26.01	26	76.09	76	0.35	0.345
Au	78.33	78.5	27.50	26	172.87	171	0.42	0.42
Ba	13.65	12.8	5.41	4.86	9.55	9.6	0.26	0.28
Bi	33.68	32	12.63	12	33.63	31	0.33	0.33
Ca	19.61	19.6	7.49	7.4	17.16	17	0.31	0.31
Co	215.93	209	82.36	82	190.33	181.5	0.31	0.31

Symbol	Young's modulus /GPa		Shear modulus /GPa		Bulk modulus /GPa		Poisson's ratio	
	Exp.	**Anp.**	**Exp.**	**Anp.**	**Exp.**	**Anp.**	**Exp.**	**Anp.**
Cr	279.65	279	115.36	115	161.87	160.2	0.21	0.21
Cu	127.35	129.8	47.31	48.3	137.83	137.8	0.35	0.343
Dy	61.43	61.4	24.63	24.7	40.45	40.5	0.25	0.247
Fe	208.19	211	80.66	82	165.67	170	0.29	0.28
Mg	44.70	44.7	17.34	17.3	35.26	35.6	0.29	0.29
Mn	197.70	191	80.15	79.5	123.56	120	0.23	0.24
Mo	319.45	324.8	122.72	125.6	268.27	261.2	0.30	0.31
Nb	105.20	104.9	37.64	37.5	171.33	170.3	0.398	0.397
Ni	209.73	199.5	79.90	76	186.33	177.3	0.31	0.31
Pb	24.14	16	8.56	5.59	44.61	45.8	0.41	0.44
Pd	131.52	121	47.43	43.6	193.06	187	0.39	0.39
Pt	161.63	170	57.47	60.9	287.16	276	0.41	0.39
Sn	48.47	49.9	17.90	18.4	55.38	58.2	0.35	0.33
Sr	14.80	15.7	5.72	6.03	11.99	12	0.29	0.28
Ta	182.66	185.7	68.09	69.2	191.92	196.3	0.34	0.34
Tb	56.99	55.7	22.69	22.1	38.89	38.7	0.26	0.261
Ti	114.62	116	43.36	45.6	107.27	108.4	0.32	0.32
V	129.14	127.6	47.39	46.7	156.49	158	0.36	0.365
W	409.56	411	160.16	160.6	308.33	311	0.28	0.28
Zn	100.96	104.5	40.74	41.9	64.48	69.4	0.24	0.249

*Table 4-8 The elastic property data for the 27 pre-defined consistent elements. Values are calculated from single crystal elastic constants at 295±5 K, except for Mn and Ca. **E** is Young's modulus, **G** is shear modulus, **K** is bulk modulus and **v** is Poisson's ratio.*

Symbol	Structure	Elastic property				Elastic constant /GPa							Ref
		E /GPa	***G*** /GPa	***K*** /GPa	***v***	C_{11}	C_{44}	C_{12}	C_{13}	C_{14}	C_{33}	C_{66}	
Ag	Cubic	80.99	29.56	103.78	0.37	123.99	46.12	93.67					[282]
Al	Cubic	70.04	26.01	76.09	0.35	106.78	28.21	60.74					[283]
Au	Cubic	78.33	27.50	172.87	0.42	192.34	41.95	163.14					[282]
Ba	Cubic	13.65	5.41	9.55	0.26	12.6	9.5	8.02					[280]
Bi	Rhom	33.68	12.63	33.63	0.33	63.5	11.3	24.7	24.5	7.23	38.1	19.4	[284]
Ca*	Cubic	19.61	7.49	17.16	0.31								[140]
Co	Hex	215.93	82.36	190.33	0.31	307.1	75.5	165	102.7		358.1	71.05	[285]
Cr	Cubic	279.65	115.36	161.87	0.21	350	100.8	67.8					[286]
Cu	Cubic	127.35	47.31	137.83	0.35	169.1	75.42	122.2					[287]
Dy	Hex	61.43	24.63	40.45	0.25	73.1	24	25.3	22.3		78.1	23.9	[288]
Fe	Cubic	208.19	80.66	165.67	0.29	229	115	134					[289]
Mg	Hex	44.70	17.34	35.26	0.29	59.4	16.4	25.61	21.44		61.6	16.9	[290]
Mn*	Complex cubic	197.70	80.15	123.56	0.23								[140]
Mo	Cubic	319.45	122.72	268.27	0.30	469.6	106.8	167.6					[291]
Nb	Cubic	105.20	37.64	171.33	0.40	246	28.7	134					[292]
Ni	Cubic	209.73	79.90	186.33	0.31	249	114	155					[289]
Pb	Cubic	24.14	8.56	44.61	0.41	49.5	14.92	42.16					[281]

Symbol	Structure	Elastic property				Elastic constant /GPa							Ref
		E /GPa	***G*** /GPa	***K*** /GPa	***v***	C_{11}	C_{44}	C_{12}	C_{13}	C_{14}	C_{33}	C_{66}	
Pd	Cubic	131.52	47.43	193.06	0.39	227.1	71.73	176.04					[293]
Pt	Cubic	161.63	57.47	287.16	0.41	338.03	75.63	261.72					[294]
Sn	tetra	48.47	17.90	55.38	0.35	72	21.9	58.5	37.4		88	24	[284]
Sr	Cubic	14.80	5.72	11.99	0.29	15.3	9.9	10.34					[280]
Ta	Cubic	182.66	68.09	191.92	0.34	260.91	81.82	157.43					[295]
Tb	Hex	56.99	22.69	38.89	0.26	69.24	21.75	24.98	21.79		74.39	22.13	[296]
Ti	Hex	114.62	43.36	107.27	0.32	162.4	46.7	92	69		180.7	35.2	[297]
V	Cubic	129.14	47.39	156.49	0.36	230.06	42.81	119.71					[298]
W	Cubic	409.56	160.16	308.33	0.28	521.4	160.4	201.8					[291]
Zn	Hex	100.96	40.74	64.48	0.24	163	39.4	30.6	48.1		60.3	65.9**	[299]

* No single crystal elastic constant is available, see text for further discussion.

** The data of C_{66} listed here is directly adopted from Ref. [299], which is the average value over a large number of studies. Although its magnitude is slightly smaller than 1/2(C_{11}-C_{12}), it is expected to be closer to the true level of pure metal.

2) Correlations captured by ANNs

Ideally, the elastic property correlations captured by ANNs should be the same as the relationships indicated by the correlating equations listed in Table 1-2. However, due to the small size of training data, this is unlikely to happen. Appendix XVIII illustrates how each elastic property (denoted by the Z axis) would change with the increasing of any other two properties (denoted by the X axis and the Y axis) in the correlating equations and the ANN models. In total, 11 figures were drawn. The changing tendency of the correlating equation is indicated as mesh grids, while the changing tendency of the ANN models is indicated as colourful surfaces, i.e. the yellow surfaces represent the ANNs predicting the Young's modulus. Similarly, the blue surfaces represent the ANNs predicting the shear modulus, the green surfaces represent the ANNs predicting the bulk modulus, and the cyan surfaces represent the ANNs predicting the Poisson's ratio. The values of 27 elements, which are used to construct the ANNs, are also plotted in the figures as red data points.

Figure 4-5 is an example that compares the ANN modelling (three-dimensional yellow shaded surface) with the correlating equations (mesh grids) in predicting the Young's modulus from the Poisson's ratio and the shear modulus. The Young's moduli (Z component) are obtained by assuming a series of values at fixed intervals for Poisson's ratio (Y component) and shear moduli (X component). As can be seen, the true values of the 27 elements are located in the overlapping areas of the yellow surface and the mesh grids. Moreover, in the data range that is bound by the maximum and minimum value of the 27 elements (i.e. shear moduli from ~4 GPa to ~170 GPa, Young's moduli from ~12 GPa to ~411 GPa, Poisson's ratio from ~0.2 to ~0.44), the trend of the ANN model are very similar to the trend of the correlating equation. For an enlarged data region, the similarity between the correlation captured by ANNs and the relationship illustrated by the correlating equation will decline. Because not enough data in such region are applied for the ANN training, the uncertainty of the ANN increases by extrapolation, making the ANN less applicable when presented with new data.

.

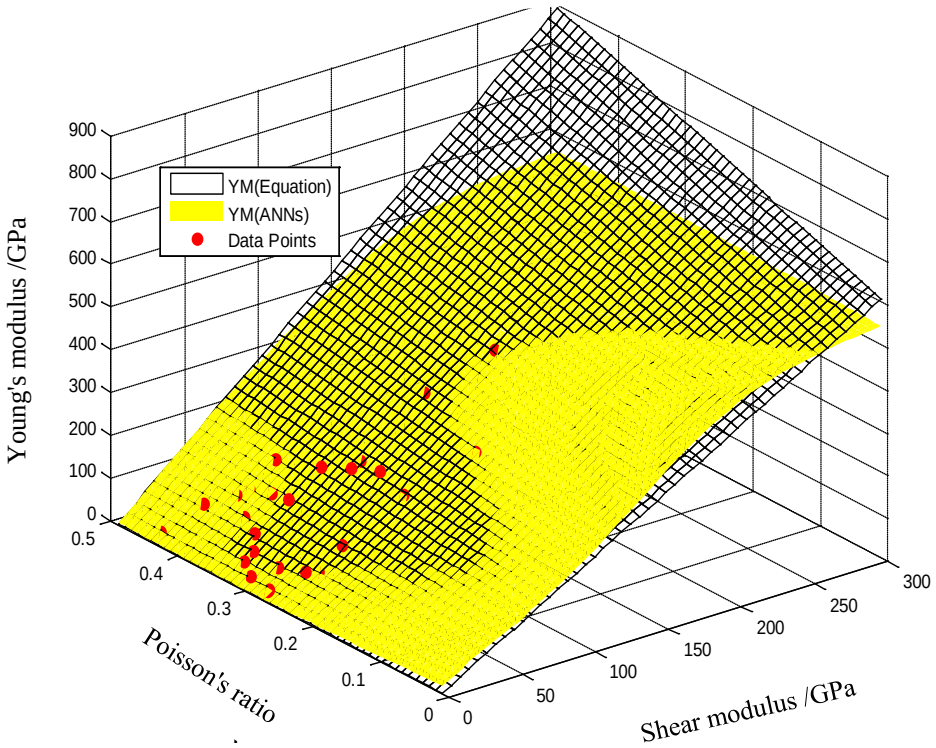


*Figure 4-5 Comparison of the ANN model and the correlating equation in predicting **E** from **ν** and **G** in 3D dimensions.*

4.3.2 Comparisons of elastic properties predicted by ANNs and the correlating equations

Table 4-9 tabulates the results that are evaluated by the ANN models and the results that are evaluated by the correlating equations (according to the same methodology that is described in Table 4-6). It also includes the data that occur most frequently in the source pool for each element. The underlined data in italic font indicate the values are inconsistent for the labelled property in the source pool, which are verified by the ANNs and by utilizing the correlating equations.

The results also show that the ANNs are equivalent to the correlating equations in a reasonable data range. For elements, whose values are located in data range of the 27 training examples, the data predicted by the ANNs are in good agreement with the values obtained from the correlating equations is evident: the average error is about 2%.

However, for elements whose elastic property data are outside the range of the training examples, the extrapolations from the ANN model are less reliable. Especially for Ru, Na, In, Tl, Eu, Rb, U and Yb, the average error is as large as ~23%.

Because the ANNs are trained with a relatively small size of data, the ANN predictions that are outside training domain provide rough estimations. However, the prediction accuracies of the ANNs might be improved by employing multiple correlations (ANN models) that mutually restrict each other. Take lithium as an example. Its largest value of shear modulus recorded in the source pool, i.e. 4.31 GPa, is smaller than the smallest input in ANNs training examples, i.e. 4.86 GPa. The result predicted by the ANN model from the consistent bulk modulus and Poisson's ratio is 7.82 GPa, which leads to 4.31 GPa to be chosen from the source pool as the closest value to the prediction. However, the data predicted within the usage of an ANN model that correlate Poisson's ratio and Young's modulus (to-be-determined), together with an ANN model that correlate bulk modulus and Young's modulus (to-be-determined), is 3.85 GPa. It is the same as the result directly obtained by employing Equation 1-21-1 and Equation 1-21-7.

It should be emphasised, however, the results predicted from a single ANN correlation and from several mutual-restrained ANN correlations should be treated equally if the magnitude of data is within the data range covered by the ANN training examples. In this case, the data recommend by ANNs are their average value (if the difference is less than 10%) or the value re-determined through the approach that is described in Table 4-6 (if elastic properties predicted by utilizing different correlations do not yield results within 10% variation).

Table 4-9 Comparison of the elastic property values yield by ANNs and the Equation 1-21.

Symbol	Most common value in the source pool				Results yield by ANNs				Results yield by Equations			
	v	***K*** /GPa	***E*** /GPa	***G*** /GPa	***v***	***K*** /GPa	***E*** /GPa	***G*** /GPa	***v***	***K*** /GPa	***E*** /GPa	***G*** /GPa
Gd	0.259	38	55	*21.9*				21.8				21.8
Nd	0.281	31.9	41.2	*16.15*				16				16
Er	0.237	*44.4*	69.95	28.3		44				44.4		
Ir	0.26	*365*	528	209.8		383				371		
Li	0.36	11.2	*4.91*	*4.23*			10.44	3.85			10.44	3.85
Pr	0.281	28.8	*37.3*	*14.8*			37	15			36	14.8
Tc	0.301	289	*351*	*134*			322	123			322	123
Rh	0.26	*275.5*	*330*	149.45		276	379			270.3	379	
Cs	*0.295*	*1.6*	1.7	0.65	0.295	1.4			0.295	1.4		
Os	*0.25*	*376.45*	558.75	222.25	0.253	379.9			0.255	379.9		
Re	*0.293*	*367.5*	463	180	0.298	365			0.289	365		
Ru*	*0.286*	*286*	432	173	0.3	348			0.25	286		
Sc	*0.279*	*56.6*	74.4	29.1	0.279	57.3			0.28	56.6		
Tm	*0.213*	*44.5*	74	30.5	0.235	45			0.217	43.6		
Y	*0.252*	*41.2*	64.4	25.6	0.246	46.9			0.242	46.9		
Th	*0.27*	54	*78.3*	30.8	0.263		75.65		0.26		78.3	

Symbol	Most common value in the source pool				Results yield by ANNs				Results yield by Equations			
	v	***K*** /GPa	***E*** /GPa	***G*** /GPa	***v***	***K*** /GPa	***E*** /GPa	***G*** /GPa	***v***	***K*** /GPa	***E*** /GPa	***G*** /GPa
Ho	*0.231*	*40.2*	*64.8*	26.3	0.255	39.7	65		0.23	40	64.8	
Zr	*0.34*	*89.8*	*94.95*	35	0.332	95.3	94		0.332	95.3	94.95	
Hf	*0.3*	109	*140*	*54.9*	0.282		140	56	0.282		143	55.8
Na*	*0.34*	6.52	*6.8*	*2.915*	0.315		6.8	2.53	0.366		5.41	1.98
Pu	*0.195*	54	*96*	*43*	0.21		87.5	34.5	0.21		91.75	34.5
Ce	0.245	*21.7*	*33.6*	*13.5*		21.5	33.5	14		22	33.6	13.53
In*	0.45	*41.1*	*10.8*	*3.68*		42	14	3.68		41.1	13.8	4.78
K	0.35	*3.14*	*3.53*	*1.3*		4.2	3.53	1.3		4.2	3.53	1.3
Tl*	0.45	*35.8*	*7.975*	*2.775*		28.5	15.42	5.4		35.7	15.42	5.4
Be	*0.032*	*117.785*	*291.8*	*139*	0.118	130	287	128.4	0.118	125.57	287.25	128.4
Cd	*0.3*	*46.7*	*58.6*	*24*	0.303	51	62.3	24.6	0.302	52.9	62.54	24
Eu*	*0.152*	*8.3*	*18.2*	*7.9*	0.286	14.7	18.2	7.9	0.152	8.7	18.2	7.9
Ga	*0.47*	*42.6*	*9.81*	*6.67*	0.235	58.2	93.2	37.8	0.233	58.2	93.2	37.8
La	*0.28*	*27.9*	*37.5*	*14.9*	0.284	30.3	36.8	14.15	0.284	28	37	14.25
Lu	*0.2605*	*47.6*	*68.6*	*27.2*	0.26	42.6	61.5	24.41	0.26	42.7	61.5	24.41
Rb*	*0.328*	*2.5*	*2.35*	*0.91*	0.3	2.3	2.4	1.02	0.374	2.3	1.73	0.63
Sm	*0.274*	*37.8*	*49.7*	*19.5*	0.268	35.7	49.7	19.55	0.268	35.7	49.7	19.6

Symbol	Most common value in the source pool				Results yield by ANNs				Results yield by Equations			
	v	***K*** /GPa	***E*** /GPa	***G*** /GPa	***v***	***K*** /GPa	***E*** /GPa	***G*** /GPa	***v***	***K*** /GPa	***E*** /GPa	***G*** /GPa
U*	*0.22*	*100*	*208*	*97.7*	0.24	97.9	175.8	73.1	0.205	100	175.8	74.05
Yb*	*0.207*	*30.5*	*23.9*	*9.9*	0.284	13.8	18	7.1	0.205	13.2	23.93	9.95

* The value predicted from the ANNs is different from the value calculated from the correlating equations.

4.3.3 Original experimental value

In order to access the uncertainty of the ANN data policing method, and to analyse the reason that causes data disparity in the source pool, an attempt is made to present all available information, referring to the original sources, for the 35 elements that have inconsistent elastic property data. The investigations of the experimental elastic property data at room temperature for each element have been summarized in Table 4-10 ~ Table 4-44, including the measurements that were conducted on single crystals and polycrystalline samples. When the measurements on single crystals are available, the elastic constants are provided as well as the elastic properties obtained through the VRH averaging method. For the measurements on polycrystalline materials, elastic property data are directly collected from the literature.

The author's choices of the best experimental values of elastic properties for the 35 elements are marked in bold and italic font in Table 4-10 ~ Table 4-44 for each element for comparison, and are summarised in Appendix XVII. In general, data obtained on single crystals are preferred, because the assumption that polycrystalline materials are isotropic is not always true [119]. If multiple studies on the elastic constants are available for an element, the values from the latest experiments that were conducted on a wider temperature range are selected, because the original sources may give extra information including the temperature dependences of elastic properties, for example, Na (see page 138). When no single crystal elastic constant was measured at room temperature, or when estimations are made in obtaining single crystal elastic constants, data obtained from polycrystalline material are used, such as the data chosen for the elements Tc, Cs and Tm. Moreover, the elastic property data of polycrystalline materials obtained from dynamic measurements are believed to be more accurate than the data obtained from static measurements [171, 174]. Consequently, the dynamic measurement results are deemed the best values.

To compare the experimental values and the values verified by the ANNs and by the correlating equations (listed in Table 1-2), the property that has sparse values in the source pool has its data underlined in Table 4-10 ~ Table 4-44 for each element. Most errors in the source pool attribute to different experimental conditions, i.e. temperature, purity, specimen structure (single crystal or polycrystalline), and measurements (static

or dynamic methods). Some errors are the results of incorrect unit conversions, e.g. gallium.

1) Gd and Nd (Inconsistent *G*)

Both gadolinium (Gd) and neodymium (Nd) are rare earth elements that adopt the hexagonal close-packed (hcp) structure at room temperature. They have a large discrepancy of shear moduli in the source pool. Except for the smallest value for Gd (12.5 GPa) which comes from Ref. [279], the other 7 data are around 22 GPa, the same value recommended by the ANNs and the correlating equations. For Nd, shear modulus data recorded in CES [228] and Ref. [279] (~14.5 GPa) differ from the data recorded in the other 6 sources (~16 GPa). Both ANNs and the correlating equations recommend that the correct value for Nd is 16 GPa.

The elastic properties of Gd and Nd have been thoroughly reviewed by Scott [119], who studied several investigations and found the experimental values of shear moduli were in excellent agreement at ~22 GPa for Gd [296, 300-304]. While for Nd, the minimum value of shear moduli, i.e. 14.5 GPa, is given by Smith *et al.* [300], but other experiments values, either by calculating from the single crystal elastic constants [305] or by determining from the polycrystalline samples [303, 306], are all close to 16 GPa.

2) Er and Ir (Inconsistent *K*)

i. Erbium (Er)

Er is the rare earth element with hcp structure. Its bulk modulus data in the source pool are in good agreement, except for the smallest value (41.1 GPa, a variation of 13.14%) recorded in CES. The smallest value seems to be obtained by Bridgman using a static method [152, 307]. Because the bulk modulus (~45 GPa) calculated through the VRH averaging method using the single crystal elastic constants, which were obtained by Fisher *et al*. [308], Rosen *et al*. [309], Palmer *et al*. [296], and Plessis [310], coincides with the polycrystalline value that were reported by Smith *et al*.[300] and Rosen [302], it is believed to be more accurate than Bridgman's data.

Table 4-10 Elastic property of gadolinium (Gd) and neodymium (Nd).

Element	C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	v	Ref
	GPa										
Gd	67.83	71.23	20.77	21.12	20.72	25.59	***54.8***	***21.8***	***37.9***	***0.259***	[296]
							56.2	22.3	38.92	0.259	[300]
							55.6	22.2	38.1	0.252	[301, 302]
							59.2	23.8	38.6	0.244	[119, 304]
							56.1	22.3	38.46	0.257	[303]
							55	21.8	38	0.259	Equations
							55	21.8	38	0.259	ANNs
Nd	54.82	60.86	15.03	15.1	16.6	24.62	***41.8***	***16.3***	***31.8***	***0.281***	[305]
							37.9	14.5	32.47	0.306	[300]
							40.7	15.8	31.35	0.283	[301, 306]
							43.1	16.8	33.11	0.283	[303]
							41.2	16	31.9	0.281	Equations
							41.2	16	31.9	0.281	ANNs

i. Iridium (Ir)

Ir is face-centred cubic (fcc) transition metal, which is found to be brittle at room temperature in both single crystal and polycrystalline state [311]. Its single crystal elastic constants were dynamically studied by Macfarlane *et al*. [131] from 4.2 K to room temperature. The elastic properties of polycrystalline Ir were thoroughly investigated by Merker *et al*. [312] from room temperature to 1500 K using the resonance technique. The substantial discrepancy of bulk moduli may be ascribed to the anisotropy associated with the as-cast microstructure of Merker's test specimen. The bulk modulus that is calculated using the single crystal elastic constants is in good agreement with the static measurement results (~355 GPa) [152, 313].

The elastic property data recorded in the source pool are identical to the values tabulated by Koster [140], excepted for bulk modulus data, which generally falls in two groups:

~320 GPa and ~370 GPa. Only the larger value corresponds relatively well with the data in Ref. [140]. The difference between the bulk modulus predicted from ANNs and the value calculated using the correlating equations is very small (about 3%), thus it is acceptable.

Table 4-11 Elastic property of erbium (Er).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	v	Ref
GPa										
86.3	85.5	28.1	27.9	22.7	30.5	71.6	28.9	45.5	0.238	[308]
82.5	84.5	27.9	27	21.1	28.5	70.0	28.4	43.4	0.231	[309] *
83.67	84.45	27.53	27.19	22.22	29.29	***69.9***	***28.3***	***44.3***	***0.237***	[296]
83.88	84.32	26.45	27.31	24.21	29.26	68.9	27.6	45.3	0.246	[310]
						73.3	29.6	46.48	0.238	[300]
						72.9	29.6	45.66	0.233	[301, 302]
						73.35	29.62	41.07	0.238	[152]
						69.95	28.3	44.4	0.237	Equations
						69.95	28.3	44	0.237	ANNs

* Elastic constants are read off from the graph.

Table 4-12 Elastic property of iridium (Ir).

C_{11}	C_{44}	C_{12}	E	G	K	v	Ref
GPa							
580	256	242	***540.19***	***216.74***	***354.67***	***0.246***	[131]
			525.5	218.2	296.06*	0.204*	[312]
			527.89	209.86	365.79	0.260	[140]
			528	209.8	371	0.26	Equations
			528	209.8	383	0.26	ANNs

* The value is obtained using Equation 1-21.

3) Li, Pr and Tc (Inconsistent *E* and *G*)

i. Lithium (Li)

Li is a very soft alkali metal with a cubic structure. Robertson and Montgomery [314] studied the Young's modulus of isotopically concentrated polycrystalline Li (~8 GPa) at room temperature through the resonant measurement, and found the isotopic mass had no effect on the elastic properties. The elastic constants of single crystal Li have been dynamically measured by Nash and Smith [315] at low temperature (78 K, 155 K and 195 K), by Jain [134] at 300 K and by Felice *et al.* [316] from 90 K to 300 K. Their data were in reasonably good agreement when extrapolated to room temperature. Data from Felice *et al.* [316] are recommended to be the best values because elastic properties were provided at a wider temperature range.

In the source pool, Young's moduli (~10 GPa) provided by CES and Ref. [115], which are close to the experimental results [134, 315, 316], seriously disagree with the value 4.9 GPa provided by other sources. The value 4.9 GPa was acquired from a static test that was first reported by Bridgman [317]. While the most common value of shear moduli in the source pool (~4.2 GPa) seems to be calculated from the single crystal data at 83 K by Gschneidner [152]. It is much larger than the recent experimental results listed in Ref. [134, 315, 316]. Even though the shear modulus and Young's modulus of Li are slightly smaller than the lower bounds of ANNs training and testing datasets, the ANNs gave the same values as the correlating equations did.

Table 4-13 Elastic property of lithium (Li).

C_{11}	C_{44}	C_{12}	E	G	K	v	Ref
GPa							
14.85	10.80	12.53	12.66	4.72	13.30	0.341	[315]*
11.51	8.56	9.60	10.10	3.78	10.23	0.335	[315]**
14.06	8.8	12.02	10.69	3.93	12.70	0.360	[134]
13.36	8.82	11.28	***10.70***	***3.96***	***11.97***	***0.351***	[316]
			4.90	1.72	11.11	0.427	[317]
			11.47	4.23	11.57	0.362	[152]&

C_{11}	C_{44}	C_{12}	E	G	K	v	Ref
			GPa				
			10.44	3.85	11.2	0.36	Equations
			10.44	3.85	11.2	0.36	ANNs

* At 78 K.

** Linearly extrapolated to 300 K.

[&] At 83 K.

ii. Praseodymium (Pr)

Pr is a rare earth element with hcp structure at room temperature and below [318]. Scott [119] made an excellent compilation for the elastic data published in literature and recommended the data that were obtained by Greiner *et al.* [318] for a future usage. Table 4-14 is the summery of elastic properties measured at room temperature through dynamic measurements. Smith *et al.* [300], Rosen [306] and Gust and Royce [303] all investigated the elastic behaviour of polycrystalline Pr. Only Greiner *et al.* [318] has determined Pr single crystals data.

In the source pool, the maximum value of Young's moduli (50 GPa) is recorded in CES, which exhibits a significant gap from the rest data (~37 GPa). The variation of shear moduli recorded in the source pool is close to 10%, and is narrowed down by the ANNs and the correlating equations. In addition, the Young's moduli and shear moduli predicted by ANNs and the correlating equations agrees well with the experimental values.

Table 4-14 Elastic property of praseodymium (Pr).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	v	Ref
					GPa					
49.35	57.4	13.6	13.2	14.3	22.95	***37.9***	***14.8***	***28.8***	***0.281***	[318]
						35.2	13.5	29.9	0.305	[300]
						36	13.9	28	0.291	[306]
						38.5	15	29.9	0.283	[303]
						36	14.8	28.8	0.281	Equations
						37	15	28.8	0.281	ANNs

iii. Technetium (Tc)

Tc crystallizes in hcp structure at room temperature [319]. Very few studies have reported the elastic properties for Tc, owing to the difficulty in acquiring a sample with high purity [320]. Therefore, it is no surprise to find that there are only two data in the source pool for Tc: the smaller data (***E***=322 GPa, ***G***=123 GPa) is collected from Ref. [115], and the larger one (***E***=380 GPa, ***G***=145 GPa) is collected from CES. The larger data can be traced back to the values of polycrystalline Tc that were dynamically measured by Love *et al.* [320]. Though there is no direct measurement on Tc single crystal, the VRH averaging results using elastic constants derived by Guillermet and Grimvall [319] from experimental phonon dispersion curves based on the assumption that C_{12} equals C_{13}, are in close agreement with Love's data.

No literature can be identified as the origin of the data recorded in CES, but Ref. [115] and CES give relatively consistent values for Poisson's ratio (~0.3) and bulk modulus (~290 GPa). The calculated values of Young's modulus and shear modulus by employing Equation 1-21 are the same as the ANN predicted data. Table 4-15 lists the experimental elastic properties [319, 320] together with the values estimated by Gschneidner [152].

Table 4-15 Elastic property of technetium (Tc).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	ν	Ref
				GPa						
433	470	117	117	199	199	313.6	119.4	280.9	0.314	[319]*
						322.0	**123.0**	**281.0**	**0.309**	[320]
						368.73	142.20	297.14	0.293	[152]**
						322	123	289	0.301	Equations
						322	123	289	0.301	ANNs

* Assuming C_{12} equals C_{13}.

** Estimated values.

4) Rh (Inconsistent *E* and *K*)

Rhodium (Rh) is a transition metal possessing fcc crystal structure at room temperature [321]. Most bulk moduli data in the source pool are close to 270 GPa, excepted the largest value (380 GPa) that is recorded in the online database [322], and the smallest value (~220 GPa) that is recorded in the article [115]. Young's moduli generally fall into three ranges: ~380 GPa, ~330 GPa and ~270 GPa. Besides the single crystal measurement that were conducted below ambient temperature (from 4.2 K to 250 K) by Maurer *et al.* [321], the only investigation on single crystal elastic constants at room temperature for Rh is performed by Walker *et al.* [323]. According to their elastic constants, the bulk modulus and the Young's modulus calculated through the VRH averaging method are 267 GPa and 378 GPa, which are very close to the values predicted by ANNs and the correlating equations. Experimental data measured by Walker *et al.* [323], Maurer *et al.* [321] and the values tabulated by Gschneidner [152] from static measurements are also listed in Table 4-16.

Table 4-16 Elastic property of rhodium (Rh).

C_{11}	C_{44}	C_{12}	*E*	*G*	*K*	*v*	Ref
			GPa				
412.6	184.1	193.5	***377.84***	***149.49***	***266.53***	***0.264***	[323]
406.03	185.05	187.45	377.10	149.81	260.31	0.259	[321]*
			372.65	147.10	270.47	0.270	[152]
			379	149.45	270.3	0.26	Equations
			379	149.45	276	0.26	ANNs

* Linear extrapolated to 300 K.

5) Cs, Os, Re, Ru, Sc, Tm and Y (Inconsistent *K* and *ν*)

i. Cesium (Cs)

Cs is a heavy alkali metal, and will transform from bcc to fcc phase under pressure ~2 GPa at room temperature [324]. In the source pool, the unique data of Poisson's ratio (0.356) in contrast to the most common value 0.295 is recorded in CES. The largest two bulk moduli data (>2 GPa) are recorded in CES and Ref. [279], while the smallest data are obtained from Ref. [115]. The only available single crystal elastic constants are determined by Kollarits and Trivisonno [325] at low temperature (2 K and 78 K) by an ultrasonic pulse echo technique. The relatively consistent Young's moduli (~1.7 GPa) and shear moduli (~0.65 GPa) in the source pool, are identical to the values that are calculated by Gschneidner [152] using the experimental bulk modulus and the estimated ratio of Young's modulus to shear modulus. The bulk modulus (~1.7 GPa) of polycrystalline Cs at close room temperature (290 K) measured by Anderson and Swenson [326] is smaller than that measured by Bridgman [327] (~2 GPa). Voronov *et al.* [324] found the shear modulus of polycrystalline Cs was roughly 0.63 GPa.

As described previously, ANN of predicting ***K*** from ***E*** and ***G*** is less accurate, thus its prediction should be used with great caution to compare with other ANNs predictions. This concern should also be extended to the ANN predictions for Os, Re, Ru, Sc, Tm and Y, where the bulk modulus predicted from the shear modulus and the Young's modulus should be excluded in the values finally determined by ANNs. Though the experimental data are smaller than the lower bound of the ANN input dataset, ANNs yield the same results as the correlating equations do. The inconsistence between ANNs predictions and the experimental values is possibly due to the different Young's moduli used as the input: a less than 5% change in the Young's modulus will result a larger than 10% change in the bulk modulus and the shear modulus.

Table 4-17 Elastic property of cesium (Cs).

C_{11}	C_{44}	C_{12}	E	G	K	ν	Ref
GPa							
2.47	1.48	2.06	1.89	0.70	2.19	0.356	[325]*

C_{11}	C_{44}	C_{12}	E	G	K	v	Ref
GPa							
2.12	1.13	1.70	1.59	0.59	1.84	0.356	[325]**
			1.76	***0.65***	***2.03***	***0.356***	[152]
			1.7	0.65	1.4	0.295	Equations
			1.7	0.65	1.4	0.295	ANNs

* At 78 K.

** Values have been linearly extrapolated to 300 K.

ii. Osmium (Os) and Ruthenium (Ru)

Os and Ru belong to the platinum group metals, and possess an hcp structure under ambient conditions of pressure and temperature. Particularly, Os has the highest Young's modulus (~650 GPa) in the Periodic Table. In the source pool, the unique Poisson's ratio value (0.285) from CES is different from the rest of data (~0.25). CES and Ref. [328] also record the unique bulk modulus value 462 GP. For Ru, data discrepancy is surprisingly large: the Poisson's ratio ranges from 0.25 to 0.3, and the bulk moduli spread from 220 GPa to 384 GPa.

Elastic properties of Os and Ru have been determined by Darling [136], and Narayana and Swamy [329]. They all indicate the Young's modulus and the shear modulus of polycrystalline Os should be ~560 GPa and ~220 GPa, respectively. However, much larger values (***E***=648 GPa and ***G***=263 GPa) were obtained by Pantea *et al.* [330] from Os single crystal. The largest controversy that causes disputes in literature is its bulk modulus. The X-ray diffraction experiment conducted by Cynn *et al.* [331] indicates the bulk modulus for Os is 462 GPa, even exceeding that of diamond (~446 GPa). Such result is challenged by Darling [136], Narayana and Swamy [329], Pandey *et al.* [332], Occelli *et al.* [333], Armentrout [127], Pantea *et al.* [330, 334], and Kenichi [335]. Their values range from 380 GPa to 420 GPa, but are all smaller than that of diamond. It is argued that the larger value of bulk modulus determined by Cynn *et al.* is attributable to the inappropriate pressure-transmitting medium (Ar) used in the experiments, introducing systematic error by giving large non-hydrostatic stress [334]. If under hydrostatic condition by replacing Ar to He, the bulk modulus will decrease to ~395 GPa [335].

Table 4-18 Elastic property of osmium (Os).

C_{11}	C_{33}	C_{44}	C_{66}	C_{13}	C_{12}	E	G	K	v	Ref
GPa										
749.5	819.9	259.2	259.9	217.8	229.9	***654.0***	***265.6***	***405.3***	***0.231***	[330]
						558.5	222.5*	380*	0.255*	[329]
						560	220	380	0.25	[136]
						558.75	222.25	379.9	0.255	Equations
						558.75	222.25	379.9	0.253	ANNs

* Estimated data.

In addition to the experiments that were conducted by Darling [136], Narayana and Swamy [329], Pandey *et al.* [332], Brown *et al.* [336] also measured polycrystalline Ru at room temperature. The only single crystal elastic constants for Ru were measured by Fisher and Dever [337] over the temperature range 4 K to 1123 K. Their data together with the results obtained through VRH averaging method using the single crystal elastic constants are reported below. Data predicted from ANNs and that calculated from the correlating equations are consistent for Os, but inconsistent for Ru, probably because the values of the elastic property of Ru are out of ANNs training domain.

Table 4-19 Elastic property of ruthenium (Ru).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	v	Ref
GPa										
562.6	624.2	180.6	187.4	168.2	187.8	***475.2***	***190.8***	***310.7***	***0.245***	[337]
						447	173	365	0.3	[336]
						420	163	327.1	0.286	[152]
						413.7	163*	298.5	0.269*	[329]
						430	172	292	0.25	[136]
						432	173	286.3	0.25	Equations
						432	173	348	0.3	ANNs

* Estimated data.

iii. Rhenium (Re)

Re is an hcp metal with the axial ratio c/a of 1.6 at 300 K, which is slightly smaller than the ideal value 1.633 [337, 338]. Its single crystal elastic constants have been measured by Shepard and Smith [339], Fisher and Dever [337], and Manghnani *et al.* [338] through the dynamic approaches. Polycrystalline data at room temperature have been reported by Brown *et al.* [336] and Darling [136].

In the source pool, the smallest data of Poisson's ratio (0.26) and bulk modulus (334 GPa) for Re are recorded in the handbook [251] and the handbook [143]. The online database [340] also reports the bulk modulus of Re being 334 GPa. The ANNs predictions are in good agreement with the correlating equation calculations. They can be traced back to Brown's experimental values [336], and are very close to the other set of experimental results [337].

Table 4-20 Elastic property of rhenium (Re).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	v	Ref
GPa										
612.6	682.7	162.5	171.4	206	270	461.7	179.2	363.5	0.288	[339]
618.2	683.5	160.6	171.4	207.8	275.3	***460.8***	***178.5***	***366.9***	***0.291***	[337]
617.7	682.8	160.5	171.4	205.5	274.9	460.8	178.6	365.6	0.290	[338]
						463	178.6	378	0.296	[336]
						470	182	379	0.293	[152, 336]
						472	180	340	0.26	[136]
						463	180	365	0.289	Equations
						463	180	365	0.298	ANNs

iv. Scandium (Sc)

Sc, Th and Y are hcp rare earth elements. A completed set of single crystal elastic constants of Sc has been provided by Fisher and Dever [341] over a wide temperature range, followed by Leisure [342] measuring at room temperature. Polycrystalline data at room temperature have been determined by Browns *et al.* [336] and Gust and Royce

[303] through dynamic methods, and tabulated by Gschneidner [152]. It is noted that elastic properties calculated by Scott [119] are slightly different from the VRH results listed below, especially for the shear modulus and the Poisson's ratio, which have 13% and 8% differences. Because Scott did not present the elastic constants he actually used, and the present VRH results are supported by Ref. [342, 343], it is believed that data in Table 4-21 have higher accuracy.

The largest Poisson's ratio (0.309) and bulk modulus (67.1 GPa) in the source pool from Ref. [279] can be traced back to the Browns' measurement [336], and the smallest data ($\boldsymbol{\nu}$=0.258, $\boldsymbol{K}$=55.1 GPa) from Ref. [115] is similar to the result obtained by Gust and Royce [303]. Results yielded by ANNs are generally consistent with results yielded by the correlating equations.

Table 4-21 Elastic property of scandium (Sc).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	ν	Ref
GPa										
98.6	106.2	27.5	26.9	29.5	44.8	75.1	29.4	56.8	0.279	[342]
99.3	106.9	27.7	26.8	29.4	45.7	***75.5***	***29.5***	***57.1***	***0.280***	[341]
						74.4	29.1	56.6	0.278	[119]
						77	29.4	67.2	0.31	[336]
						80.9	31.9	44.4	0.269	[152, 336]
						79.9	31.8	55.2	0.257	[303]
						74.4	29.1	56.6	0.28	Equations
						74.4	29.1	57.3	0.279	ANNs

v. Thulium (Tm)

Very few experiments have been performed to study the elastic property of hcp Tm. Rosen [344] measured high purity polycrystalline Tm at the temperature ranging from 4.2 K to 300 K. Lim *et al.* [345] measured the single crystal elastic constants of Tm as a function of temperature. They both use an ultrasonic pulse technique. However, due to the small sample size, Lim *et al.* [345] did not measure $\boldsymbol{C_{13}}$ because of the difficulty in

preparing parallel faces required for the propagation of ultrasonic waves. Instead, they used an interpolated data. In this case, the VRH averaging values listed in Table 4-22 must be taken with some reservation. The smallest bulk modulus (~37 GPa) [228, 279] recorded in the source pool may be derived from the Bridgman's compression data [119]. The ANN determined data are close to the experimental values recorded in Ref. [344].

Table 4-22 Elastic property of thulium (Tm).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	v	Ref
GPa										
92.5	81.5	28.2	29.5	25*	33.5	73.1	29.3	47.9	0.246	[345]
						74	***30.5***	***44.5***	***0.213***	[344]
						74	30.5	43.6	0.217	Equations
						74	30.5	45	0.235	ANNs

* Interpolated data.

vi. Yttrium (Y)

Yttrium is an hcp rare earth element. Poisson's ratio data in the source pool for Y can be divided in two categories with a variation of 10.42%: ~0.24 and ~0.265. The smallest value of bulk modulus (36.6 GPa) recorded in CES and the largest value (46.9 GPa) recorded in Ref. [279] differ from the rest data listed in the source pool(~41 GPa).

Smith *et al.* [300], Gust and Royce [303], Smith and Gjevre [346] and Savage *et al.* [347] all measured the elastic properties of Y, but only the last two experiments determined the single crystal elastic constants of Y over a wide temperature range. Scott [119] also calculated elastic properties using single crystal elastic constants from Ref. [346]. Because his calculation is very similar to the present results (see Table 4-23), probably his claim that C_{44} and C_{66} were interchanged in Ref. [346] is questionable since no correction is made in this calculation.

It is also noted that the elastic constants presented in Ref. [347] are consistent with the data listed in Ref. [346], excepted for C_{13}, leading significant changes in the calculation of bulk modulus and Poisson's ratio. Though Savage *et al.* [347] attribute the variation

in C_{13} to their higher sample purity, it should be emphasised that this does not necessarily mean their data are more reliable, because their six elastic constants were collected from five different samples containing different impurities. It makes more sense to use data collected from the consistent purity sample as presented by Ref. [346]. The elastic properties predicted by ANNs are equal to the values calculated from the correlating equations, and close to the experimental values in Ref. [300, 346].

Table 4-23 Elastic property of yttrium (Y).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	ν	Ref
GPa										
77.9	76.9	24.31	24.7	21	28.5	***63.3***	***25.4***	***41.5***	***0.246***	[346]
79.10	78.70	24.66	24.80	32.60	29.47	62.2	24.3	47.4	0.281	[347]
						66.3	26.2	46.92	0.265	[300]
						60.9	23.5	49.75	0.296	[303]
						63.5	25.6	41.2	0.243	[119, 346]
						64.4	25.6	46.9	0.242	Equations
						64.4	25.6	46.9	0.246	ANNS

6) Th (Inconsistent *E* and *ν*)

Thorium (Th) has the fcc structure at room temperature. The two properties recorded in source pool that have sparse data are Young's modulus and Poisson's ratio. Ref. [348] presents a much smaller value for Young's modulus (58.6 GPa), comparing to the most common value 78.3 GPa, while the largest Poisson's ratio 0.3 is recorded in CES.

The elastic properties of polycrystalline Th were first dynamically measured by Reynolds [349]. Carlson *et al.* [350] adopted the bulk modulus data from Reynolds' result, but used different values for the other three elastic properties, so his data do not obey the relationships between the elastic property (as shown in Table 1-2). The single crystal elastic constants of Th were first determined by Armstrong *et al.* [351], followed by Greiner *et al.* [352] measuring over temperature range 4.2 K ~ 300 K. Their data together with the VRH averaging results are listed in Table 4-24. Though ANNs yields the similar results as employing the correlating equations, it should be mentioned that

the experimental values in Ref. [352] are recommended to be the best values to describe the elastic behaviour of Th. The differences between ANNs predictions and the experimental values are due to the different shear moduli and bulk moduli using as inputs.

Table 4-24 Elastic property of thorium (Th).

C_{11}	C_{44}	C_{12}	E	G	K	ν	Ref
GPa							
75.3	47.8	48.9	73.74	28.65	57.70	0.287	[351]
77.02	45.54	50.88	***71.94***	***27.70***	***59.59***	***0.30***	[352]
			72.8	28	60.67	0.3	[349]
			72.4	27.6	60.67	0.27	[350]
			78.3	30.80	54.00	0.26	Equations
			75.65	30.80	54.00	0.263	ANNs

7) Ho and Zr (Inconsistent *E*, *K* and *ν*)

i. Holmium (Ho)

Ho is a heavy rare earth metal that crystallizes in hcp structure at room temperature. Apart from the dynamic measurement carried out on polycrystalline Ho [301, 302], Palmer and Lee [288], Salama *et al.* [353] and Rosen *et al.* [354] determined the elastic constants of Ho over a wide temperature range. Their bulk modulus are larger than the adiabatic bulk modulus (39.6 GPa) converted by Scott [119] from Bridgeman's isothermal compressibility data obtained by the static method [307]. Because Palmer and Lee made corrections for density and acoustic path length, their results are believed more accurate [119].

In the source pool, only the shear moduli are documented in accordance with each other (~26.3 GPa). Poisson's ratio can be generally divided into three categories: ~0.231, ~0.255 and ~0.272. The largest bulk modulus (48.9 GPa) recorded in Ref. [115] is approximately 20% larger than the smallest value (39.7 GPa) from CES. CES also

provides the largest Young's modulus (72 GPa), which is close to the value obtained at 0 K rather than 300 K [288].

Table 4-25 Elastic property of holmium (Ho).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	ν	Ref
GPa										
76.1	77.6	25.7	25.65	20.6	24.8	***64.8***	***26.3***	***40.2***	***0.231***	[288]
76.12	80.15	25.92	25.06	20.72	26	65.0	26.3	40.8	0.235	[353]
77.31	81.24	26.13	25.63	26.11	26.84	65.3	26.1	43.8	0.251	[354]*
						66.3	27.3	39.06	0.216	[301]
						66.9**	26.4**	48.4***	0.269***	[302]
						67.1	26.7	45.83	0.255	[300]
						64.8	26.3	40	0.23	Equations
						65	26.3	39.7	0.255	ANNs

* Elastic constants are all read off from graph.

** Read from graph.

*** Calculated from the correlating equations.

ii. Zirconium (Zr)

Zr transforms from the bcc structure (β phase) to the hcp structure (α phase) at 1135 K. The elastic properties of polycrystalline Zr have been measured dynamically by Reynolds [349] and Myers [355]. Fisher and Renken [297, 356, 357] conducted a series of experiments over a wide temperature ranging from 4 K up to 1155 K on Zr single crystal. The results from their last measurements are taken as the best values.

Examination in the source pool indicates the Poisson's ratio, bulk moduli and Young's moduli of Zr drop in two categories: ~0.34 and ~0.38 for ν, ~89.9 GPa and ~95.3 GPa for K, and ~68 GPa and ~98 GPa for E, respectively. Only the shear moduli are generally consistent within 10% variation (~35 GPa). As shown in Table 4-26, the ANN predictions agree well with the experimental values and the calculations utilizing the correlating equations.

Table 4-26 Elastic property of zirconium (Zr).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	v	Ref
GPa										
143.5	164.9	32.07	35.5	65.4	72.5	96.2	36.1	95.3	0.332	[356]
143.4	164.8	32	35.3	65.3	72.8	96.0	36.0	95.3	0.332	[297]
143.68	165.17	32.14	35.32	65.88	73.04	***96.1***	***36.1***	***95.7***	***0.333***	[357]
						92.05	34.45	91.1	0.34	[349, 356]
						91	34.3	94.9	0.339	[355]
						94.95	35	95.3	0.332	Equations
						95.3	35	95.3	0.332	ANNs

8) Hf, Na and Pu (Inconsistent *E*, *G* and *v*)

i. Hafnium (Hf)

Hf possesses an hcp structure at room temperature. Gschneidner [152] summarized the elastic properties of polycrystalline Hf in 1964, later Fisher and Renken [297] measured single crystal elastic constants of Hf from 4 K to 300 K. In the source pool, the bulk moduli are in excellent agreement (~109 GPa). However, Poisson's ratio varies from 0.26 to 0.37, the corresponding Young's moduli varies from 141 GPa to 78 GPa, and the shear moduli varies from 56 GPa to 30 GPa. The values predicted from ANNs are consistent with the values obtained from the correlating equations, and are close to the experimental results [297].

Table 4-27 Elastic property of hafnium (Hf).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	v	Ref
GPa										
181.1	196.9	55.7	52	66.1	77.2	***143.0***	***55.8***	***108.6***	***0.281***	[297]
						137.29	52.96	108.85	0.3	[152]
						143	55.8	109	0.282	Equations
						140	56	109	0.282	ANNs

ii. Sodium (Na)

Na is an alkali metal with a bcc structure. The earliest experimental measurement on single crystal Na was probably carried out by Quimby and Siegel at 80 K [358]. Daniels [359] extrapolated their elastic constants to 300 K, and found good agreement on C_{44} with his own investigation but large differences in C_{11} and C_{12}, owing to the possible errors introduced by the extrapolation. Martinson [360] studied the elastic constants of Na at a wider temperature range, and presented similar values. Hence, his data are chose to the best values. It should be noted that the elastic properties of Na compiled by Gschneidner [152] are based on single crystal measurements at 90 K (excepted for the bulk modulus).

In the source pool, the values of Poisson's ratio for Na were 0.315, 0.34 and 0.366. Though the 9.5% variation of bulk modulus is somewhat smaller than the inconsistent judgment criterion (10% variation), it ranges from 6.3 GPa to 6.8 GPa. The data of Young's moduli and shear moduli can generally be divided into three categories: 5 GPa, 6.8 GPa and 10 GPa for Young's moduli, and 1.98 GPa, 2.53 GPa, 3.3 GPa for shear moduli.

As shown in Table 4-28, elastic properties predicted from ANNs, which are close to the experimental value measured at 80 K [358] rather than that at room temperature, are different from the values calculated using the correlating equations. The differences are caused by two factors: i) ANN models are less accurate than the correlating equations listed in Table 1-2; ii) The elastic data of Na are smaller than the lower bound of dataset that was used to construct ANNs, thus, ANNs extrapolation is less accurate.

Table 4-28 Elastic property of sodium (Na).

C_{11}	C_{44}	C_{12}	E	G	K	ν	Ref
GPa							
7.36	5.68	5.94	6.86	2.60	6.41	0.322	[358]*
5.21	4.17	4.01	5.18	1.99	4.41	0.304	[358, 359]**
7.38	4.19	6.21	5.40	1.98	6.60	0.364	[359]
7.69	4.31	6.47	***5.58***	***2.05***	***6.88***	***0.365***	[360]

C_{11}	C_{44}	C_{12}	E	G	K	v	Ref
			GPa				
			8.94	3.43	6.81	0.315	[152]&
			5.41	1.98	6.52	0.366	Equations
			6.8	2.53	6.52	0.315	ANNs

* At 80 K.

& At 90 K.

** Values are extrapolated to 300 K.

iii. Plutonium (Pu)

Plutonium is in alpha phase with monoclinic structure at room temperature [361]. Because α-Pu is brittle during the tension test and relatively ductile in the compression test [362], elastic property data obtained through a dynamic method is much more reliable. A summary of experimental measurements on polycrystalline α-Pu elastic properties has been reported by Migliori *et al.* [361]. The completed sets of experimental data are quoted in Table 4-29. A recent measurement [363] over a wider temperature range (300 K to 750 K) is also included in the table and is deemed to provide the best values.

Compared to the unique bulk modulus data for Pu in the source pool, i.e. 54 GPa (from CES), the other three elastic properties are supplied by more sources, though large disparities exist. Roughly, the Young's moduli of Pu are recorded close to 87.5 GPa or 96 GPa, the shear moduli are recorded close to 34.5 GPa or 43 GPa, and Poisson's ratio are recorded close to 0.15, 0.18 or 0.21, respectively. The ANNs predictions agree well with the calculations.

Table 4-29 Elastic property of plutonium (Pu).

E	G	K	v	Ref
GPa				
98.3	40.9	54.3	0.199	[361, 364]
100.7	42.3	53.4	0.186	[361, 365]

E	*G*	*K*	*v*	Ref
GPa				
99.3	43.3	46.7	0.15	[361, 366]
107	44.5	59.9	0.202	[361, 367]
107.5	46	54.1	0.169	[361, 368]
92.8	39.1	49.4	0.187	[361, 369]
108	45.9	55.9	0.178	[361]
103.09	***43.46***	***54.72***	***0.186***	[363]
91.75	34.5	54	0.21	Equations
87.5	34.5	54	0.21	ANNs

9) Ce, In, K and Tl (Inconsistent *E*, *G* and *K*)

i. Cerium (Ce)

Ce is in γ phase with fcc structure at room temperature [300]. Most elastic properties measurements are conducted on polycrystalline Ce [300, 301, 303]. Only one experiment is undertaken to determine the single crystal elastic constants of γ-Ce, which was using ultrasonic pulse technique. The values of bulk modulus and Poisson's ratio derived from the single crystal elastic constants through the VRH averaging method are somewhat smaller than the values reported for polycrystalline γ-Ce.

Table 4-30 Elastic property of cerium (Ce).

C_{11}	C_{44}	C_{12}	*E*	*G*	*K*	*v*	Ref
GPa							
26.01	17.30	14.26	***27.94***	***11.23***	***18.18***	***0.244***	[370]
			30	12	19.81	0.248	[300]
			33.7	13.6	21.55	0.24	[301]
			30	11.8	21.74	0.269	[303]
			33.6	13.53	22	0.245	Equations
			33.5	14	21.5	0.245	ANNs

Poisson's ratio data in the source pool are in great agreement for Ce, i.e. 0.245. However, the largest bulk modulus (26.2 GPa) from CES is about 30% higher than the smallest data (19.8 GPa) that is recorded in Ref. [279], and is about 20% higher than the most common value 22 GPa. The unique data of Young's modulus (30 GPa) and shear modulus (~12 GPa) from CES and Ref. [279], are the smallest values recorded in the source pool. As shown in Table 4-30, the ANN predictions are consistent with the values calculated from the correlating equations.

ii. Indium (In)

Indium has an unusual face-centred-tetragonal structure with axial ratio c/a extended to 1.08 [371]. The single crystal elastic constants of indium were measured through ultrasonic pulse method by Winder and Smith [371] at room temperature, Chandrasekhar and Rayne [372] from 1.4 K to 300 K, and Vold *et al.* [281] from room temperature to the melting point. Their VRH averaging results are in satisfactory agreement with the elastic properties directly measured using high purity (99.99%) polycrystalline indium by Kim [373]. However, the values are quite different from the data compiled by Gschneidner [152], who was using data measured by Koster [141, 142] and Bridgman [374].

Table 4-31 Elastic property of indium (In).

C_{11}	C_{33}	C_{44}	C_{66}	C_{13}	C_{12}	E	G	K	ν	Ref
GPa										
44.5	44.4	6.55	12.2	40.51	39.44	13.8	4.8	41.6	0.445	[371]
45.35	45.15	6.51	12.07	41.51	40.06	***13.6***	***4.7***	***42.4***	***0.447***	[372]
45.1	45.3	6.53	11.9	41.1	39.7	13.9	4.8	42.1	0.445	[281]
						12.74	4.394	42.33	0.4498	[373]
						10.49	3.73	41.08	0.460	[152]
						13.8	4.78	41.1	0.45	Equations
						14	3.68	42	0.45	ANNs

In the source pool, Poisson's ratio data for indium is very consistent and equals to 0.45. However, the other three elastic properties data roughly fall into two categories: 36.4 GPa and 42 GPa for bulk modulus, 11 GPa and 14 GPa for Young's modulus, and 3.68 GPa and 4.78 GPa for shear modulus, respectively.

iii. Potassium (K)

Bender [375] first measured the single crystal elastic constants bcc potassium at 83 K through a static method. The elastic constants measured by Smith *et al.* [376] at room temperature using ultrasonic pulse technique are in excellent agreement with the lower temperature values (195 K to 4.2 K) when extrapolated to room temperature [377]. The difficulty of using dynamic methods lies in the determination of the specimen length due to the mechanical softness of potassium. The bulk modulus calculated from the VRH averaging method using Smith's data agrees well with Bridgman compressibility data obtained from the hydrostatic measurements on polycrystalline samples [152].

Table 4-32 Elastic property of potassium (K).

C_{11}	C_{44}	C_{12}	E	G	K	ν	Ref
			GPa				
4.59	2.63	3.72	3.55	1.31	4.01	0.35	[375]*
3.7	1.88	3.14	***2.49***	***0.91***	***3.33***	***0.38***	[376]
					3.18		[152]
			3.53	1.3	4.2	0.35	Equations
			3.53	1.3	4.2	0.35	ANNs

* At 83 K.

In the source pool, three sources provide Poisson's ratio data for potassium, and they are relatively consistent, i.e. 0.35. Excepted the largest value (4.2 GPa) in Ref. [279], the bulk moduli provided by other sources are almost the same (~3.1 GPa). The Young's moduli recorded in CES and Ref. [115] (~2.4 GPa) are much smaller than the rest of data (3.53 GPa). The smallest value of shear modulus (0.9 GPa) also comes from Ref. [115], comparing to the most common value 1.3 GPa. The use of Poisson's ratio 0.35 is

possibly the main reason that the ANN prediction and the calculation from correlating equations are all close to the elastic property data obtained in Ref. [375].

iv. Thallium (Tl)

Tl crystallizes in an hcp structure at room structure. To the author's knowledge, the only single crystal elastic constants measurement were conducted by Ferris *et al.* [378] at temperature ranging 4.2 K to 300 K. The bulk modulus calculated through the VRH averaging method is very close to the value derived from Bridgman compressibility data [374]. However, the Young's modulus, shear modulus and Poisson's ratio calculated from the elastic constants are substantially different from the data obtained by Koster [141, 142] from static measurement on polycrystalline samples. A compilation made by Gschneidner's [152] contains the data referring to Bridgman and Koster's results.

The Poisson's ratio data for Tl recorded in the source pool is quite consistent, but the other three elastic properties are not. The bulk moduli for Tl range from 28.5 GPa to 43 GPa, and the Young's moduli range from 8 GPa to 15.42 GPa. The largest shear modulus 5.4 GPa from Ref. [115] is almost twice larger than the rest of the available values. The large Poisson's ratio (outside of the ANN training domain) is probably the reason that leads to the variation of bulk modulus predicted from ANNs and the bulk modulus calculated using the correlating equations.

Table 4-33 Elastic property of thallium (Tl).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	ν	Ref
GPa										
40.8	52.8	7.26	2.7	29	35.4	***15.3***	***5.3***	***35.6***	***0.429***	[378]
						7.94	2.75	35.92	0.460	[141, 142, 152, 374]
						15.42	5.4	35.7	0.45	Equations
						15.42	5.4	28.5	0.45	ANNs

10) Be, Cd, Eu, Ga, La, Lu, Rb, Sm, U and Yb (Inconsistent *E*, *G*, *K* and *v*)

i. Beryllium (Be)

The five independent elastic constants of beryllium hcp single crystal have been determined through dynamic approach by Gold [379], Smith and Arbogast [380], Silversmith and Averbach [381], Rowlands and White [382], and Migliori *et al.* [383]. Only Migliori *et al.* investigated the elastic properties of polycrystalline Be as well, thus their results are deemed to be the best values with internal consistencies. The marked disparity between the values reported by Gold and others is attributable to Gold's highly questionable assumption that the specimen is sufficiently isotropic [380]. With the exception of Gold's data, major variances lie in C_{12} and C_{13}, the accuracy of which depend on the established values of C_{11}, C_{33} and C_{44} and the propagation angle in the measurements [382].

Table 4-34 Elastic property of beryllium (Be).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	v	Ref
				GPa						
308	357	110	183	87	-58	293.2	131.4	127.1	0.115	[379]
292.3	336.4	162.5	132.8	14	26.7	311.1	148.6	114.4	0.047	[380]
295.4	356.1	170.6	134.8	-1	25.9	317.6	155.6	110.4	0.020	[381]
288.8	354.2	154.9	134.35	4.7	20.1	306.9	148.4	109.8	0.034	[382]
293.6	356.7	162.2	133.4	14	26.8	***315.2***	***150.1***	***116.8***	***0.050***	[383]
						313.8	149.2	116.6	0.0516	[383]
						287.25	128.4	125.57	0.118	Equations
						287	128.4	130	0.118	ANNs

The variation in Poisson's ratio is probably the result of specimen's anisotropy [384]. The fact that large discrepancies of elastic properties exist in original literature is reflected in the source pool. The Poisson's ratio of Be varies from 0.02 to 0.118, its bulk moduli vary from 110 GPa to 130 GPa, its Young's moduli vary from 287 GPa to 318

GPa, and its shear moduli vary from 128.4 GPa to 156 GPa. Due to the small Poisson's ratio (outside of the ANN training domain), the values predicted from ANNs point to Gold's less accurate data, though they are consistent with the values calculated utilizing the correlating equations.

ii. Cadmium (Cd)

The atoms of cadmium are arranged in hexagonal lattices at room temperature, exhibiting anomalously large axial ratio (c/a=1.886) [385]. The single crystal elastic constants of Cd measured by Bridgman [386] through the static approach should only serve as the preliminary results due to his intention to discard the inconsistent data on comparing with those obtained by Gruneisen and Goens [387]. The other two sets of single crystal elastic constants of Cd were measured through an ultrasonic pulse technique by Garland and Silverman [385] from 4.2 K to 300 K, and by Chang and Himmel [388] from 300 K to its melting point 575 K. Their data are in good agreement with polycrystalline data compiled by Gschneidner [152], except for Poisson's ratio and bulk modulus. Because Garland and Silverman used a less accurate density value in their calculation, data obtained by Chang and Himmel are recommended to be the best values.

Table 4-35 Elastic property of cadmium (Cd).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	v	Ref
GPa										
109	45.9	15.6	34.5	37.5	40	54.9	20.9	49.6	0.316	[386]
121	51.3	18.5	36.5	44.2	48	60.5	22.9	56.6	0.322	[385, 387]
115.8	51.4	20.39	38.01	40.6	39.75	63.5	24.4	53.6	0.303	[385]
114.5	50.85	19.85	37.5	39.9	39.5	***62.6***	***24.0***	***53.0***	***0.303***	[388]
						62.27	24.12	46.74	0.300	[152]
						62.54	24	52.9	0.302	Equations
						62.3	24.6	51	0.303	ANNs

The unique value of Poisson's ratio 0.39 compared to the most common values 0.3 in the source pool is recorded in Ref. [279]. The other three elastic properties generally fall into two or three categories: 19 GPa and 24 GPa for shear modulus; 42 GPa, 46.7 GPa, and 51 GPa for bulk modulus; and 50 GPa, 55 GPa and 62 GPa for Young's modulus. The values predicted from ANNs are in good agreement with the values calculated from the correlating equations and the experimental data.

iii. Europium (Eu)

In contrast to most rare earth metals crystalizing in an hcp structure, Eu has a bcc structure. Though no single crystal elastic constant data has been found for Eu, extensive elastic properties data of polycrystalline Eu have been measured through ultrasonic wave propagation technique over a wide temperature range by Rosen [301, 389], Burkhanov *et al.* [390], and Gust and Royce [303]. Their results are in good agreement with the data of Young's modulus measured by Bodryakov and Nikitin [391] through the vibration method. As Scott [119] pointing out, the smallest data obtained by Rosen appear to be obtained from the highest purity specimen. Scott's recommendation is reflected in the source pool excepted Ref. [279] and CES. The results predicted by the ANNs for bulk modulus and Poisson's ratio are almost twice larger than the values obtained in the experiments or that obtained from the correlating equations. Again, it indicates that ANN extrapolation outside its problem domain is less reliable, when the values are smaller than the lowest bound of the dataset that was used to construct ANNs.

Table 4-36 Elastic property of europium (Eu).

E	***G***	***K***	***v***	**Ref**
	GPa			
18.2	***7.9***	***8.85***	***0.155***	[301, 389]
21.40	8.70	13.16	0.230	[390]
19.6	7.5	17.06	0.308	[303]
20.8				[391]
18.2	7.9	8.7	0.152	Equations
18.2	7.9	14.7	0.286	ANNs

iv. Gallium (Ga)

Due to the low melting temperature 303 K of Ga, the elastic properties of the base-centred orthorhombic Ga were usually measured below room temperature [392]. A completed set of single crystal elastic constants for Ga has been dynamically measured by Lyall and Cochran [393] at 273 K, 77 K, and 4.2 K. It is in good agreement with the extrapolation obtained by Langill and Trivisonno [394] from 4.2 K to 190 K. Lyapin *et al.* [395] investigated the pressure dependence of polycrystalline gallium at temperature range 240 K to 360 K, in which region contains orthorhombic Ga, bcc Ga and liquid Ga. Their data for orthorhombic Ga are slightly larger than the VRH averaging values obtained using Lyall and Cochran's results, or the data compiled by Gschneidner [152]. Further study on the elastic properties of Ga is needed in order to narrow down the discrepancies.

In the source pool, the most common data of Young's modulus is equivalent to ~9.8 GPa, approximately 10 times less than the experimental value and data recorded in Ref. [115], hence is highly suspect as a result of wrongly placed decimal point or unit, i.e. should be dynes/cm^2 rather than GPa. In this case, no surprise that only the shear modulus recorded in Ref. [115] as well as the Poisson's ratio recorded in Ref. [115] and CES are close to the experimental data 40 GPa and 0.23, respectively, not to mention bulk moduli varying from 35 GPa to 58.2 GPa.

v. Lanthanum (La)

La is stable in an hcp structure at ambient conditions, but will transfer into fcc or bcc structure with increased temperature or pressure [396]. No single crystal elastic constant data are available for hcp La, but the elastic properties of polycrystalline La at room temperature have been dynamically measured by Smith *et al.* [300], Rosen [301], and Gust and Royce [303]. Literature values are summarized in Table 4-37. As suggested by Scott [119], Gust and Royce gave the best experimental values. Most elastic data in the source pool are identical to the values obtained by Gust and Royce or Gschneidner, except the bizarre value recorded in Ref. [115].

Table 4-37 Elastic property of lanthanum (La).

E	***G***	***K***	***v***	**Ref**
	GPa			
33.7	13.6	16	0.166	[119, 301]
38.4	14.9	30.3	0.288	[119, 300]
36.6	***14.3***	***27.9***	***0.28***	[119, 303]
37.95	14.91	24.29	0.288	[152]
37	14.25	28	0.284	Equations
36.8	14.15	30.3	0.284	ANNs

Table 4-38 Elastic property of gallium (Ga).

C_{11}	C_{22}	C_{33}	C_{44}	C_{55}	C_{66}	C_{12}	C_{13}	C_{23}	E	G	K	ν	Ref
GPa													
101.9	92.2	137.6	35.6	42	41.3	44.5	27.6	24	***90.99***	***38.4***	***48.1***	***0.23***	[393, 395]*
103.74	93.39	140.38	36.78	42.85	41.69								[394]*
									103.43	42.38	61.63	0.21	[395]**
									102.33	41.73	62.25	0.23	[395]§
									92.57	37.46	56.88	0.235	[152]*
									93.2	37.8	58.2	0.233	Equations
									93.2	37.8	58.2	0.235	ANNs

* Data are obtained at 273 K.

** Data are obtained at 268 K, but the values of C_{12}, C_{12}, and C_{23} are not available.

§ Read from graph at 283 K.

vi. Lutetium (Lu)

The single crystal elastic constants of hcp Lu at room temperature have been measured by Tonnies *et al.* [397] and Greiner *et al.* [129]. Tonnies *et al.* also investigated the temperatures dependences, and Greiner *et al.* focused on the study of impurities influence, especially hydrogen. It should be noted that Greiner *et al.* did not determine the off-diagonal constant C_{13}. Therefore, the value obtained by Tonnies *et al.* [397] is used for the VRH averaging method. However, other values reported are estimated by Gschneidner [152] based on the assumption that the Young's modulus is proportional to the shear modulus and utilizing Bridgman compressibility data [307]. The Young's modulus given by Gschneidner is about 14% smaller than the VRH averaging value. Apparently, the data in the source pool are close to one of two set data. To be specific, CES seems to use the estimated data from Gschneidner, others all adopt the results provided by Tonnies *et al.* [397].

Table 4-39 Elastic property of lutetium (Lu).

C_{11}	C_{33}	C_{44}	$C_{66}=(C_{11}-C_{12})/2$	C_{13}	C_{12}	E	G	K	ν	Ref
GPa										
86.2	80.9	26.8	27.1	28	32	***68.4***	***27.2***	***47.6***	***0.261***	[397]
88.19	82.59	27.11	27.4	28	33.39	69.6	27.6	48.6	0.261	[129]
						84.34*	33.83*	41.12	0.233*	[152]
						61.5	24.41	42.7	0.26	Equations
						61.5	24.41	42.6	0.26	ANNs

* Estimated data.

vii. Rubidium (Rb)

A completed set of single crystal elastic constants of Rb with a bcc structure has been determined by Roberts and Meister [398] at 80 K, and Gutman and Trivisonno [399] from 78 K to 170 K. The mechanical softness of Rb made it difficult to generate shear waves at room temperature, i.e. lack of sufficiently rigid bond between transducer and specimen in the acoustic study. Those data are listed in Table 4-40 including data that are linearly extrapolated to 293 K.

It is worth to point out that the extrapolated data of C_{11} at 293 K (2.58 GPa) is in good agreement with the experimental data obtained by Roberts and Meister (2.41±0.12 GPa), and the value of bulk modulus calculated from the VRH averaging method also agrees well with the experimental compression data given by Anderson and Swenson at 295 K [400]. With the data of bulk modulus determined by Bridgeman in 1931's, Koster [140] interpolated the elastic properties of Rb by assuming a ratio of bulk modulus to Young's modulus equals to 0.8. However, the data given by Bridgman before 1940's were found to be incorrect. The estimated values given by Gschneidner [152] were based on the bulk modulus data derived from Bridgeman 1948's measurement and an assumption of a constant ratio (2.71) of Young's modulus against shear modulus. Consequently, the linearly extrapolated data at 293 K from single crystal elastic constants are probably more reliable. In the source pool, excepted the data recorded in Ref. [115], which is very close to my extrapolated value, other values either refer to the low temperature experimental results or use estimate data from Koster or Gschneidner.

Table 4-40 Elastic property of rubidium (Rb).

C_{11}	C_{44}	C_{12}	E	G	K	ν	Ref
			GPa				
2.96	1.6	2.44	2.16	0.79	2.61	0.362	[398]*
3.25	1.98	2.73	2.49	0.92	2.90	0.357	[399]§
2.96	1.71	2.5	2.18	0.80	2.65	0.363	[399]**
2.58	1.36	2.19	***1.78***	***0.65***	***2.32***	***0.372***	[399]§§
			2.35	0.91	1.86	0.29	[140]
			2.72	1.00	3.14	0.356	[152]
			1.73	0.63	2.3	0.374	Equations
			2.4	1.02	2.3	0.3	ANNs

* At 80 K.

§ At 78 K.

** At 170 K

§§ Extrapolated to 293 K

viii. Samarium (Sm)

No single crystal elastic constants data available for Sm, probably due to its low symmetry rhombohedral structure [401]. However, polycrystalline elastic properties of Sm have been extensively studied through dynamic approach by Smith *et al.* [300], Rosen [301, 306] (from 4 K to 300 K), and Gust and Royce [303]. The compilation made by Gschneidner [152] adopted the elastic property data obtained by Smith *et al.* except for the bulk modulus, which was obtained by Bridgman through a static approach.

The inconsistent four experimental results lead to the discrepancies of elastic property data recorded in the source pool. It seems that CES and Ref. [279] are using Gschneidner's data, while others are using data listed by Gust and Royce or Rosen. Nevertheless, Scott [119] suggested using the Rosen' values, because there are more experimental details available in literature.

Table 4-41 Elastic property of samarium (Sm).

E	***G***	***K***	***v***	**Ref**
	GPa			
34.1	12.6	38.31	0.352	[300]
48	***18.7***	***37.88***	***0.284***	[301, 306]
50.6	19.9	37.31	0.272	[303]
34.13	12.65	29.40	0.352	[152]
49.7	19.6	35.7	0.268	Equations
49.7	19.55	35.7	0.268	ANNs

ix. Uranium (U)

α-U with a orthorhombic structure is stable up to 935 K then changes to tetragonal β-U [402]. Polycrystalline elastic properties of α-U have been investigated by Rosen [403], Armstrong *et al.* [404] and Abey and Bonner [405] over a wide temperature and pressure range. Fisher *et al.* [402, 406-408] conducted a series of dynamic study on the temperature dependence of single crystal elastic constants of U from 2 K to 923 K. A

variation as large as 30% in the experimental data is the main cause of the disparity that is noticed in the source pool.

Table 4-42 Elastic constants of uranium (U).

Unit	C_{11}	C_{12}	C_{13}	C_{22}	C_{23}	C_{33}	C_{44}	C_{55}	C_{66}	Ref
(GPa)	214.7	46.5	21.8	198.6	107.6	267.1	124.44	73.42	74.33	[402, 406-408]

Table 4-43 Elastic property of uranium (U).

E	*G*	*K*	*v*	Ref
GPa				
217.09	***79.90***	***255.80***	***0.359***	[402, 406-408]
186.01*	80.25*	90.88*	0.159**	[403]
201.00	81.40	126.25**	0.23	[404]
204.66**	86.00	110.00	0.190**	[405]
186.33	73.55	98.75	0.245	[152]
175.8	74.05	100	0.205	Equations
175.8	73.1	97.9	0.24	ANNs

* Read off from graph.

**Calculated using Equation 1-21.

x. Ytterbium (Yb)

Like Eu, Yb has an fcc structure in contrast to the common hcp structure for rare earth elements. However, there is no single crystal elastic constant measurement for Yb. The elastic property data obtained by Smith *et al.* [300] and Gust and Royce [303] through wave propagation technique for polycrystalline Yb, are very close to the data obtained by Bridgman [307], but quite different from the values obtained by Rosen [301, 344]. Such disagreement is evidently reflected in the source pool.

Data in CES and Ref. [279] refer to the results of Smith *et al.*, others in the source pool may be traced back to Rosen's experimental values, except for the bulk modulus (~31 GPa), which probably mistook 31 for 13. It should be emphasised that this mistake has been published repeatedly in several sources including CRC handbook. It is a typical example where errors passing along like genetic mutation to the subsequent publications. Scott [119] recommended to use Rosen's data because it is the only work included the temperature dependence. Nevertheless, elastic property recently investigated by Boguslavskii *et al.*[409] suggests their room temperature values are in good agreement with the experimental results obtained by Smith *et al.* [300]. Because Boguslavskii *et al.* described their results in a graph rather than in the form of numeric data, the experimental data given by Rosen are chose to be the best values.

Table 4-44 Elastic property of ytterbium (Yb).

E	***G***	***K***	***v***	**Ref**
	GPa			
17.8	7	13.77	0.284	[300]
23.8	***9.9***	***13.61***	***0.207***	[301, 344]
18.8	7.3	14.97	0.29	[303]
17.77*	6.93**	13.54**	0.281*	[409]
23.93	9.95	13.2	0.205	Equations
18	7.1	13.8	0.284	ANNs

* Read from graph.

** Calculated using Equation 1-21.

4.3.4 Factors that influence elastic properties

1) Theoretical factors

The magnitudes of ***E***, ***G***, ***K*** and ***v*** ultimately depend on the element electronic configuration and position in the Periodic Table. The most important assumption made in determining elastic property of polycrystalline is that the grains are sufficiently numerous and randomly oriented so the specimen may be approximated as elastic

isotropic [300]. However, elastic properties can show remarkable variation with crystal orientation [410].

As mentioned in Section 1.6, the ratio of Young's modulus to the shear modulus is roughly a constant which equals to 3/8 [154]. Another method to estimate an elastic property is to utilize the value of bulk modulus and the ratio of Young's modulus to shear modulus, when only the bulk modulus are available [152]. However, this method not only introduces errors by assuming Young's modulus is proportional to shear modulus, but also highly depends on the accuracy of bulk modulus. In fact, a large number of compressibility data and corresponding bulk modulus data in literature are collected from Bridgman's static measurements. However, very few people are aware that Bridgman made corrections on his previous data, as errors were still retained in recent publications [152].

2) Experimental factors

i. Purity

Addition of boron reduces osmium's bulk modulus from 421 to 365 GPa [127]. Greiner *et al.* investigated the effect of the carbon addition to Th single crystal, and found it strengthened dilatational deformation and torsional deformation along <110> directions, but weaken torsional deformation along <100> directions [352]. Ashida *et al.* [128] studied the elastic property changes on Zr caused by hydrogen concentration, and found the Young's modulus and the shear modulus decrease linearly with increasing hydrogen concentration in the hcp phase (below 1135 K) but increase linearly in the bcc phase. Comparisons made by Greiner *et al.* on Lu single crystals indicate that the impact of interstitial impurities (e.g. H, C, N, and O) on the elastic constants attributes to lattice stiffen by increasing bounding [129].

ii. Temperature

At atmosphere pressure, the elastic constants as well as the Young's modulus and the shear modulus appear to decrease linearly with the increasing temperature, and may exhibit an abruptly decrease at phase transition temperature [360]. At lower temperature, the non-linear variation with temperature is not related to the effects of lattice defects [411], but ascribed to anharmonic vibration of metal atoms [128]. Vold *et al.* [281]

noted the elastic anisotropy increase dramatically with increasing temperature in In single crystal. Burkhanov *et al.* [390] observed a sudden change in the elastic property of Eu took place at antiferromagnetic transitions temperature due to the existence of considerable magnetostriction stress and possible crystallographic symmetry change. Rosen [389] believed the peak anomalies in ultrasonic attenuation near the Neel point are the result of interaction between critical fluctuations of the spins and acoustical phonons. Apart from antiferromagnetic ordering, electron-type transition and temperature-dependent crystallographic phase change may also cause abnormal changes in the temperature dependence of elastic properties [306].

iii. Mechanical processing

The elastic constants of single crystal are quite sensitive to its past processing history, which will affect specimen geometry, texture, surface condition, and internal defects. Elastic property measurement of γ-cerium exhibits differences if the specimen has been previously cooling down to liquid helium temperature and then warming up to room temperature. Such process will produce a small amount of β phase content [412]. In order to satisfactorily removed a finite amount of residual strain introduced in specimen preparation, crystals should be annealed before the tests [380].

iv. Static or dynamic measurements

Experimental uncertainties might lie in the transit-time determination and possible dislocation effects [372]. A systematic error (known as 'transit – time error') may occur in the ultrasonic pulse-echo method due to the acoustic reflections. The acoustic reflections gradually change the shape of a pulse and make it difficult to identify correct cycles in the successive echoes at the quartz-specimen interfaces [371]. Adoption of a improved technique that avoids the direct measurement of delay time [381], or a modified experimental apparatus (such as the buffer technique [377]), could eliminate the potential error. Transit-time may need to be corrected when actual directions in measurement substantially deviated from the propagation and polarization directions indicated in sound velocity calculations [372]. Transit time reading obtained during the heating cycle may also differ from that obtained on cooling due to the specimen temperature hysteresis effects [382]. A method to correct transit-time error arising from transducer loading is outlined in Ref. [413].

Data accuracy also depends on the precision in measuring the sample dimensions which control the density and propagation path length in wave propagation technique [382]. The sample length [381], density [388], as well as angle of wave velocity propagation [385], may need appropriate correction using thermal expansion data over a comparable temperature range when measured at different temperatures [377]. A significant uncertainty is always expected in length measurement of mechanically soft metal, i.e. potassium [376]. Phase transition trigged by temperature changing will cause difficulty in preserving the ultrasonic coupling between specimen and transducer [412]. Thus, coupling between specimen and transducer need to be tight enough to hold them together without cause structure change in specimen.

4.4 Conclusion

A systematic and potentially automatic method has been applied in this work to identify outliers and correct errors that exist in handbooks and databases. This is demonstrated by employing ANNs to explore the ternary order correlations of the elastic properties of 68 metallic elements, and by utilizing such correlations to identify and correct the suspect data. The ANN models are built on the premise that correct values are provided by most sources for the most commonly known pure metals. Such presupposition is proved to be essentially correct.

The elastic property data that were recorded in prestigious handbooks and databases for elements in the Periodic Table were thought to be very reliable, and errors were likely to be lower than 5%. However, a carefully conducted inspection for 68 metallic elements in the Periodic Table shows that large discrepancies exist in the 12 sources. Only five metallic elements (Co, Dy, Fe, Ta and Tb) have the variances of the four elastic properties smaller than 10%, apart from 6 elements (Ac, Fr, Np, Pa, Ra and Pm) that have very limited data available in the source pool.

Totally, 11 ANN models are successfully constructed using the consistent values for 5 elements (Co, Dy, Fe, Np, Ta and Tb) plus the most common values for 22 elements (Ag, Al, Au, Ba, Bi, Ca, Cr, Cu, Mg, Mn, Mo, Nb, Ni, Pb, Pd, Pt, Sn, Sr, Ti, V, W and Zn). The correlations that are captured from the 27 training examples are then used to evaluate the inconsistent data for the rest 35 elements. In general, the obtained ANN models agree well with the correlating equations (established from the physical models)

in the data ranges that are determined by the maximum and minimum values of the 27 training examples. For an enlarged data range, the similarity between the ANN models and the correlating equations declines.

For ternary order correlations, one missing property data can be predicted from several relationships. When elastic properties predicted by utilizing different ANN models do not yield results within 10% variation, a systematic methodology is proposed to resolve the possible discrepancy. When suspect data lie in the ANN training domain, the data verified by ANNs are very close to the data verified by the correlating equations. For suspect data that exceed the training domain, the accuracy of ANN estimation can be improved by employing multiple mutually constrained correlations.

A detailed literature review providing a completed set of available experimental values has also been conducted to evaluate the elastic properties of 35 elements, including elastic constants for single crystals and the corresponding elastic properties obtained through the VRH averaging method. Factors that would affect the experimental values have been discussed. Meanwhile, reasons for data variances have been analysed case by case for each element. Most data discrepancies attribute to the different experimental conditions, i.e. experimental temperature, specimen structure, purity, and processing history, and measurement methods. Incorrect unit conversions or publication typos also result some errors. Finally, a considerable effort has gone into making sure elements' elastic data provided in this chapter much reliable, and thus enables readers to use them with more confidence.

5. Capturing materials properties correlations using artificial neural networks: an example in hardening of pure metals by high pressure torsion

5.1 Introduction

There has been considerable interest in material processing by high pressure torsion, a metal forming method that produces improvements in properties, especially in strength, by imposing a very high strain on a bulk solid [191]. Extensive work in the field by a range of researchers indicates that the hardness saturates to steady-state levels at high strains and remains unchanged with further straining [184-187, 189]. The saturated values correlated to several mechanical properties and atomic bond parameters. However, little attention is given to establishing whether a limited set of physical properties can explain all measured changes in hardness due to HPT.

In the following section, ANNs are employed to capture the statistical relations between hardness increment of pure HPT processed metals and 13 physical, mechanical, and electronic properties. To propose the most promising variables for the construction of a physically based model, the ANN combinatorial search method and the forward selection method are both applied to identify the properties that constitute the strongest correlations. Because the available data on the hardness of HPT-processed pure metals are of limited supply (only 17 pure metal data), the question that whether fruitful ANN models can be obtained when dealing with a small dataset is addressed.

5.2 Methodology

5.2.1 Data collection

Among the 70 metals in the periodical table, roughly 30 pure metals and semi-metals with various crystal structures have been processed by HPT, and their Vickers hardness at the saturated level ($\boldsymbol{H_{Vs}}$) have been reported [191, 414-417]. Thus, the absolute harness increments ($\Delta\boldsymbol{H_{V}}$) due to HPT can be calculated by subtracting the hardness at

annealed state (H_{Vi}) from the hardness at the saturated level (H_{Vs}), i.e. $\Delta H_V = H_{Vs} - H_{Vi}$. However, no systematic measurement of H_{Vi} has been conducted for all the metals. Through the exhaustive searches in various databases including various journals, handbooks and online databases such as **Reaxys** and **Landolt-Börnstein**, the H_{Vi} data were obtained in different experimental conditions, for example, 'as-received' or 'annealed'. Consequently, the smallest reported value for each element was adopted, as it is believed to better reflect the initial pure un-deformed condition (see Table 5-1). The influences of H_{Vi} data on model accuracy are discussed in Section 5.3.3.

Table 5-1 Vickers hardness data of 17 elements before and after HPT (H_{Vs} and H_{Vi}) and the absolute increment of hardness (ΔH_V).

Elements	Structure	H_{Vs} /GPa	H_{Vi} /GPa	ΔH_V /GPa	Reference for H_{Vi}
Mg	hcp	0.342	0.285	0.057	[414]
Al	fcc	0.313	0.167	0.146	[279]
Ti	hcp	2.599	0.971	1.628	[279]
V	bcc	2.354	0.628	1.726	[279]
Cr	bcc	4.756	1.060	3.696	[279]
Fe	bcc	3.020	0.608	2.412	[279]
Co	hcp	3.544	1.043	2.501	[279]
Ni	fcc	3.021	0.638	2.383	[279]
Cu	fcc	1.298	0.369	0.929	[279]
Zn	hcp	0.362	0.353	0.009	[186]
Zr	hcp	2.532	0.903	1.629	[279]
Nb	bcc	2.354	0.354	2.000	[418]
Pd	fcc	2.127	0.461	1.666	[279]
Ag	fcc	0.941	0.251	0.690	[279]
Ta	bcc	4.132	0.873	3.259	[279]
Pt	fcc	2.525	0.549	1.976	[279]
Au	fcc	0.804	0.216	0.588	[279]

To allow a comparison with the physical model [116], the same 17 pure metals (Mg, Al, Ti, V, Cr, Fe, Co, Ni, Cu, Zn, Zr, Nb, Pd, Ag, Ta, Pt and Au), which are considered the most reliable, are used to conduct the ANN analysis. The input arguments are selected initially from 12 properties given in CES [228] together with Burgers vector data (***b***) from Edalati and Cullity [185, 419]. The 13 properties plus the target property ΔH_V are listed in Table 5-2, and the entire dataset is listed in Appendix XII. Though some of these 12 properties may have strong internal correlations, ANNs are capable of highlighting the most important parameters.

Table 5-2 The 14 properties used in ANNs: ΔH_V is the target output, while the input variables are chosen from the rest 13 properties.

Property symbol	Property	Unit
ΔH_V	Hardness increment after HPT	GPa
b	Burgers vector	nm
A_n	Atomic number	
BE/A	Binding energy per nucleon	keV
E_{coh}	Cohesive energy	kJ/mol
ρ	Density at 300 K	kg/m^3
H_{fus}	Heat of fusion	kJ/mol
a	Lattice parameter, a	nm
T_m	Melting temperature	K
V_m	Molar volume	m^3/kmol
G	Shear modulus at 300 K	GPa
C_p	Specific heat capacity	J/kg°C
α_L	Thermal expansion coefficient at 300 K	μstrain/°C
W	Work function	eV

5.2.2 The inputs and output

Artificial neural networks are a biologically motivated computing paradigm that is capable of mapping a set of input data values to the associated output data within a desired accuracy [4, 5, 78]. The target output of the ANN and the physical model is the

absolute increment of hardness after HPT ($\Delta \boldsymbol{H_V}$). The performance of ANNs depends on the various combinations of input variables, which indicates if the correlation exists or not.

The combinatorial correlations search started from the prediction of $\Delta \boldsymbol{H_V}$ using one input parameter (1st order search). The results of 1st order search is the binary order correlations between $\Delta \boldsymbol{H_V}$ and the other property chosen from the 13 properties prepared for the 17 metals. Subsequently, the combinations of any two parameters selected from the 13 properties were used as input variables for the prediction of 2nd order search. Similarly, searches were performed up to the 4th order. The results of 2nd order search and 3rd order search are ternary order correlations and quaternary order correlations, respectively. The results will show the 4th order search is not an appropriate choice.

In the forward selection, the key variable that have the potential to deliver satisfactory explanations of phenomena are initially determined from the 1st order search. In the next step (2nd order search), such variable is fixed as one of the input properties, and the second input properties is chosen from rest 12 properties. Therefore, 12 ANNs representing 12 ternary order correlations are constructed. After analysing the 12 ANN models, the two key variables that have the most predictive power are fixed as the input properties, a third property is selected from the rest 11 properties to conduct a 3rd order search for quaternary correlations.

5.2.3 Neural network analysis method

Figure 5-1 shows a 2nd order ANN predicting $\Delta \boldsymbol{H_V}$ from shear modulus ($\boldsymbol{G}$) and melting temperature ($\boldsymbol{T_m}$). For an accurate prediction, the solid red and dot blue line should coincide, and both the slope $\boldsymbol{S_l}$ and correlation coefficient $\boldsymbol{R_l}$ should approach 1. Element symbol is used to label the data point if the prediction deviation is greater than 10%. For Ag, Ti and Zr, $\Delta \boldsymbol{H_V}$ is underestimated by ~0.2 GPa, while for Zn, Mg and Al, ANNs show a slight overestimation of ~0.05 GPa, comparing to the experimental values. For the other studied metals, the ANN model give quite accurate results.

To numerically evaluate the performance of the ANNs, two methods can be used. The first one is $\boldsymbol{\delta}$, which is a comprehensive evaluation of $\boldsymbol{S_l}$ and $\boldsymbol{R_l}$ for both the training data and test data [10]. The smaller $\boldsymbol{\delta}$, the better ANNs perform, and the stronger the correlations. Typically, ANNs start to reveal important correlations when $\boldsymbol{\delta}$ is smaller

than 0.1. Another way is to simply compare the difference between the experimental data and the prediction data by scrutinizing both individual difference and the average difference. When the difference of $\boldsymbol{\delta}$ is small between two similar ANNs, the average error percentage is also used to justify the performance of the ANNs.

In assessing model performance, it is important to note that it has been shown that on average 5% or more of the data in handbooks and databases are incorrect or to a significant extent inaccurate. Though they may not have a big impact on the performance of ANNs [7], a small percentage of incorrect data could change the average error percentage dramatically. As suggested by Chapman [234] and Chrisman [235], it is anticipated that about 5% of the data are incorrect in the prepared data, and hence 5% of data with the biggest error were not included during the calculation of average error percentage to provide a "fair" comparison. This average error percentage is considered as the second criterion. All ANN results are ranked based on the two criteria. The top properties correlations will come forth as the priority attributes that can be used to reveal the underlying physical principles.

5.2.4 Knowledge extraction

The major challenge in extracting knowledge (rather than just correlations) from the ANN modelling results is to find a way to determine which input variables significantly affect the output variable, finally leading to the construction of analytical theory. When the number of the input variables is small, it is instructive to conduct the exhaustive search to identify the optimal set of input variables for a given ANN model [108].

However, when the search space of candidate sets is highly dimensional (system has high complexity), and exhibits a computational time dependence which scales as $\boldsymbol{N^d}$, where $\boldsymbol{N}$ is some measure of the size of the problem, $\boldsymbol{d}$ is the search space dimension, evaluating all the possible combinations of input variables to select the best set according to a predetermined optimality criteria may not always be possible [420]. A forward selection strategy can be applied as a solution.

The forward selection method is a linear incremental search strategy to efficiently locate the key parameters that have the potential to deliver satisfactory explanations of phenomena [108]. It selects candidate variables one at a time and is terminated either

when the result has yielded to a desired level of accuracy or when it has reached a pre-set number of parameter values that one is interested in exploring in the blind search.

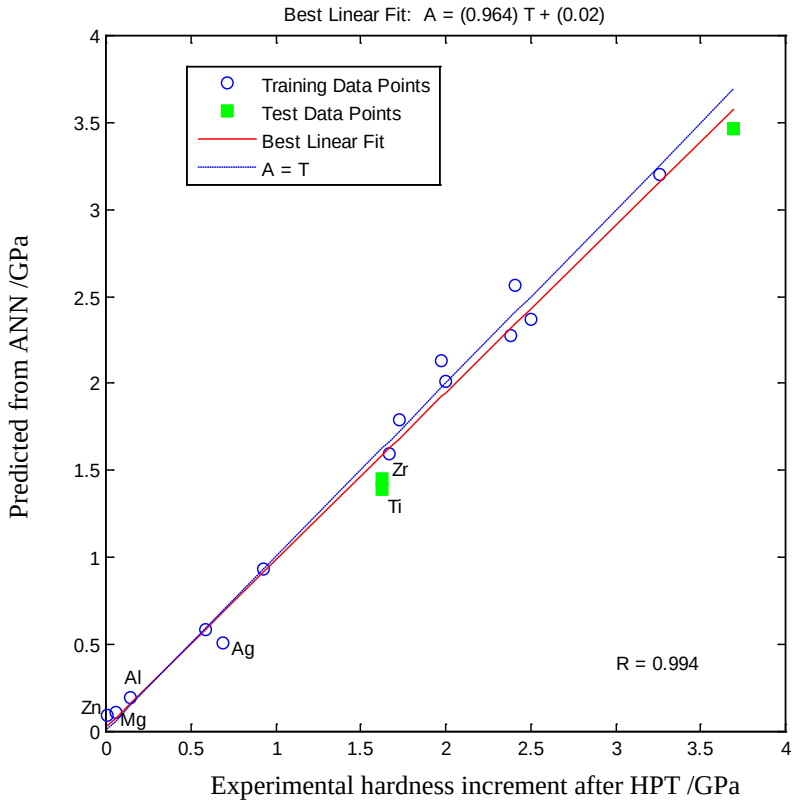


Figure 5-1 The result of ANNs in predicting ΔH_V from G and T_m, where δ equals to 0.04 and the correlation error equals to 12.67%.

5.3 Results and discussion

5.3.1 Comparison of ANN curves

The influence of each input variable on the output data is determined by both linear regression and 1st order search ANNs using only one input variable to predict the ΔH_V. Totally, 13 ANNs have been searched during which the largest correlation coefficient (R_l) 0.88 obtained from the prediction of ΔH_V from G, followed by T_m, H_{fus} and α_L with R_l equalling to 0.85, 0.82 and 0.81, respectively. For all the other nine properties, R_l is much smaller than 0.80. Linear regressions between a possible physical property and the hardness increment exhibit similar trends. Although R_l is similar for the ANN results

predicting from G or T_m, the first evaluation criteria δ indicates G can be considered to have the strongest impact on ΔH_V (see Table 5-3).

Table 5-3 Comparison of 1st order ANN models and linear regression results using 13 properties. R_l is the relative coefficient of the experimental result and the ANN prediction using each property as the input, δ is the criteria to evaluate the ANN performance. R_2 is the relative coefficient between the input property and the target output ΔH_V from the linear regression analysis.

Property symbol	ANNs evaluation parameters			Linear regression
	R_l	δ	E_c	R_2
b	0.46	1.31	179.0%	0.15
A_n	0.17	1.83	290.3%	0.17
BE/A	0.43	1.42	189.3%	0.18
E_{coh}	0.67	0.91	87.5%	0.66
ρ	0.28	1.54	189.9%	0.25
H_{fus}	0.82	0.44	44.5%	0.72
a	0.46	1.27	284.7%	0.36
T_m	0.85	0.41	60.6%	0.81
V_m	0.41	1.37	99.9%	0.40
G	0.88	0.32	54.1%	0.87
C_p	0.46	1.44	219.1%	0.38
α_L	0.81	0.55	48.6%	0.8
W	0.36	1.47	136.1%	0.26

In the next step (2nd order search), G is fixed as one of the input properties. For the 2nd order ANNs, there are 12 properties left to be chosen as the second input properties. The results are shown in Table 5-4. ANNs provide the best predictions by using G and T_m, G and α_L, and G and H_{fus} with correlation coefficient R_l all larger than 0.98 and δ all smaller than 0.1. However, the average error percentage is larger than 10% in all ternary order correlations. Though G and T_m make the optimal input variable set for the 2nd

order ANN modelling, it should be noticed that the predictions, $\boldsymbol{G}$ and $\boldsymbol{\alpha_L}$ and from $\boldsymbol{G}$ and $\boldsymbol{H_{fus}}$ give very similar results. Given the fact that $\boldsymbol{T_m}$ and $\boldsymbol{H_{fus}}$ as well as $\boldsymbol{T_m}$ and $\boldsymbol{\alpha_L}$ are correlated to some extent [421, 422], these similar results predicted by ANNs may simply be related to these correlations.

In the 3rd order ANNs search, $\boldsymbol{G}$ is fixed as first input properties, and $\boldsymbol{T_m}$, $\boldsymbol{H_{fus}}$ or $\boldsymbol{\alpha_L}$ is fixed as the second input property, and 11 properties are left to be chosen as the third input property. Table 5-5 tabulates the most accurate ANNs for predictions of the quaternary order correlations. $\Delta\boldsymbol{H_V}$ predicted by using $\boldsymbol{b}$, $\boldsymbol{G}$ and $\boldsymbol{T_m}$ is found to be a very good match to the experimental result as shown in Figure 5-2. In fact, $\boldsymbol{b}$, $\boldsymbol{G}$ and $\boldsymbol{T_m}$ are the key parameters in the suggested physical model [116].

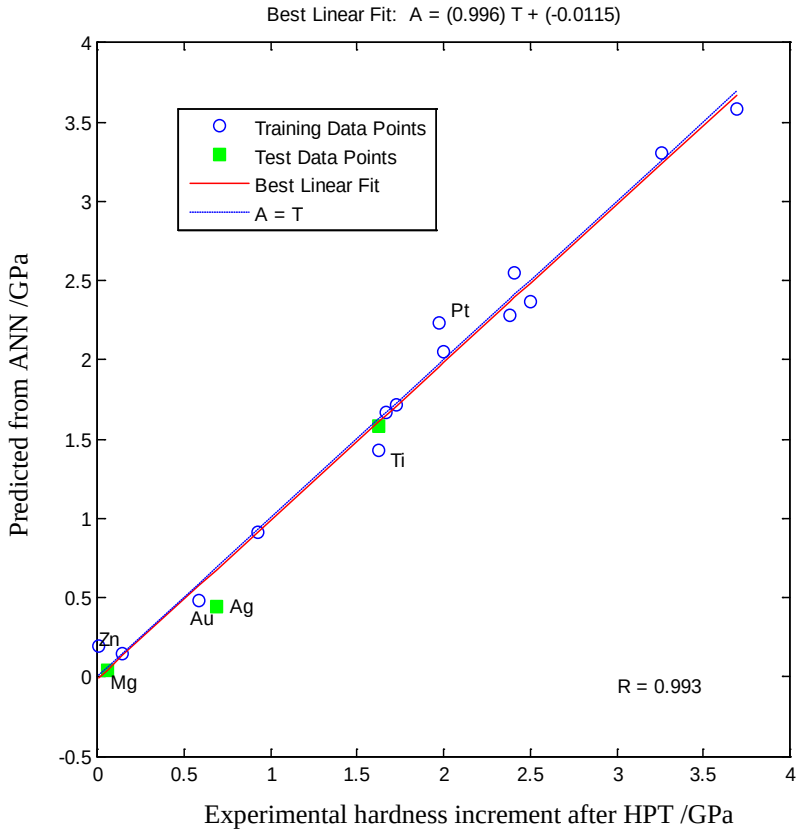


Figure 5-2 Result of ANNs in predicting $\Delta\boldsymbol{H_V}$ from $\boldsymbol{G}$, $\boldsymbol{T_m}$ and $\boldsymbol{b}$ with $\boldsymbol{\delta}$ equals to 0.01 and 8.05% error.

Table 5-4 The 12 ANNs correlations to predict ΔH_V using 2 properties in the forward selection results.

Importance	ANNs evaluation parameters			Input properties
	ψ	δ	E_c	
1	0.13	0.04	12.67%	T_m, G
2	0.19	0.03	19.13%	α_L, G
3	0.27	0.02	27.03%	H_{fus}, G
4	0.36	0.06	35.85%	E_{coh}, G
5	0.42	0.15	38.72%	V_m, G
6	0.43	0.33	27.04%	C_p, G
7	0.56	0.31	46.78%	BE/A, G
8	0.67	0.35	57.68%	a, G
9	0.71	0.35	61.58%	A_n, G
10	0.75	0.28	69.70%	W, G
11	0.80	0.22	77.41%	b, G
12	1.06	0.47	95.13%	ρ, G

The 4th order ANNs search has also been conducted. Not surprisingly, the results from top-ranked combinations are almost same as the 2nd and the 3rd order searches, consisting of G plus T_m, H_{fus} or α_L within R_l approaching 0.98, δ smaller than 0.1 and error ranging from 5% to 15%. This may be due to the relatively small available data set combined with large degrees of freedom introduced by the 4th order search, therefore loosing generalization [423, 424]. As mentioned before, it is generally recommended having more than 10 examples per input variable, though the size of data required in practice depends on the complexity of the problem and amount of noise in the data [190]. Because the number of records available for training and testing was relatively small, only 17 in total, ANNs are more reliable when constructed with a fewer degrees of freedom. Thus, the search space of 4th order would be too large for 17 data to generate appropriate ANNs. Another possible reason is that, as shown in Table 5-4 and Table 5-5, ΔH_V can already be predicted quite well by using only two parameters, G and T_m, as b only plays a minor role in the optimal function.

Table 5-5 Top 12 ANNs correlations to predict ΔH_V using 3 properties in the forward selection results.

Importance	ANNs evaluation parameters			Input properties
	ψ	δ	E_c	
1	0.08	0.01	8.05%	b, T_m, G
2	0.09	0.02	8.59%	V_m, T_m, G
3	0.11	0.02	10.56%	α_L, T_m, G
4	0.13	0.04	12.61%	a, T_m, G
5	0.13	0.03	13.00%	E_{coh}, T_m, G
6	0.14	0.03	13.26%	A_n, T_m, G
7	0.15	0.04	14.52%	W, T_m, G
8	0.16	0.03	15.46%	ρ, T_m, G
9	0.16	0.03	15.97%	BE/A, α_L, G,
10	0.17	0.06	15.94%	H_{fus}, T_m, G
11	0.17	0.02	16.89%	α_L, H_{fus}, G
12	0.21	0.03	20.51%	BE/A, T_m, G

Whilst strengthening theories generally confirm that G is a key parameter, it could be argued that in a pure adaptive modelling sense the choice of G to be fixed as input variables for the 2^{nd} ~ 4^{th} order ANNs search. However, other parameters may be confounded and their effects on the prediction could be larger than the pre-set input variable G. In order to avoid any possible missing parameters that might be important for interpretation in the physical theory and to prove the forward selection method is effective in this study, an exhaustive search of input variables of the 2^{nd} and 3^{rd} order was employed. The top 12 sets according to the same criteria described before are selected, in contrast to the 12 input candidates that can be chose to construct ANNs in the forward selection method.

Recognizing that when the differences of δ and the average error percentage between the ANNs are small in the top ranked ANNs, the absolute ranking of the ANNs may not reflect the genuine physical model due to noise and possible errors from the data, hence a statistic of frequency of the input variables in the top ANNs are used. ANN programs using all possible input variable combinations were analysed to determine how many

times each input appears in the 12 top models that yield the optimal generalisation performance of the trained ANNs. The results, shown in Figure 5-3, indicate that in both the 2nd and the 3rd order of searches, input variable ***G*** always has the highest frequency, followed by T_m, then α_L and H_{fus}. It indicates that input parameters of ANNs should include ***G*** as the fundamental parameter plus either T_m, α_L or H_{fus}. Apart from these three properties, all the others appear at very low frequency more or less randomly. In spite of strong correlations in the input parameters, such as between T_m and H_{fus}, and between T_m and α_L, ANNs adequately identify the most salient input variable being T_m.

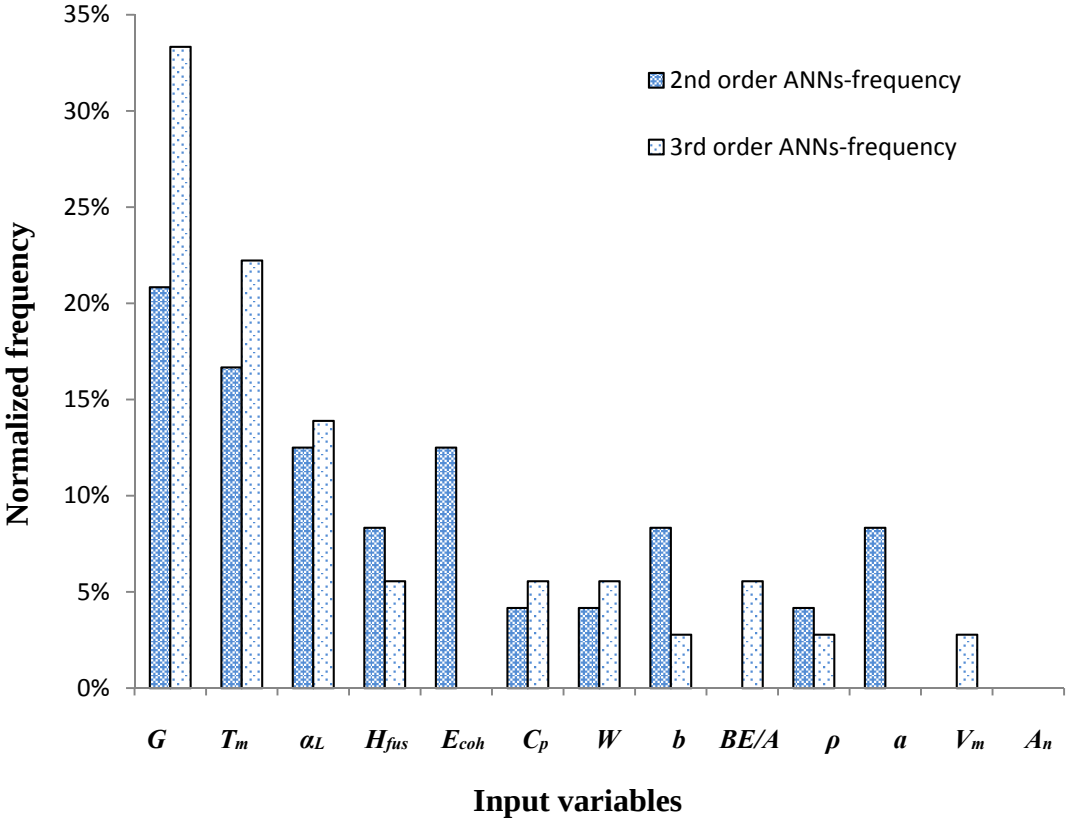


Figure 5-3 The property importance revealed in the top 12 correlations of 2nd and 3rd order in predicting ΔH_V from the ANN combinatorial search.

5.3.2 Modelling with a limited data supply

Most studies and comparisons of the effectiveness of ANN techniques attempted to use a large database, which has hundreds or even thousands of data points. Experiments using a database of 10-20 data points are not rare, but raise controversy on the models' validity for using inadequate quantities of data to create sophisticated models [5, 425-

428]. However, the ANN models presented in this study should not be perceived as impaired, owing to only 17 input-output data were used for establishing the networks.

One argument to justify the validity of a model is that, there should be more data than fitting points [428]. For a neural network with one hidden layer, its structure can be denoted as $\boldsymbol{N_{inp}}$—$\boldsymbol{N_{hid}}$—$\boldsymbol{N_{out}}$, where $\boldsymbol{N_{inp}}$ represents the number of input variables, $\boldsymbol{N_{hid}}$ represents the number of neurons in the hidden layer, and $\boldsymbol{N_{out}}$ represents the number of output parameters. The total number of weights to be determined, i.e. $(\boldsymbol{N_{inp}}+1)\times \boldsymbol{N_{hid}}+(\boldsymbol{N_{hid}}+\boldsymbol{1})\times \boldsymbol{N_{out}}$, should be smaller than the number of data pairs available for training.

As described in Chapter 2, at a ratio of 4:1, 14 data pairs are randomly selected to train ANNs, and 3 data pairs are used to test the generation ability of the trained ANNs. The number of hidden neurons is strictly constrained to the maximum connections between neurons in the input, hidden, and output layers. Because the only output variable is the hardness increment due to HPT, there is only one output neuron in an ANN model. The maximum number of hidden neurons deceases with the increase of input neurons. For the 1st order search targeting at the binary order correlations, there is one input neuron. Therefore, the number of hidden neurons is chosen from one to four. Similarly, for the 2nd order ANN searcher targeting at the ternary order correlations, the number of hidden neurons is chosen from one to three, and for the 3rd order ANN searcher targeting at the quaternary order correlations, the number of hidden neurons is chosen from one to two. Therefore, the amount of data used in the present work is enough to determine the number of fitting parameters in the networks. Moreover, the constructed ANN models achieve good performances with three testing data.

The second argument deals with the uncertainty of modelling [5, 78]. In light of the limited amount of data available, 30 ANN models are created based on the methodology described in the Chapter 2. These models are not identical, but they all reasonably explain the experimental data. The reproduction accuracy for trained ANNs is shown in Figure 5-4 as a quantitative measure of the modelling uncertainty. It displays the minimum, average and maximum value of $\Delta \boldsymbol{H_V}$ predicted from $\boldsymbol{G}$ and $\boldsymbol{T_m}$ obtained from the 30 independent ANNs. The x-axis corresponds to the experimental value of $\Delta \boldsymbol{H_V}$, the y-axis corresponds to the predicted value from the ANNs. All the elements were found to be located within the reasonable uncertainty range, meaning the ANNs capture

the correlation that exists between ΔH_V, G and T_m while retaining the modelling accuracy. The uncertainty could be the result of noise in the data or lacking of other properties that have small impacts on ΔH_V, *i.e.* b.

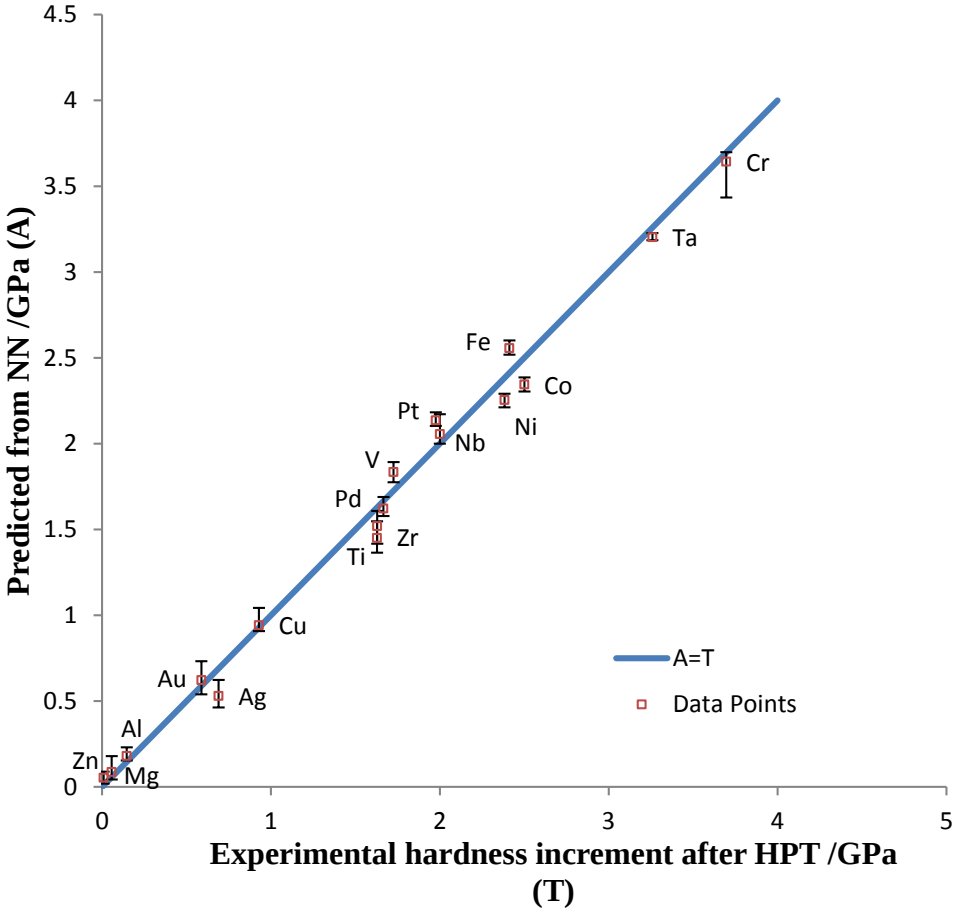


Figure 5-4 Modelling uncertainty: predicting ΔH_V from G and T_m from 30 independent ANNs.

5.3.3 Factors that affect the accuracy in the prediction

Apart from the ANN modelling error tolerance, the prediction accuracy is mostly influenced by the accuracy of data used to build the model, which are G, T_m and ΔH_V. The accuracy of T_m and G has been discussed in Ref. [116], and a more detailed discussion on G can be found in Chapter 4 and Appendix XIII. In brief, the melting temperature of all metals shows excellent accuracies. T_m of 10 metals (Al, Co, Ni, Cu, Zn, Nb, Pd, Ag, Pt and Au) is the defining fixed temperature of the International Temperature Scale of 1990 (ITS-90). The different sources of data on T_m of the other 8

elements (Mg, Ti, V, Cr, Fe, Zr, Hf and Ta) are in good agreement with variations typically smaller than 2%. Data on $\boldsymbol{G}$ of pure metals are not very consistent, and the largest spread in reported values of $\boldsymbol{G}$ for pure metals occurs for Pd (a variation of ~20%). However, the model accuracy is not significantly influenced if using different reported $\boldsymbol{G}$ values. As the determination of corresponding $\boldsymbol{Hv_i}$ for each element has been explained previously, the accuracy of $\boldsymbol{H_{Vs}}$ is discussed in detail here.

The purity, grain size, dislocation density, texture, thermo-mechanical history, recrystallization effects of samples, as well as deformation technique, testing atmosphere, and other experiment conditions are all important factors that may cause variation of mechanical prosperities in nominally "pure" metals. Impurities can adversely affect the grain size and fraction of high-angle boundaries [429]. Gludovatz *et al.* [430] observed that both grain size and microstructure have a strong impact on the fracture behaviour for materials with a low impurity concentration. Wei *et al.* [417] also noticed a variation of hardness on top and bottom faces of the HPT disk. Recent study reports that hardness after HPT of some metals begins to decrease during storage at room temperature, such as Ag and Au [431]. Errors could be introduced if the hardness measurement has not been performed in time.

In Table 5-6, $\boldsymbol{H_{Vs}}$ of Al, Ti, Fe, Ni, Cu, Zr, Ta and Nb, the 8 metals processed under the similar HPT experimental condition that can be found out of the 17 metals in other references [432-440], were compared with the results reported by Edalati *et al.* [183-185, 187]. The absolute discrepancies of $\boldsymbol{H_{Vs}}$ data are mainly within ±0.15 GPa if disc samples were processed under pressure 2~6 GPa for 4~10 revolutions at room temperature avoiding phase change [184], with variation from a few percentage up to near 30%. The influence of $\boldsymbol{H_{Vs}}$ on $\Delta\boldsymbol{H_V}$ was also been reported in the last column of Table 5-6.

*Table 5-6 Hardness at steady state after HPT in different sources, * indicated the value used for each element.*

Metal	$\boldsymbol{H_{Vs}}$ **/GPa**	**Reference for** $\boldsymbol{H_{Vs}}$	$\boldsymbol{H_{Vs}}$ **Difference/GPa**	$\boldsymbol{H_{Vs}}$ **Variation**	$\Delta\boldsymbol{H_V}$*** /GPa**	$\Delta\boldsymbol{H_V}$ **variation**
Al	0.313* 0.403	[183-185, 187] [437]	+0.090	28.75%	0.15	61.64%

Metal	H_{Vs} /GPa	Reference for H_{Vs}	H_{Vs} Difference/GPa	H_{Vs} Variation	ΔH_V* /GPa	ΔH_V variation
Ti	2.599*	[183-185, 187]	+0.098	3.77%	1.63	6.02%
	2.697	[433]				
Fe	3.020*	[183-185, 187]	+0.580	19.21%	2.41	24.05%
	3.600	[440]				
Ni	3.021*	[183-185, 187]	-0.168	5.56%	2.38	7.05%
	2.853	[441]				
Cu	1.298*	[183-185, 187]	+0.049	3.78%	0.93	5.27%
	1.347	[182]				
Zr	2.532*	[183-185, 187]	-0.325	12.84%	1.63	19.95%
	2.207	[438]				
Ta	4.132*	[183-185, 187]	-0.132	3.19%	3.26	4.05%
	4.000	[435]				
Nb	2.354*	[183-185, 187]	+0.176	7.48%	0.35	8.80%
	2.530	[434]				

The largest variations of H_{Vs} of Al, Fe and Zr cause more than 10% variations on their calculated ΔH_V. Harai *et al.*[442] observed a maximum hardness about 0.4 GPa in Al at an equivalent strain of ~2, similar to the result obtained by Xu *et al.* [437], which then decreased to about 0.3 GPa when equivalent strain reach ~6, and remained unchanged with respect to imposed strain, due to the mutual interaction between dislocation density and grain boundaries. For Fe, the difference in H_{Vs} is possibly due to the different carbon contents in the experimental materials [443], while for Zr, the discrepancy of H_{Vs} is possibly due to the impurities in the experimental materials and the different test pressures [438]. Thus, the overestimates of Al and Fe and underestimate of Zr in ANNs as shown in Figure 5-4 may actually reflect the true experimental values.

5.3.4 Underlying physical principle of parameters extracted by ANNs

The results shown in Figure 5-2 indicate the confounding effect of shear modulus, melting point and Burgers vector on the increment of hardness. Comparing with Figure 5-1, the Burgers vector plays a minor role in the ANN models. It is desirable to explore the physical principles behind the information revealed by the ANNs, in order to verify the feasibility of adopting this method to explore other unknown cross-property relationships in the future.

Edalati *et al.* [187] pointed out that $\boldsymbol{G}$ and $\boldsymbol{T_m}$ are important parameters to scale $\boldsymbol{H_{Vs}}$ attained by HPT because "$\boldsymbol{G}$ is a parameter to explain dislocation interactions, deformation processes and hardening rate" and "$\boldsymbol{T_m}$ is a parameter to explain the activation energy for diffusion, recovery processes and softening rate". Since $\boldsymbol{H_{Vs}}$ is a characteristic property of each metal, and so does $\boldsymbol{H_{Vi}}$, it is reasonable to believe the increment of hardness after HPT, $\Delta\boldsymbol{H_V}$, can also be represented by $\boldsymbol{G}$ and $\boldsymbol{T_m}$, or better expressed as homologous temperature to correlate temperature dependence of deformation for different metals [186]. Because the HPT processing used here occurs at room temperature, homologous temperature varies inversely with $\boldsymbol{T_m}$. Burgers vector is well known in describing dislocation of crystal lattice and dislocations at grain boundaries [444], which is connected to the main aspect of hardening and recovery during HPT process. Because the magnitudes of Burgers vector of 17 metals are very close, the impact of $\boldsymbol{b}$ on the prediction of $\Delta\boldsymbol{H_V}$ is expected to be quite small, which is consistent with the ANN results.

Atomic bond energy and activation energy for self-diffusion are the two properties that recently have been reported to show a correlation with $\boldsymbol{H_{Vs}}$ of HPT processed metal [185]. The ANN models utilizing the two properties have also been conducted. However, they show no significant impacts in predicting $\Delta\boldsymbol{H_V}$. Because the data of the two properties in CES for the 17 metals reveal a moderate variance from the data listed in the Ref. [185], they are not included in the systematic ANNs search.

5.3.5 Comparison with the physical model

In parallel with the present study, a physically-based model (using the same 17 metals) incorporating volume-averaged thermally activated dislocation annihilation and grain boundary formation to predict the increment of hardness and grain refinement of pure

metals due to HPT, is proposed in Ref. [116]. The ANN results presented in this study is beneficial to the construction of the physical model for illuminating the most important parameters (i.e. $\boldsymbol{b}$, $\boldsymbol{T_m}$ and $\boldsymbol{G}$). From the point view of the physically-based model [116], the hardness increment due to HPT is the result of dislocation hardening and grain boundary strengthening, which can be mainly determined by shear modulus and Burgers vector. Melting temperature is linearly correlated to the activation energy for a relaxation process. The physical model exhibits the almost same prediction accuracy as the ANN model constructed with the three parameters. The RMSE for $\Delta\boldsymbol{H_V}$ for 17 elements of the physically based model is 0.14 GPa (9% of the average $\Delta\boldsymbol{H_V}$) and the RMSE of the ANN is 0.13 GPa.

It should be emphasised that: 1) Due to the small input data set (17 metals), it is impossible to establish reliable ANNs with the same variables as used in the physical model. ANNs constructed with four variables are suffering from over-fitting, producing bad generalization to new data. 2) The physical model is a simplified model that ignores some secondary effects. For example, the treatment of thermally activated dislocation annihilation ignores multiple interacting thermally activated processes impact on the rate of dislocation annihilation. 3) Instead of dynamic recovery employed in the physically based model, dynamic recrystallization may be a dominant mechanism in strengthening HPT processed metals [185]. Therefore, the physical model utilizing five variables is not the only model that can be used to explain the phenomena of hardness increment. 4) Strong correlations with the melting temperature, such as heat of fusion, cohesive energy and thermal expansion coefficient, are the reasons why they show a correlation with $\boldsymbol{\Delta H_V}$ or $\boldsymbol{\Delta H_V/G}$ when plotted, however there is no suggestion of a causal relation. None of the other physical, chemical, and mechanical parameters assessed by ANN has a statistically significant influence on hardening due to HPT separate from the parameters identified. 5) Nevertheless, the ANNs successfully reveal that shear modulus and melting point are the significant properties that influence the hardening of pure metals most, and Burgers vector has a relatively small impact. The correlations captured by the ANNs shed a light on the construction of a physically based model by identifying the salient input parameters.

5.3.6 The applicability of the forward selection method

The forward selection method is expected to find the most important input parameters that have maximum influence on the predicted results, and is particularly effective when the database is relatively small but the search space of inputs is quite large. For a database that is large enough compared to the relatively small search space of inputs, variable contributions can be analysed by utilizing the weight vectors of ANNs (see the discussions in Section 6.2.1 and Section 6.3.2).

In a general situation, for n^{th} order correlations with m properties already determined as the significant parameters in previous $(n-1)^{th}$ order correlations, by analysing the ANN results with linearly increased inputs, a minimal acceptable model that can characterise the system with optimal prediction ability is identified. In this chapter, ANNs are generated for correlation modelling up to 4^{th} order, and G, T_m and b are identified as the most important physical properties to explain the hardness increase mechanism, within the average error less than 10% and the relative coefficient (between the experimental data and the ANNs predictions) larger than 0.99.

5.4 Conclusion

The main finding of this chapter is that an ANN model can be devised to explore an underlying physical principle with minimum or no prior knowledge. As an example, the increment of Vickers hardness of pure metals due to HPT is accurately predicted from a limited set of properties. The limited set of properties, which are determined by ANNs from 13 physical properties that may explain the measured changes, constitutes very strong correlations with hardness increment.

By conducting the ANNs input variables forward selection and the ANN combinatorial search, shear modulus and melting temperature are identified to have the largest impacts on hardening. They can be used to quantitatively predict the ΔH_V due to HPT at an accuracy of 87%. Adding Burgers vector as the third input variable, the new ANN model produces the optimal performance. The quaternary correlation between b, T_m, G and ΔH_V outperform any other ANN models (i.e. correlations of the same or lower order) that an accuracy of 92% is finally reached. The problems related to the accuracy

of input data mainly depended on $\boldsymbol{H_{Vs}}$. However, using different $\boldsymbol{H_{Vs}}$ data does not heavily impair the predict power of the ANNs.

Nevertheless, even using a small database of 17 data points, the ANNs presented in this study should be perceived as valid models. Because the number of hidden neurons is strictly restricted to a maximum value, the amount of data used in the experiment is large enough to determine the number of fitting parameters in the networks. Moreover, to ensure the generalization ability of ANN models, testing data are always used. Meanwhile, the modelling uncertainty of the best binary order correlation has also been analysed.

A proposed physically based model corroborates the dominant correlations revealed by the ANNs. The correlations found by ANNs shed a light on the construction of physically based model by pointing out the input parameters of greatest importance. It suggests that, even with a limited supply of data, ANNs can be applied to explore property correlations in materials science where a physically based model is not readily available.

6. The discovery of materials properties correlations through artificial neural networks and symbolic regression

6.1 Introduction

Data mapping is not a new concept. Various data-driven techniques are available to create models, including ANNs and genetic programming. However, research in automatically generating the analytic forms of correlations (mathematical equations) is rare in materials science. In this chapter, a combined ANN-SR (symbolic regression) method is employed to determine the enthalpies of vaporization of 175 organic and inorganic compounds. ANNs are applied to capture the correlations hidden in the data and to identify the important input variables. SR is then employed to derive the correlating equations.

Though ANNs have long been regarded as 'black-boxes' that are encoded with weight vectors, a number of methods studying variable contributions in ANNs have been developed in recent years. The 'PaD' method [220] and the 'CW' approach [221] are the two methods providing explanatory insight of the 'black-boxes' and receive the greatest attention in many disciplines. A comparison study of the two methods is performed to evaluate the contributions of the five input variables, namely, $\boldsymbol{M_w}$, $\boldsymbol{T_b}$, $\boldsymbol{T_c}$, $\boldsymbol{P_c}$ and $\boldsymbol{D_m}$. Meanwhile, the contribution of each input variable is also accessed through the ANN combinatorial search and the SR approach.

6.2 Experiment

Since the basic architecture of ANNs and the corresponding performance evaluation process have been covered in Chapter 2, the description of ANNs herein is limited to the needs of this chapter. The six properties collected from CRC Handbook [229] were: i) normal boiling point ($\boldsymbol{T_b}$), ii) critical point ($\boldsymbol{T_c}$), iii) critical pressure ($\boldsymbol{P_c}$), iv) dipole moment ($\boldsymbol{D_m}$), v) molecular weight ($\boldsymbol{M_w}$), and vi) enthalpy of vaporization at boiling point ($\boldsymbol{\Delta H_{vb}}$). The five properties were selected because they are considered to correlate with the enthalpy of vaporization to a certain degree. In total, the experimental data

consist of 175 organic and inorganic compounds that have full records of all six properties. All data are listed in Appendix XIV, among which T_b ranges from 82 K to 613 K, T_c ranges from 133 K to 850 K, P_c ranges from 2 MPa to 22 MPa, D_m ranges from 0.023 to 4.270 D, M_w ranges from 17 to 278, and ΔH_{vb} ranges from 6.04 kJ/mol to 79.20 kJ/mol.

The target property is ΔH_{vb}, i.e. the output of the ANNs and the SR models. Any combination of input variables chosen from the remaining five properties (T_b, T_c, P_c, D_m and M_w) is used as the inputs. Both the 'PaD' method and the 'CW' method are employed on the ANN models, which are constructed using the entire input candidates. Once ANNs determine the predominant input variables, SR is applied to derive the mathematical expression represented by the same parameters. Furthermore, to compare the prediction accuracy of the SR approach with the ANN results, and to prove the necessity of additional feature selections, SR models utilizing all 5 variables are constructed and the contribution of each input variable is analysed.

The platform of ANNs is Matlab 2010a, while SR is performed on Discipulus Genetic Programming System. The SR models are evaluated by taking the average of the absolute error over the total 175 data. A combination of input variables is considered valuable in predicting ΔH_{vb} if the average absolute error is less than 5% (the lower error, the better predictive power). In the meantime, a good combination should contain as few variables as possible.

The experiment is systematically performed in four steps: 1) performing an ANN combinatorial search; 2) analysing the input variable contribution through the 'PaD' method and the 'CW' method; 3) conducting symbolic regression within all available features; and 4) conducting symbolic regression within the features selected based on the first two steps.

6.2.1 The contribution of the different variables in ANNs

1) The 'PaD' method

The 'PaD' method computes the partial derivatives of the ANN output with respect to the inputs. It was first proposed by Dimopoulos *et al.* [220] using a logistic sigmoid function as the transfer function. Based on the same principle, Equation 6-1 is derived

for a network with 'Tansig-Purelin' as the transfer function. For a network with $\{X_i\}$ inputs (i=1, 2…n), one hidden layer with m neurones, and one output, if the transfer function between the input layer and the hidden layer is 'Tansig', and transfer function between the hidden layer and the output layer is 'Purelin', the partial derivatives of the output Y with respect to the input X_i is denoted as d_i (see Equation 6-1). The relative contribution of the input variables can be accessed by ranking their squared partial derivatives.

$$d_i = \sum_{j=1}^{m} IW_{i,j} LW_j \, sech^2\left(\sum_{i=1}^{n} IW_{i,j} X_i + b_j\right) \qquad \textit{Equation 6-1}$$

where:

IW — the weight vector between the input layer and the hidden layer, $IW_{i,j}$ is the weight assigned to the hidden neurone j with respect to the input X_i;

LW — the weight vector between the hidden layer and the output layer, LW_j is the weight assigned to the output Y with respect to the hidden neurone j;

b_j — the bias added to the hidden neurone j;

X_i — the i^{th} input;

2) The 'CW' method

The idea of 'CW' method is using the weight vector to reveal the relative importance of input variables [445]. According to Gevrey *et al.* [221], the weight partitioning process is calculated as below:

1) Construct a network with $\{X_i\}$ inputs (with i=1, 2…n), one hidden layer with m neurones, and one output. Obtain the weight vector IW between the input layer and the hidden layer, and the weight vector LW between the hidden layer and the output layer.
2) For each hidden neuron j, multiply the absolute value of the hidden-output layer connection weight by the absolute value of the hidden input layer connection weight for each input variable X_i.

$$P_{i,j=}\left|IW_{i,j}\right| \times \left|LW_j\right| \qquad \textit{Equation 6-2}$$

3) For each hidden neuron $\boldsymbol{j}$, divide $\boldsymbol{P_{i,j}}$ by the sum of the absolute value of the input-hidden layer connection weight of all input neurons.

$$Q_{i,j} = \frac{P_{i,j}}{\sum_{i=1}^{n} P_{i,j}} \qquad \textit{Equation 6-3}$$

4) For each input variable $\boldsymbol{X_i}$, sum up the $\mathbf{Q_{i,j}}$.

$$S_i = \sum_{j=1}^{m} Q_{i,j} \qquad \textit{Equation 6-4}$$

5) The relative importance of input variable $\boldsymbol{X_i}$, i.e. $\boldsymbol{CW_i}$, is obtained by dividing $\boldsymbol{S_i}$ with the sum of the $\boldsymbol{S_i}$.

$$CW_i = \frac{S_i}{\sum_{i=1}^{n} S_i} \qquad \textit{Equation 6-5}$$

6.2.2 SR modelling to obtain the mathematical expression

When the ANNs have been used to quantify the variable's importance, variables that contribute most to the prediction performance are selected as inputs to obtain SR results through Discipulus [446], a commercial genetic programming software package. It should be noted that the SR models are constructed with the same 175 compounds as in ANN models to predict the enthalpy of vaporization at boiling point. Similarly, Discipulus requires both training data and validation data sets in the model creation. Unlike ANNs where training data and validation data (i.e. test data) are split at a ratio of 4:1, training data and validation data in Discipulus are randomly but equally divided: 88 data for training and 87 data for validation, as recommended in Ref. [447]. The initial population size is set to be 500. The mutation frequency ($\boldsymbol{M_f}$), crossover frequency ($\boldsymbol{C_f}$) and reproduction frequency ($\boldsymbol{R_f}$) is set to be 95%, 50% and 2.5% respectively. Be noted, mutation is applied regardless whether the programme has been selected for crossover or not, thus the reproduction rate is calculated (in percentages) by Equation 6-6 [447]:

$$R_f = 1 - M_f - \left[C_f \times \left(1 - M_f\right)\right] \qquad \textit{Equation 6-6}$$

The operators function set consists of arithmetic operations (+, -, *, /), mathematical functions (exponential), comparison and exchange functions. Data values are converted into logarithm values as the standard procedure in performing SR [448], due to the lacking of logarithm operator in Discipulus. A single run of SR terminates at 300 generations without improvement in fitness. A total of 500 runs are performed in this

study. The SR models that have the least square error are selected and further interpreted into mathematical expressions.

6.3 Results and discussion

6.3.1 ANN combinatorial search

The ANN combinatorial search was conducted by using all the possible combinations of input variables (T_b, T_c, P_c, D_m and M_w) as inputs. The exhaustive search for the prediction of ΔH_{vb} stated by using one input variables (1st order search) to construct ANNs and increased to use more variables, until all the five variables were used together. The predictive capacity for each ANN model is tabulated in Table 6-1 at the descending order of the general ANN modelling performance (ψ). A proportion of 5% data that have the largest errors in predicting ΔH_{vb} are excluded in the average error calculation, in order to reduce the possible influence of noise data [234]. The evaluation criteria ψ, δ and (E_c) are described in Chapter 2 and Chapter 3. ANNs have better generalization ability when δ is smaller, and have good prediction accuracy when (E_c) is small.

As shown in Table 6-1, the error of the top five ANNs is around 2.6%, all of which contain input variables T_b, T_c and P_c, except Row 4. In contrast to the first row, because the addition of D_m in the third row has only 0.2% contribution on the modelling accuracy, D_m is regarded as an insignificant input. Similarly, by comparing Row 1 and Row 5, M_w is considered to be a less important variable. Furthermore, the comparison between Row 1 and Row 4 indicates P_c and M_w have similar contributions in predicting ΔH_{vb}. However, if only T_b and T_c are used as input variables, the prediction error dramatically increases from 2.7% (Row 1) to 4.3% (Row 11). In addition, the ANN model constructed using T_b overrides any other model constructed using only one input variable (Row 12), and the model constructed using T_c is the second best (Row 20). In this case, P_c or M_w is a necessary input variable though neither of them is as important as T_b or T_c. Thus, according to the general ANN modelling performance (ψ), which has taken the modelling accuracy (A) and the ANN generalization ability (δ) into consideration, the three input variables, namely, T_b, T_c and P_c, constitute the most promising correlation with ΔH_{vb}. The second promising combination of inputs utilizing three variables is M_w, T_b and T_c.

Figure 6-1 presents the normalized frequency of variables in the top 20% ANN models during the ANN combinatorial search of 1st order, 2nd order, 3rd order and 4th order (the I^{th} order model refers to the model constructed using I different input variables). As can be seen from the sum of normalized frequency, T_b is the largest contributed variable, and followed by T_c, P_c and M_w. The last candidate, i.e. D_m, has the least contribution in these top 20% predictions, thus are regarded to be the less-significant variable.

Figure 6-2 shows the results of ANNs in predicting ΔH_{vb} from three different input variable combinations: i) T_b, ii) T_b, T_c and P_c, and iii) M_w, T_b and T_c. As can be seen, the prediction accuracy of testing data (indicated by the green square points) is slightly less than the training data (indicated by the blue circle points). However, the determination coefficient R_l of the total 175 data is 0.964, 0.985, and 0.980 respectively, thus these models are believed to give very reliable predictions.

6.3.2 The 'PaD' 'and 'CW' method

Figure 6-3 presents the relative contributions of the five variables with maximum-minimum range that are accessed by the 'PaD' method and the 'CW' method on three independent ANN models. All data are listed in Appendix XV. The two methods reach similar conclusions that they both indicate T_b is the variable that has the largest contribution, and then followed by T_c, P_c, M_w and D_m. There is no significant contribution difference between P_c and M_w, and it shows D_m has the least impact. The results are in good agreement with the discussions in Section 6.3.1. Compared to the result of the 'CW' method, the large variation in the evaluation of P_c indicates the 'PaD' method is less stable. However, the variable contribution difference based on the 'PaD' method is larger than that based on the 'CW' method, making it more visible in terms of contribution evaluation.

Table 6-1 ANN models in the prediction of ΔH_{vb} using different input variables.

Importance	ANNs evaluation parameters			Input properties
	ψ	δ	E_c	
1	0.04	0.029	2.7%	T_b, T_c, P_c
2	0.05	0.039	2.7%	M_w, T_b, T_c, P_c, D_m

Importance	ANNs evaluation parameters			Input properties
	ψ	δ	E_c	
3	0.05	0.042	2.5%	T_b, T_c, P_c, D_m
4	0.05	0.044	2.6%	M_w, T_b, T_c
5	0.06	0.049	2.5%	M_w, T_b, T_c, P_c
6	0.06	0.049	3.5%	M_w, T_b, D_m
7	0.06	0.054	3.6%	M_w, T_b, T_c, D_m
8	0.07	0.063	3.8%	M_w, T_b, P_c, D_m
9	0.08	0.068	4.3%	M_w, T_b
10	0.08	0.071	4.2%	M_w, T_b, P_c
11	0.08	0.071	4.3%	T_b, T_c
12	0.09	0.079	3.8%	T_b
13	0.09	0.082	4.1%	T_b, T_c, D_m
14	0.10	0.089	4.1%	T_b, P_c
15	0.10	0.087	4.7%	T_b, D_m
16	0.10	0.091	4.1%	T_b, P_c, D_m
17	0.11	0.103	4.9%	T_c, P_c, D_m
18	0.12	0.109	4.9%	M_w, T_c, P_c, D_m
19	0.12	0.108	5.6%	M_w, T_c
20	0.12	0.11	5.2%	T_c
21	0.14	0.127	5.6%	M_w, T_c, P_c
22	0.15	0.134	5.7%	T_c, P_c
23	0.16	0.148	6.2%	M_w, T_c, D_m
24	0.16	0.151	6.3%	T_c, D_m
25	0.66	0.649	14.0%	M_w, P_c, D_m
26	0.70	0.689	14.9%	D_m
27	0.71	0.688	15.8%	P_c, D_m
28	0.76	0.746	15.3%	M_w, D_m
29	0.95	0.94	15.6%	M_w, P_c
30	0.97	0.955	15.1%	M_w

Importance	ANNs evaluation parameters			Input properties
	ψ	δ	E_c	
31	1.15	1.138	18.4%	P_c

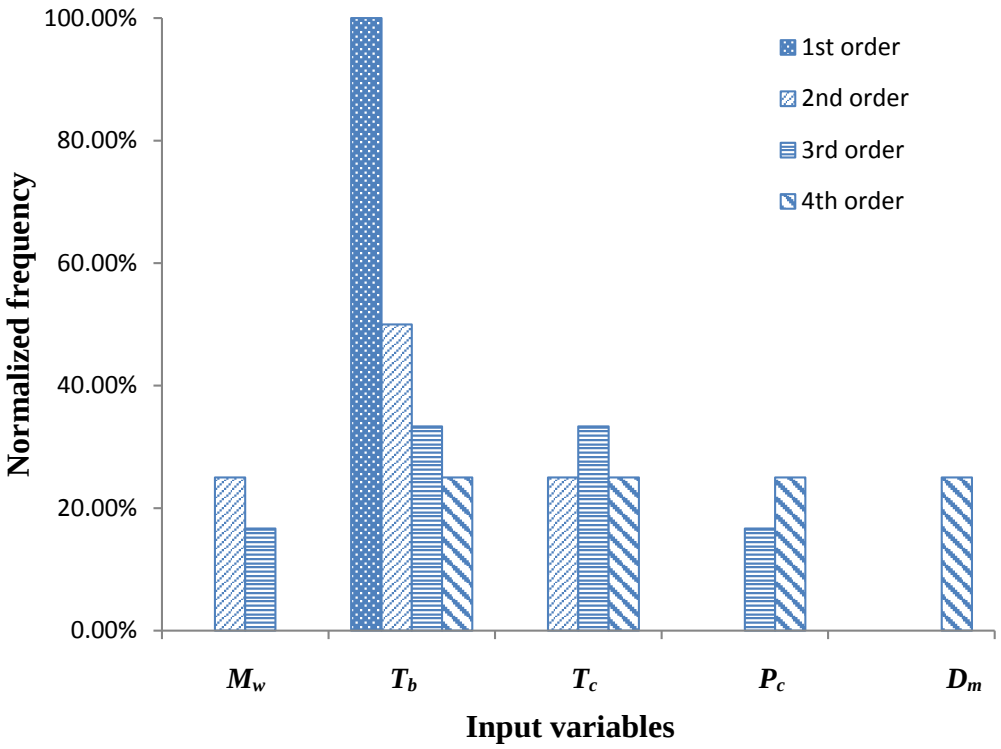
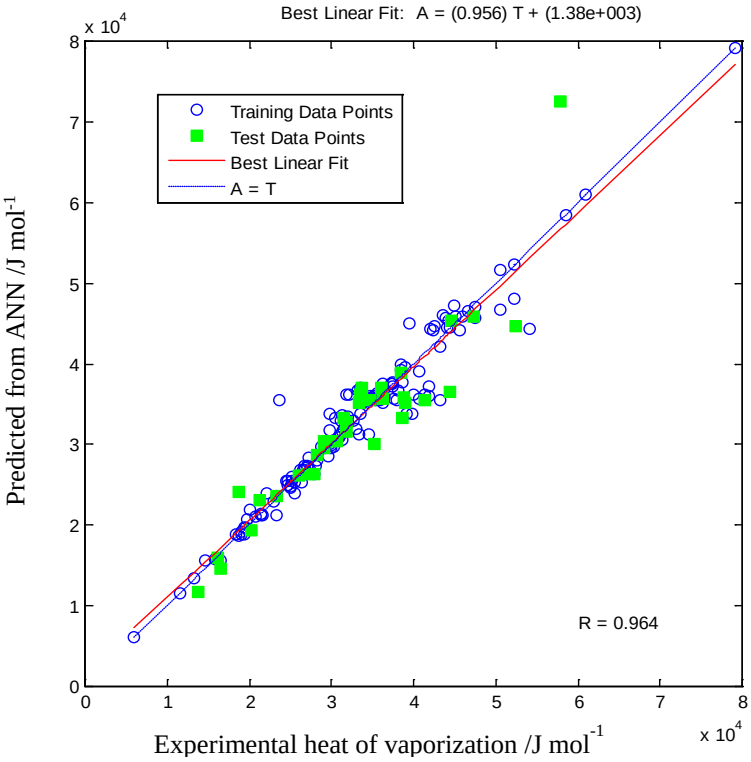
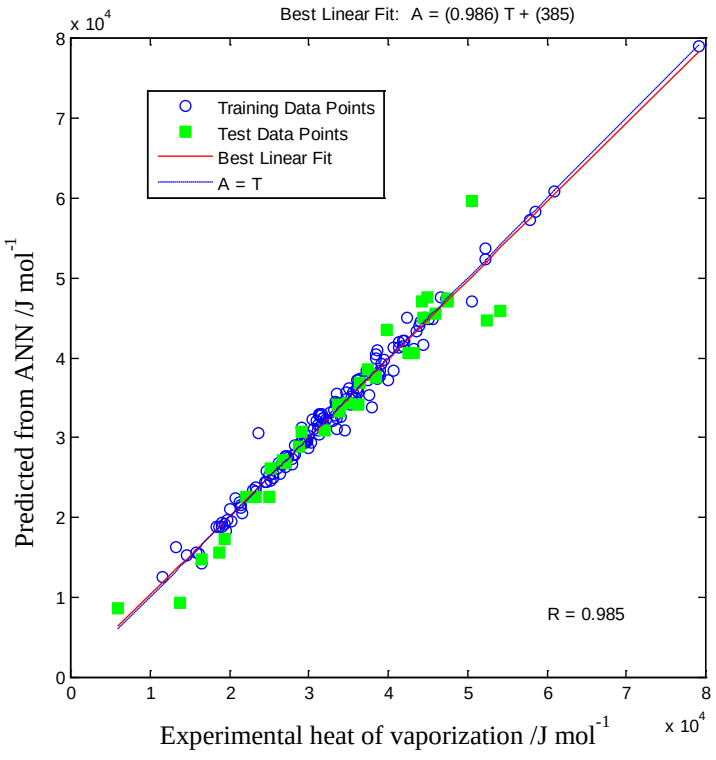


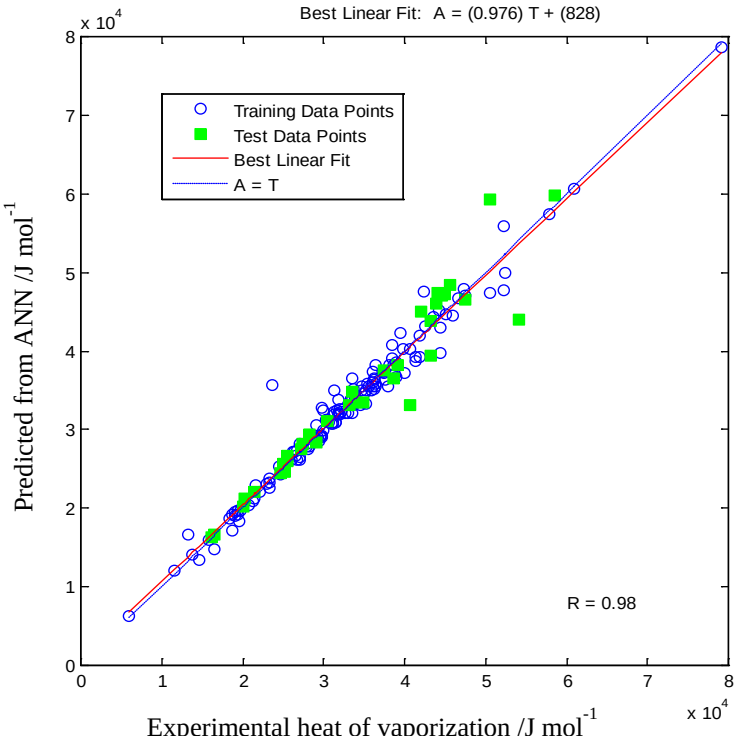
Figure 6-1 The property importance revealed in the top 20% correlations of 1st, 2nd, 3rd and 4th order in predicting ΔH_{vb} from the ANN combinatorial search.


Best Linear Fit: A = (0.956) T + (1.38e+003)
x 10^4
Training Data Points
Test Data Points
Best Linear Fit
A = T
R = 0.964
Predicted from ANN /J mol^-1
Experimental heat of vaporization /J mol^-1
x 10^4

(a)


Best Linear Fit: A = (0.986) T + (385)
x 10^4
Training Data Points
Test Data Points
Best Linear Fit
A = T
R = 0.985
Predicted from ANN /J mol^-1
Experimental heat of vaporization /J mol^-1
x 10^4

(b)



(c)

Figure 6-2 Results of ANNs in predicting ΔH_{vb} from a) T_b with δ equals to 0.079 and an error of 3.8%; b) T_b, T_c, P_c with δ equals to 0.029 and an error of 2.7%; and c) M_w, T_b, T_c with δ equals to 0.044 and an error of 2.6%.

6.3.3 The important input variables - T_b, T_c and P_c

ANN results show T_b is the most important variable among the five input candiadates. In fact, Kistiakowsky [208] has proposed a simple equation to make a fast estimation of ΔH_{vb}. As shown in Equation 6-7, the predicted value from this approach generally falls within ±5% of the experiment value [211].

$$\frac{\Delta H_{vb}}{T_b} = 36.6 + R\ ln\ T_b \qquad \textit{Equation 6-7}$$

The ANN study also shows that a combination of input varibles T_b, T_c and P_c produces the good performance for the prediction of ΔH_{vb}. The finding is in agreement with the empirical results of other researchers [208, 211, 449, 450]. Equation 6-8 to Equation 6-11 are the ones that have been widely employed to make good estimates of ΔH_{vb}. Their predicted values generally fall within ±3% of the experiment values [211]. It is interesting to note that ANNs have obtained a similar accuracy.

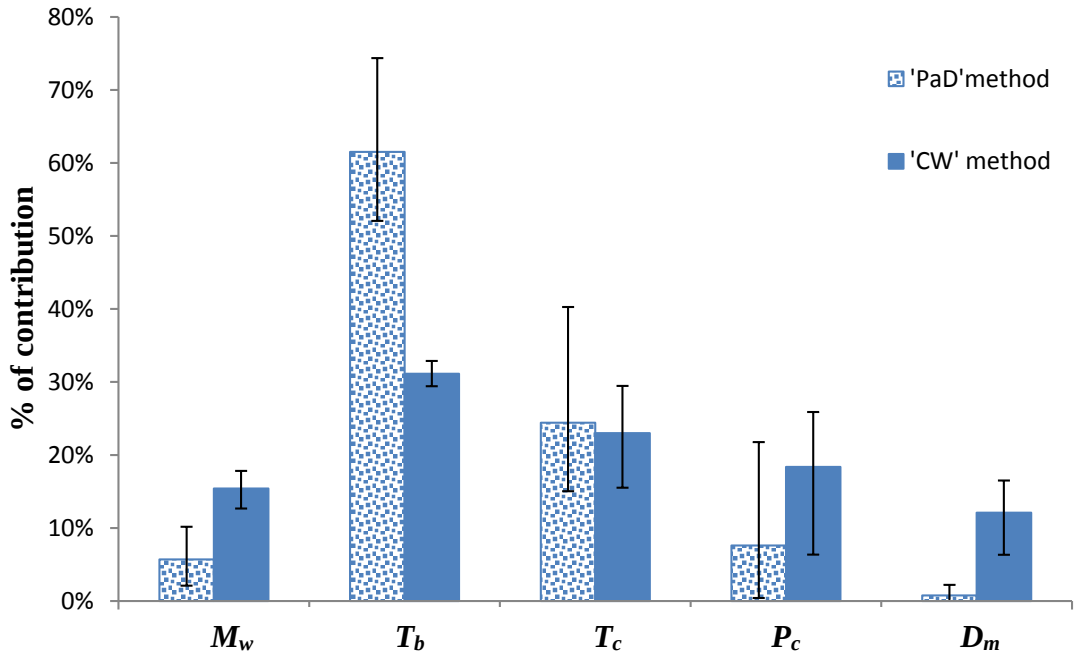


Figure 6-3 The contributions of the 5 variables accessed by the 'PaD' method and the 'CW' method in the prediction of ΔH_{vb}. The error bars indicate the maximum and minimum values obtained in 3 independent ANN models.

The Giacalone's equation [449]:

$$\Delta H_{vb} = RT_cT_{br}\left[\frac{ln(P_c/1.01325)}{1-T_{br}}\right] \qquad \textit{Equation 6-8}$$

The Riedel's equation [208]:

$$\Delta H_{vb} = 1.093RT_cT_{br}\left(\frac{ln\,P_c-1.013}{0.93-T_{br}}\right) \qquad \textit{Equation 6-9}$$

The Chen's equation [450]:

$$\Delta H_{vb} = RT_cT_{br}\left(\frac{1.555ln\,P_c+3.978T_{br}-3.958}{1.07-T_{br}}\right) \qquad \textit{Equation 6-10}$$

The Vetere's equation [211]:

$$\Delta H_{vb} = RT_cT_{br}\left(\frac{0.4343ln\,P_c+0.89584T_{br}-0.69431}{0.37691-0.37306T_{br}+0.15075P_c{}^{-1}T_{br}{}^{-2}}\right) \qquad \textit{Equation 6-11}$$

where:

T_b = boiling point in K;

T_c = critical point in K;

P_c = critical pressure in bar;

T_{br} = reduced boiling point, which equals to T_b/T_c;

R = gas constant, which equals to 8.314 J mol^{-1} K^{-1}.

6.3.4 SR model analysis

1) SR models using all five input variables

Though GP might be superior to other data-driven techniques with or without feature selection in respect to the prediction accuracy [451], to identify the optimal functional form of statistical models, the determination of input variables is a very important and challengeable task for all data-driven techniques [108]. Generally, to obtain a robust SR model, the input variables not only need to be diverse enough to describe the knowledge hidden in the experimental data, but also succinct enough to acquire a parsimonious equation [452]. It is the author's experience that the largest difficulty of SR lies in obtaining the parsimonious equations from hundreds or even thousands 'evolved' programmes without substantially sacrificing the accuracy. Generating SR programmes within different input variables only exacerbate the problem as more 'evolved' programmes need to be analysed.

In this study, the ANN approach is used as an alternative method to identify the most promising input variables without any prior knowledge. The applications of ANNs to predict material properties are readily available in many studies, and explanatory methods, such as the 'PaD' method and the 'CW' method, have been developed to analyse the contribution of each input variable. Because all the five available variables (M_w, T_b, T_c, P_c and D_m) are used as inputs in both methods (the 'PaD' method and the 'CW' method), it is interesting to address the performance of SR utilizing the same five variables as inputs.

Discipulus provides a way to look through all of the best 30 SR models generated in one experiment [447]. By analysing the variable appearances in models, and the average and maximum impacts of variables when they are removed and replaced by the average value of that input, the contribution of each input variable to the fitness of models can be accessed. Figure 6-4 presents the frequency, the average impact, and the maximum impact of the five variables in the best 30 SR models obtained in three independent

experiments. All data are listed in Appendix XVI. The error bars indicate the maximum and minimum values obtained in the 3 experiments. The average prediction accuracy of SR models is around 95%. It should be noted that frequency and the impact are two independent characters to describe the behaviours of the input variables. A value of 100% of frequency indicates that the input variable appeared in 100% of the best 30 SR models. To be specific, $\boldsymbol{T_b}$, $\boldsymbol{T_c}$, $\boldsymbol{M_w}$ and $\boldsymbol{D_m}$ are most commonly used, that they are found in more than ~90% of models. Even the last variable $\boldsymbol{P_c}$ is likely to occur in half of the models.

Neither the average impact nor the maximum impact of $\boldsymbol{T_b}$, $\boldsymbol{T_c}$, $\boldsymbol{P_c}$, $\boldsymbol{M_w}$ and $\boldsymbol{D_m}$ shows identical trends with the changes of the frequency. A higher frequency does not necessarily imply a larger influence on the output. This is probably because programs generated by SR usually contain instructions that utilize all the inputs including redundant inputs, for example, inputs within small changes might be used as constants in SR models [447]. Moreover, the variation trend of the average impact of the five variables does not correspond to that of the maximum impact, indicating different components in the SR models. Like the 'PaD' and the 'CW' method, $\boldsymbol{T_b}$ and $\boldsymbol{T_c}$ are identified to be the variables with largest impacts. However, it is difficult to distingush the contribution made by $\boldsymbol{P_c}$, $\boldsymbol{M_w}$ and $\boldsymbol{D_m}$. The average impacts of $\boldsymbol{P_c}$, $\boldsymbol{M_w}$ and $\boldsymbol{D_m}$ are smaller than 3%, while the maximum influence of the three input variables varies. The maximum impact of $\boldsymbol{M_w}$ and $\boldsymbol{D_m}$ can reach as large as ~60% and ~90% respectively.

As can be seen, without a feature selection in prior, SR tends to use all available input variables in the terminals set to build the model, in which some variables may have very limited contributions. Huge time and effort are required to interpret a complicated SR model into a parsimonious correlating equation due to the existing redundancy. On the other hand, by analysing SR models, though it is possible to identify the primary variables that have the largest impacts, the ability to identify the secondary important variables is restricted. As a result, it is important for ANNs to act as an additional yet automatic method to determine necessary input variables to feed into SR models.

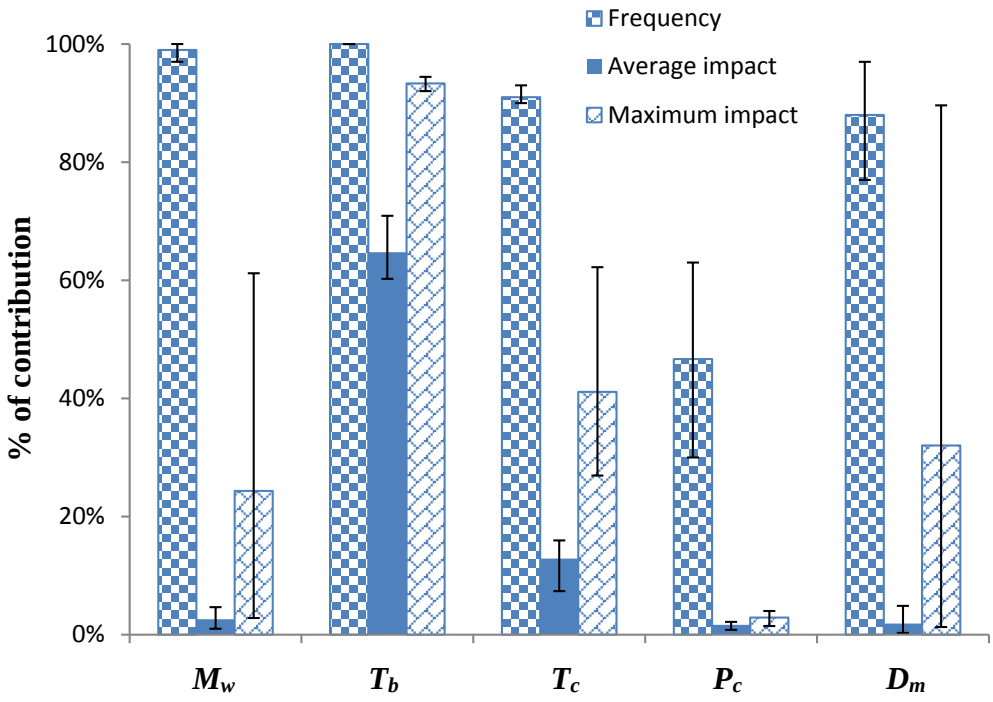


Figure 6-4 The contributions of the 5 variables accessed by Discipulus (SR method). The error bars indicate the maximum and minimum values obtained in 3 independent experiments.

2) Mathematical expression using variables selected by ANNs

As discussed in the previous section, without any prior knowledge, the ANN study including both the combinatorial search and the 'PaD'/'CW' method, suggests that ΔH_{vb} can be successfully predicted by using T_b, and the model constructed by T_b, T_c and P_c results in higher accuracy and better generalization capability. Though not all the equations intelligently generated by SR appear to offer advantages over Equation 6-8 to Equation 6-11, Equation 6-12 and Equation 6-13 are the ones the author obtained that are at least as accurate as the well-known equations mentioned in Section 6.3.3.

Table 6-2 and Table 6-3 compare the predicted and experimental values of ΔH_{vb} by applying these two equations (Equation 6-12 and Equation 6-13) to the same 175 compounds used in constructing the ANNs and the SR models. As shown in Table 6-2, the average error percentage decrease from 7.45% to 6.33% by using the new SR equation utilizing only one variable, i.e. T_b. More than 20% data have predicted values fall within ±5% of the experiment values compared to the results calculated by the Kistiakowsky's equation. The SR equation obtained using three input variables, i.e. T_b,

T_c and P_c, is generally simpler than any other equations readily available in literature. In addition, it produces more accurate predictions than the Riedel's equation, and yields the same error (~ 3.8%) as the remaining equations do.

Nevertheless, SR has limitations in finding the optimal constant in the expression [192]. Especially for Equation 6-13, it is hard to attribute real world meaning to interpret the constant of power assigned to each variable. It might be resolved by specifying constant form at the beginning of SR modelling, for example, only integers can be assigned to the power operator. Another concern about SR model is its complexity and its analytic form. Due to the uncertainty of the evolution, the 'perfect' mathematical equations to be distinguished from a large number of equivalent models still need human effort [28, 452]. However, there is no doubt that SR will show superiority in deriving the corresponding mathematical equations to describe the observed phenomena that not yet accessible for physical modelling.

$$ln\,\Delta H_{vb} = 1.11\,ln\,T_b + 3.82 \qquad \textit{Equation 6-12}$$

$$\Delta H_{vb} = \frac{1.056{T_b}^{3.152}{P_c}^{0.342}}{{T_c}^{2.131}} \qquad \textit{Equation 6-13}$$

*Table 6-2 Comparison of the calculation accuracy of the Kistiakowsky's equation and the equation obtained by SR (indicated by *) using one input variable (T_b) for 175 compounds.*

Equation formula	**Average error**	**Number of compounds**		
		Error%<5%	**Error% ∈(5%, 10%]**	**Error%>10%**
$\frac{\Delta H_{vb}}{T_b} = 36.6 + R\ln T_b$	7.45%	90	38	47
$\ln \Delta H_{vb} = 1.11 \ln T_b + 3.82^*$	6.33%	112	26	37

*Table 6-3 Comparison of the various equations available in literature and the equation obtained by SR (indicated by *) using three input variables, i.e. T_b, T_c and P_c.*

Equation	Formula	Average error
Giacalone's equation [449]	$\Delta H_{vb} = RT_cT_{br}\left[\frac{\ln(P_c/1.01325)}{1-T_{br}}\right]$	3.80%
Riedel's equation [208]	$\Delta H_{vb} = 1.093RT_cT_{br}\left(\frac{\ln P_c - 1.013}{0.93-T_{br}}\right)$	5.65%
Chen's equation [450]	$\Delta H_{vb} = RT_cT_{br}\left(\frac{1.555\ln P_c + 3.978T_{br} - 3.958}{1.07-T_{br}}\right)$	3.54%
Vetere's equation [211]	$\Delta H_{vb} = RT_cT_{br}\left(\frac{0.4343\ln P_c + 0.89584T_{br} - 0.69431}{0.37691 - 0.37306T_{br} + 0.15075{P_c}^{-1}{T_{br}}^{-2}}\right)$	3.50%
New SR equation*	$\Delta H_{vb} = \frac{1.056{T_b}^{3.152}{P_c}^{0.342}}{{T_c}^{2.131}}$	3.78%

6.4 Conclusion

This chapter discussed a method of determining materials property correlations by the use of the combination of artificial neural networks and symbolic regression. ANNs are applied to determine the smallest number of important input variables that are sufficient to represent the correlation hidden in the data. Then SR is employed to derive the mathematical equations that represent these relationships using the variables selected by ANNs. As an example, the determination of the enthalpy of vaporization at boiling point is presented. By applying both ANNs and SR, ΔH_{vb} was successfully predicted through the application of both approaches: in the case of 175 compounds tested, the present method gives an average error of ~3% and ~4% respectively.

To provide an explanatory insight of the ANN 'black-boxes', a comparison study of the 'PaD' method and the 'CW' method is performed on the contribution-evaluation of five input variables, namely, M_w, T_b, T_c, P_c and D_m. Though the 'CW' method is more stable, the variable contribution difference based on the 'PaD' method is more visible. The two methods reach similar conclusions that T_b is the variable that has the largest contribution, and then followed by T_c, P_c, M_w and D_m. As the total number of input

variables (5) is much smaller than the example data points (175), it is appropaite to conduct a combinatorial search that gurantte the optimal solution. Results of combinatorial search are in good agreement with the 'CW' method and the 'PaD' method, that T_b is the most important variable among the five input candiadates. A combination of input varibles T_b, T_c and P_c produces the best performance for the prediction of ΔH_{vb}, followed by a combination of input varibles M_w, T_b and T_c.

Meanwhile, the contribution of each input variable is also accessed based on the SR approach using all the five variables. Results indicate that without the feature selection in advance, SR tends to use all the input variables to build the model regardless how much they contribute. Though it is possible for SR to identify the primary variables that have the largest impacts, SR has rendered ability to identify the secondary important variables. As a result, ANNs act as an additional yet automatic method to determine necessary input variables to feed into SR models. In this experiment, the SR models are thus constructed utilizing T_b and a combination of input varibles T_b, T_c and P_c. Compared to previously reported models, the equations obtained by SR are simpler and are at least as accurate as other models in providing a rapid estimation of ΔH_{vb}.

In conclusion, the ultimate goal of the combinatorial application of ANNs and SR is to accelerate the human's pace to discover new materials and physical laws. This study demonstrates that ANNs are capable of exploring correlations that might exist between different properties in materials without needing the knowledge of the direct structure-property relationships. By employing SR, the corresponding mathematical equations can be derived to explicate the prior phenomena observed by ANNs.

7. General conclusion and future work

7.1 General conclusion

To summarise, the applications of ANNs in exploring materials property-property correlations are demonstrated by four distinct examples, where no prior knowledge of the form of the relationship is required:

1) In Chapter 3, binary, ternary and quaternary order correlations are mined from the database that contains 24 properties for 37 pure metals through a combinatorial search.
2) In Chapter 4, the ternary order correlations between Young's modulus, shear modulus, bulk modulus and Poisson's ratio are captured from the database that contains 68 pure metals.
3) In Chapter 5, the relationships between the hardness increment due to HPT and 13 physical, mechanical, and electronic properties are investigated for 17 pure metals.
4) In Chapter 6, the relationships between the enthalpies of vaporization of 175 compounds and five other properties, namely, M_w, T_b, T_c, P_c and D_m are determined.

In order to assess the correlation importance (i.e. how strong the correlation is), two criteria are introduced to identify meaningful and nontrivial property correlations. They are: model generalization ability (δ) and correlation error (E_c). In particular, as ~5% data are considered contaminated in most databases, error calculations excluded 5% data with the highest error. For a correlation-group constituted by N correlations, the group performance (ψ_N) can be evaluated according to a quadratic superposition rule.

The correlations captured by ANNs provide a way of illuminating the facts hidden in the data, and can be applied to:

1) The materials properties prediction. For example, shear modulus, melting temperature and Burgers vector can be used to predict the ΔH_V due to HPT at an

accuracy of ~92%. The error to predict ΔH_{vb} using the values of T_b, T_c and P_c is less than 3%.

2) The errors correction in materials properties databases. For example, ANNs are successfully employed to resolve the large discrepancies existing in the elastic property data, which are recorded in prestigious handbooks and databases.
3) The identification of important input variables. Input variables need to be determined for the establishing of physical models (as described in Chapter 5) and the construction of explicit correlating equations (as illustrated in Chapter 6). The variable contribution can be accessed through the ANN combinatorial search, the forward selection method, and the methods utilizing the weight vectors of ANNs, such as the 'PaD' method and the 'CW' method.

For a discrete, irregularly distributed database that is subject to unknown error, performing ANN combinatorial searches is the only method that evaluates all of the possible correlations. By utilizing the evaluation criterion to rank the correlation based on its model performance, the combinatorial search using ANNs is guaranteed to determine the optimal set of input variables. Because ANNs can be performed in parallel, the combinatorial approach can be high-throughput and effective simultaneously. When the workload cannot be aligned with parallel processing, the forward selection method is suitable to find the most important input parameters that have maximum influence on the predicted results. The forward selection is particularly effective when the database is relatively small but the search space of inputs is quite large. For a database that is large enough compared to the relatively small search space of inputs, variables' contributions can be analysed by utilizing the weight vectors of ANNs through the 'PaD' method and the 'CW' method.

The ANN method can not only be used for complex system modelling with a large input dataset, but is also feasible in a situation where the size of dataset has an inherent limitation. As can be seen, the datasets that are used in the construction of ANNs vary from a few data points (17 data points in Chapter 5) to hundreds of data points (175 data points in Chapter 6). As long as the amount of data is enough to determine the number of fitting parameters in the networks, and the generalization ability of ANN models is tested, researchers can still benefit from ANNs when dealing with a small dataset.

7.2 Original contribution of the thesis

1) ANNs have been used extensively in materials science for predicting a target property from materials' composition or processing parameters. Studies in exploring materials property-property correlations (excluding compositional information or processing parameters) are relatively rare. The work presented in this thesis employed ANNs to explore the correlations of different materials properties. It has extended the application of ANNs in the fields of data corrections, property predictions and identifications of variables' contributions.
2) The traditional idea of combinatorial methods applied to materials science is the identification of the causal pathways that link composition and structure to properties. A combinatorial approach has seldom been applied to correlate materials' properties. In this thesis, ANNs combinatorial searches are employed as high-throughput screening methods to produce possible property- property correlations for analysis. Instead of creating composition or structure 'libraries', property 'libraries' are prepared and analysed. Because ANNs are capable of evaluating correlations in parallel, the advantages of combinatorial approach that guarantee to find the optimal correlations in the shortest possible time are retained.
3) An evaluation criterion to rank correlations and correlation-groups accessed by ANNs, is proposed and demonstrated to be useful in identifying meaningful and nontrivial property-property correlations. The criterion utilises the parameters of model performance, ANNs generalization ability and correlation error. In addition, such evaluation criterion can be applied according to a quadratic superposition rule to justify the importance of an observed correlation-group. It is especially effective when the difficulty of correlations exploration is exacerbated with the increase of complexity that exists in the correlations and the growing data size.
4) A large number of strong correlations are obtained through the ANN combinatorial search, among which the typical examples were discussed emphatically in Chapter 3. Some correlations have already been proposed in literature and used extensively in materials sciences, such as the relation between atomic weight and specific heat capacity. While, some correlations

have not been previously attended to, now are brought to researchers' attention, for example, the relation between cohesive energy of solid and heat of vaporization. The correlations captured by ANNs indicate that the ANN combinatorial search without prior knowledge, as the work presented in this thesis, is a systematic and effective way of illuminating useful information.

5) Because ANN combinatorial searches test all possible property combinations, the strong correlations obtained by ANNs may have redundant variables, especially for correlations of higher order. By comparing with the correlations of lower order, binary, ternary and quaternary order correlation-groups are summarized into different categories in Chapter 3. Such classifications could help to identify the fundamental materials property-property correlations.
6) Young's modulus, shear modulus, bulk modulus and Poisson's ratio of pure metals are important mechanical properties in materials research, but the scatter of their measurement results has drawn very little attention. The situation is now changed by the work presented in Chapter 4, where the substantial discrepancies in several prestigious handbooks and databases are clarified for the first time. Meanwhile, a latest compilation of experimental elastic properties data for pure metals is provided, including a comprehensive review of single crystal elastic constants and polycrystalline elastic data.
7) Extending the study using binary order correlations to correct suspect data, a systematic ANN protocol has been developed in Chapter 4 to capture and utilize ternary order correlations. The ANN method is tested against the known correlating equations, and is found to be reliable and effective in verifying suspect data. Because this method does not require any prior knowledge, it would also help to solve the problem of identifying outliers and correcting errors in a more general situation.
8) In Chapter 5, the increments of Vickers hardness of pure metals due to HPT are quantitatively and accurately predicted from a limited set of properties, i.e. shear modulus, melting temperature and Burgers vector. The first two variables are identified to have the largest impacts on hardening, while the last variable has a minor contribution. The correlations obtained by ANNs shed a light on the construction of a physically based model by pointing out the input parameters of greatest importance.

9) ANNs are generally used for complex system modelling with a large input dataset, but scientists would also like to benefit from ANNs in a situation where the size of data has an inherent limitation. Because only 17 data on the hardness of HPT-processed pure metals are available, the question that whether fruitful ANN models can be obtained when dealing with a small dataset is addressed. The work presented in Chapter 5 suggests that effective ANNs can be established even with a limited supply of data.
10) Without feature selection in advance, the equation model generated by SR tends to grow in complexity over the number of input variables. A combined ANN-SR method is proposed to yield more parsimonious model structures by ruling out redundant variables. An example to determine the enthalpies of vaporization of 175 organic and inorganic compounds is presented in Chapter 6. Two simple equations obtained by the combined method are found to be at least as accurate as other models in providing a rapid estimation of ΔH_{vb}.
11) A comparison study of the 'PaD' method and the 'CW' method is also performed in Chapter 6 to evaluate the contributions of five input variables. T_b is identified to be the variable that has the largest contribution to the prediction of ΔH_{vb}, and then followed by T_c, P_c, M_w and D_m. Though the result of the 'CW' method is more stable, the variable contribution difference based on the 'PaD' method is found to be more visible.

7.3 Future work

The following aspects are suggested:

1) Comparison of ANNs and other data mining methods.

Despite ANNs being probably the most well-known data driven method, other data mining techniques have become available for correlation modelling, property prediction and data forecasting, such as random forests [194], decision trees [195], support vector machines [196], and genetic programming (GP) [197]. It is interesting to perform comparative studies of these techniques and decide which would be most appropriate for future model applications over a range of experimental conditions. The comparative studies would be challenging but rewarding if a large number of modelling techniques

are evaluated, and conclusions are drawn based on a wide range of datasets that have unique characteristics.

2) Exploration of property correlations hidden in various data

It has been demonstrated that noises and errors exist in many databases, which sometimes are even considered as a fundamental dimension of data [235]. However, with the ever-increasing drive for accurate data, it is recommended to compile different databases and employ ANNs to detect outliers and correct errors for properties that have not been verified in this study.

It is evident that the property correlations captured by ANNs strongly depend on the characteristic of available data. This study mainly focused on the properties of pure metals, it would be interesting to apply the ANNs combinatorial method to explore the property correlations of alloys and functional ceramic materials.

3) Modification of the ANN architecture.

A relatively simple but sufficient enough three-layer feed-forward neural network with the back-propagation learning method is devised to explore correlations in this study. It is interesting to investigate the influence of different ANN architectures in order to enhance the performance of ANNs. This is can be done by adding one or more hidden layers, changing the training function/transfer function, and utilizing different learning rules.

4) Development of an integrated ANN-SR programme.

In this study, a commercial software package (Discipulus) is used for implementing the standard symbolic regression method. Though Discipulus shows several advantages in terms of execution time [447], prediction accuracy [214] and user-friendly interface, essentially, it is a program-based GP that executes machine code directly. Thus it is difficult to derive explicit equations from Discipulus due to the absence of adjustable parameters such as tree size restriction [205]. As the ANN models in this study are constructed based on Matlab, and GPLAB [453], a GP toolbox for Matlab that provides the evolved equation in the tree form, has become available, it is desirable to have an integrated ANN-SR programme developed in Matlab environment to capture correlations and obtain explicit correlating equations simultaneously.

References

1. Churchill, S.W., *The art of correlation.* Industrial & Engineering Chemistry Research, 2000. **39**(6): pp. 1850-1877.
2. Ashby, M.F., *Checks and estimates for material properties. I. Ranges and simple correlations.* Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences, 1998. **454**(1973): pp. 1301-1321.
3. Linoff, G.S. and M.J. Berry, *Data Mining Techniques : For Marketing, Sales, and Customer Relationship Management*. 2011: John Wiley & Sons Inc, pp.211-254.
4. Bhadeshia, H.K.D.H., *Neural networks in materials science.* ISIJ International, 1999. **39**(10): pp. 966-979.
5. Bhadeshia, H.K.D.H., *Neural networks and information in materials science.* Statistical Analysis and Data Mining, 2009. **1**(5): pp. 296-305.
6. Presnell, S.R. and F.E. Cohen, *Artificial neural networks for pattern-recognition in biochemical sequences.* Annual Review of Biophysics and Biomolecular Structure, 1993. **22**: pp. 283-298.
7. Zhang, Y., J.R.G. Evans, and S. Yang, *Detection of material property errors in handbooks and databases using artificial neural networks with hidden correlations.* Philosophical Magazine, 2010. **90**(33): pp. 4453-4474.
8. Zhang, Y., S. Yang, and J.R.G. Evans, *Revisiting Hume-Rothery's rules with artificial neural networks.* Acta Materialia, 2008. **56**(5): pp. 1094-1105.
9. Rühle, M., H. Dosch, E.J. Mittemeijer, and M.H. Van de Voorde, *European White Book on Fundamental Research in Materials Science*. 2001, Stuttgart: Max-Planck-Institut für Metallforschung, p.10.
10. Zhang, Y., *Applications of artificial neural networks (ANNs) in several different materials research fields,* PhD Thesis, School of Engineering and Materials Science, 2010, Queen Mary, University of London: London. pp. 2-67.
11. Moore, D. and W.I. Notz, *Statistics: Concepts and Controversies*. 7th Edn. 2009: W. H. Freeman, pp.311-336.
12. Han, J. and J.H.M. Kamber, *Data Mining: Concepts and Techniques*. 2nd Edn. 2006: Elsevier Books, Oxford, pp.1-5.
13. Goebel, M. and L. Gruenwald, *A survey of data mining and knowledge discovery software tools.* SIGKDD Exploration Newsletter, 1999. **1**(1): pp. 20-33.
14. Maojo, V. and J. Sanandrés, *A Survey of Data Mining Techniques Medical Data Analysis*, R. Brause and E. Hanisch, Editors. 2000, Springer Berlin / Heidelberg, pp. 77-92.
15. Mhamdi, F. and M. Elloumi. *A new survey on knowledge discovery and data mining*. Proceedings of the IEEE International Conference on Research Challenges in Information Science. 2008. Marrakech, Morocco: IEEE, pp. 427-

432.

16. Huang, J.Z., L. Cao, and J. Srivastava, *Advances in Knowledge Discovery and Data Mining*, in: 15th Pacific-Asia Conference, Shenzhen, China, May 24-27. 2011: Springer, pp.246-296.
17. Chen, M.-S., J. Han, and P.S. Yu, *Data mining: an overview from a database perspective.* IEEE Transactions on Knowledge and Data Engineering, 1996. **8**(6): pp. 866-883.
18. Cao, L., P.S. Yu, C. Zhang, and Y. Zhao, *Domain Driven Data Mining*. 2010: Springer, pp.17-25.
19. Cios, K.J., *Data Mining: A Knowledge Discovery Approach*. 2007: Springer, pp.5-6.
20. Turing, A.M., *Computing machinery and intelligence.* Mind, 1950. **59**: pp. 433-460.
21. Russell, S.J. and S.R.P. Norvig, *Artificial Intelligence: A Modern Approach*. 2010: Prentice Hall, pp.563-596, 822-831.
22. Rajpara, S.M., A.P. Botello, J. Townend, and A.D. Ormerod, *Systematic review of dermoscopy and digital dermoscopy/artificial intelligence for the diagnosis of melanoma.* British Journal of Dermatology, 2009. **161**(3): pp. 591-604.
23. Ezziane, Z., *Applications of artificial intelligence in bioinformatics: a review.* Expert Systems with Applications, 2006. **30**(1): pp. 2-10.
24. Park, K.S. and S.H. Kim, *Artificial intelligence approaches to determination of CNC machining parameters in manufacturing: a review.* Artificial Intelligence in Engineering, 1998. **12**(1-2): pp. 127-134.
25. Klopman, G., *Artificial intelligence approach to structure activity studies - computer automated structure evaluation of biological-activity of organic-molecules.* Journal of the American Chemical Society, 1984. **106**(24): pp. 7315-7321.
26. Chau, K.W., *A review on integration of artificial intelligence into water quality modelling.* Marine Pollution Bulletin, 2006. **52**(7): pp. 726-733.
27. Kalogirou, S.A., *Artificial intelligence for the modeling and control of combustion processes: a review.* Progress in Energy and Combustion Science, 2003. **29**(6): pp. 515-566.
28. Schmidt, M. and H. Lipson, *Distilling free-form natural laws from experimental data.* Science, 2009. **324**(5923): pp. 81-85.
29. Zahedi, F., *An introduction to neural networks and a comparison with artificial-intelligence and expert systems.* Interfaces, 1991. **21**(2): pp. 25-38.
30. *Biased and non-biased neurons.* [accessed 02/11/2012]; Available from: http://galaxy.agh.edu.pl/~vlsi/AI/bias/bias_eng.html.
31. Palmer, A., J. José Montaño, and A. Sesé, *Designing an artificial neural network for forecasting tourism time series.* Tourism Management, 2006. **27**(5): pp. 781-790.

32. Yonaba, H., F. Anctil, and V. Fortin, *Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting.* Journal of Hydrologic Engineering, 2010. **15**(4): pp. 275-283.

33. Harrington, P.d.B., *Sigmoid transfer functions in backpropagation neural networks.* Analytical Chemistry, 1993. **65**(15): pp. 2167-2168.

34. Hassoun, M.H., *Fundamentals of Artificial Neural Networks*. 1995, Cambridge, Mass: MIT Press, pp.197-275.

35. Attoh-Okine, N.O., I.A. Basheer, and D.H. Chen, *Use of Artificial Neural Networks in Geomechanical and Pavement Systems*, 1999, Transportation Research Board, National Research Council: Washington. p. 8.

36. Hopfield, J.J., *Neurons with graded response have collective computational properties like those of 2-state neurons.* Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences, 1984. **81**(10): pp. 3088-3092.

37. Hopfield, J.J. and D.W. Tank, *Computing with neural circuits - a model.* Science, 1986. **233**(4764): pp. 625-633.

38. Kohonen, T., *The self-organizing map.* Neurocomputing, 1998. **21**(1-3): pp. 1-6.

39. Gatti, C.J. and M.J. Embrechts, *Reinforcement learning with neural networks: tricks of the trade*, Advances in Intelligent Signal Processing and Data Mining, P. Georgieva, L. Mihaylova, and L.C. Jain, Editors. 2013, Springer Berlin Heidelberg, pp. 275-310.

40. Rumelhart, D.E., G.E. Hinton, and R.J. Williams, *Learning internal representation by error propagation*, Parallel Distributed Processing: Exploration in the Microstructure of Cognition, D.E. Rumelhart and J.L. Mcclleland, Editors. 1986, MIT Press, pp. 318-362.

41. Velten, K., R. Reinicke, and K. Friedrich, *Wear volume prediction with artificial neural networks.* Tribology International, 2000. **33**(10): pp. 731-736.

42. Jones, S.P., R. Jansen, and R.L. Fusaro, *Preliminary investigation of neural network techniques to predict tribological properties.* Tribology Transactions, 1997. **40**(2): pp. 312-320.

43. Paliwal, M. and U.A. Kumar, *Neural networks and statistical techniques: a review of applications.* Expert Systems with Applications, 2009. **36**(1): pp. 2-17.

44. McCulloch, W.S. and W. Pitts, *A logical calculus of the ideas immanent in nervous activity.* Bulletin of Mathematical Biology, 1990. **52**(1-2): pp. 99-115.

45. Rosenblatt, F., *The perceptron: a probabilistic model for information storage and organization in the brain.* Psychological review, 1958. **65**(6): pp. 386-408.

46. Selfridge, O.G. and U. Neisser, *Pattern recognition by machine.* Scientific American, 1960. **203**(2): pp. 60-68.

47. Widrow, B. and E. Hoff, *Adaptive switching circuits*. IRE WESCON Convention Record. 1960, New York: IRE, pp.96-104.

48. Minsky, M.L. and S. Papert, *Perceptrons: An Introduction to Computational Geometry*. 1969, Cambridge, Mass: MIT Press, pp.1-258.

49. Rumelhart, D.E., G.E. Hinton, and R.J. Williams, *Learning representations by back-propagating errors.* Nature, 1986. **323**(6088): pp. 533-536.

50. Carpenter, G.A. and S. Grossberg, *A massively parallel architecture for a self-organizing neural pattern-recognition machine.* Computer Vision Graphics and Image Processing, 1987. **37**(1): pp. 54-115.

51. Klopf, A.H., *A drive-reinforcement model of single neuron function: An alternative to the Hebbian neuronal model.* AIP Conference Proceedings, 1986. **151**(1): pp. 265-270.

52. Sumathi, S. and P. Surekha, *Computational Intelligence paradigms: Theory & Applications Using Matlab*. 2010: CRC Press, pp.29-32.

53. Jain, A.K., J.C. Mao, and K.M. Mohiuddin, *Artificial neural networks: a tutorial.* Computer, 1996. **29**(3): pp. 31-44.

54. Bishop, C.M., *Neural networks and their applications.* Review of Scientific Instruments, 1994. **65**(6): pp. 1803-1832.

55. Burns, J.A. and G.M. Whitesides, *Feedforward neural networks in chemistry - mathematical systems for classification and pattern recognition.* Chemical Reviews, 1993. **93**(8): pp. 2583-2601.

56. Lippmann, R.P., *Review of neural networks for speech recognition.* Neural Computation, 1989. **1**(1): pp. 1-38.

57. Ruiz, M.E. and P. Srinivasan, *Hierarchical text categorization using neural networks.* Information Retrieval, 2002. **5**(1): pp. 87-118.

58. Er, M.J., S.Q. Wu, J.W. Lu, and H.L. Toh, *Face recognition with radial basis function (RBF) neural networks.* IEEE Transactions on Neural Networks, 2002. **13**(3): pp. 697-710.

59. Rowley, H.A., S. Baluja, and T. Kanade, *Neural network-based face detection.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998. **20**(1): pp. 23-38.

60. Egmont-Petersen, M., D.d. Ridder, and H. Handels, *Image processing with neural networks - a review.* Pattern Recognition, 2002. **35**(10): pp. 2279-2301.

61. Hunt, K.J., D. Sbarbaro, R. Żbikowski, and P.J. Gawthrop, *Neural networks for control systems: a survey.* Automatica, 1992. **28**(6): pp. 1083-1112.

62. Hagan, M.T., H.B. Demuth, and O. De Jesus, *An introduction to the use of neural networks in control systems.* International Journal of Robust and Nonlinear Control, 2002. **12**(11): pp. 959-985.

63. Harris, T.J., R.R. Stroud, and K. Taylor-Burge, *Neural networks in a weld control system.* Proceedings of the IASTED International Symposium Artificial Intelligence Application and Neural Networks, 1990: pp. 172-175.

64. Larsen, G.A., S. Cetinkunt, and A. Donmez, *CMAC neural network control for high precision motion control in the presence of large friction.* Journal of Dynamic Systems Measurement and Control, 1995. **117**(3): pp. 415-420.

65. Forcellese, A., F. Gabrielli, and R. Ruffini, *Effect of the training set size on springback control by neural network in an air bending process.* Journal of Materials Processing Technology, 1998. **80**(1): pp. 493-500.

66. Stich, T.J., J.K. Spoerre, and T. Velasco, *The application of artificial neural networks to monitoring and control of an induction hardening process.* Journal of Industrial Technology, 2000. **16**(1): pp. 1-11.

67. Wang, X. and D. Wang, *Optimizing control of laser surface strengthening parameters for processing 20CrMo steel based on BP neural network.* Infrared Laser Engineering, 2004. **33**(3): pp. 269-287.

68. Wang, D.H. and W.H. Liao, *Modeling and control of magnetorheological fluid dampers using neural networks.* Smart Materials & Structures, 2005. **14**(1): pp. 111-126.

69. Karayel, D., *Prediction and control of surface roughness in CNC lathe using artificial neural network.* Journal of Materials Processing Technology, 2009. **209**(7): pp. 3125-3137.

70. Chansanroj, K., J. Petrovic, S. Ibric, and G. Betz, *Drug release control and system understanding of sucrose esters matrix tablets by artificial neural networks.* European Journal of Pharmaceutical Sciences, 2011. **44**(3): pp. 321-331.

71. Kou, X. and L. Gu, *Application of PID control based on RBF neural network in electro-hydraulic servo system for steel strip deviation.* Advanced Materials Research, 2012. **468-471**: pp. 434-438.

72. Rodemerck, U., M. Baerns, M. Holena, and D. Wolf, *Application of a genetic algorithm and a neural network for the discovery and optimization of new solid catalytic materials.* Applied Surface Science, 2004. **223**(1-3): pp. 168-174.

73. Song, R.G. and Q.Z. Zhang, *Heat treatment technique optimization for 7175 aluminum alloy by an artificial neural network and a genetic algorithm.* Journal of Materials Processing Technology, 2001. **117**(1-2): pp. 84-88.

74. Ootao, Y., Y. Tanigawa, and T. Nakamura, *Optimization of material composition of FGM hollow circular cylinder under thermal loading: a neural network approach.* Composites Part B-Engineering, 1999. **30**(4): pp. 415-422.

75. Somashekhar, K.P., N. Ramachandran, and J. Mathew, *Optimization of material removal rate in Micro-EDM using artificial neural network and genetic algorithms.* Materials and Manufacturing Processes, 2010. **25**(6): pp. 467-475.

76. Sun, Y., W. Zeng, X. Ma, B. Xu, X. Liang, and J. Zhang, *A hybrid approach for processing parameters optimization of Ti-22Al-25Nb alloy during hot deformation using artificial neural network and genetic algorithm.* Intermetallics, 2011. **19**(7): pp. 1014-1019.

77. Elsayed, K. and C. Lacor, *Modeling and Pareto optimization of gas cyclone separator performance using RBF type artificial neural networks and genetic algorithms.* Powder Technology, 2012. **217**: pp. 84-99.

78. Bhadeshia, H.K.D.H., R.C. Dimitriu, S. Forsik, J.H. Pak, and J.H. Ryu, *Performance of neural networks in materials science.* Materials Science and Technology, 2009. **25**(4): pp. 504-510.

79. Reed, P.A.S., M.J. Starink, S.R. Gunn, and I. Sinclair, *Adaptive numerical modelling and hybrid physically based ANM approaches in materials engineering - a survey.* Materials Science and Technology, 2009. **25**(4): pp. 488-503.

80. Qian, H.C., B.C. Xia, S.Z. Li, and F.G. Wang, *Fuzzy neural network modeling of material properties.* Journal of Materials Processing Technology, 2002. **122**(2-3): pp. 196-200.

81. Zhang, Z. and K. Friedrich, *Artificial neural networks applied to polymer composites: a review.* Composites Science and Technology, 2003. **63**(14): pp. 2029-2044.

82. Abendroth, M. and M. Kuna, *Determination of deformation and failure properties of ductile materials by means of the small punch test and neural networks.* Computational Materials Science, 2003. **28**(3-4): pp. 633-644.

83. Scott, D.J., P.V. Coveney, J.A. Kilner, J.C.H. Rossiny, and N.M.N. Alford, *Prediction of the functional properties of ceramic materials from composition using artificial neural networks.* Journal of the European Ceramic Society, 2007. **27**(16): pp. 4425-4435.

84. Wen, M., Y. Yue, H. Zhang, and Y. Li, *Prediction of heating processing-properties of Ti2AlC/TiAl compound materials based o BP neural network.* Advanced Materials Research, 2011. **306-307**: pp. 823-826.

85. Zhang, Z., K. Friedrich, and K. Velten, *Prediction on tribological properties of short fibre composites using artificial neural networks.* Wear, 2002. **252**(7–8): pp. 668-675.

86. Genel, K., S.C. Kurnaz, and M. Durman, *Modeling of tribological properties of alumina fiber reinforced zinc–aluminum composites using artificial neural network.* Materials Science and Engineering: A, 2003. **363**(1–2): pp. 203-210.

87. Jiang, Z., Z. Zhang, and K. Friedrich, *Prediction on wear properties of polymer composites with artificial neural networks.* Composites Science and Technology, 2007. **67**(2): pp. 168-176.

88. Hayajneh, M., A.M. Hassan, A. Alrashdan, and A.T. Mayyas, *Prediction of tribological behavior of aluminum–copper based composite using artificial neural network.* Journal of Alloys and Compounds, 2009. **470**(1–2): pp. 584-588.

89. Aleksendric, D., *Neural network prediction of brake friction materials wear.* Wear, 2010. **268**(1-2): pp. 117-125.

90. Gyurova, L.A. and K. Friedrich, *Artificial neural networks for predicting sliding friction and wear properties of polyphenylene sulfide composites.* Tribology International, 2011. **44**(5): pp. 603-609.

91. Younesi, M., M.E. Bahrololoom, and M. Ahmadzadeh, *Prediction of wear behaviors of nickel free stainless steel-hydroxyapatite bio-composites using*

artificial neural network. Computational Materials Science, 2010. **47**(3): pp. 645-654.

92. Han, J.H. and Q.S. Wu, *Artificial neural network for predicting wear properties of brake lining materials,* Mechatronics and Materials Processing I. 2011, Trans Tech Publications Ltd: Stafa-Zurich, pp. 237-240.

93. Luo, X., S.J. Wang, and X.R. Lv, *Prediction of wear volumes to 45 steel of NBR based on neural network,* Mechanical, Materials and Manufacturing Engineering. 2011, Trans Tech Publications Ltd: Stafa-Zurich, pp. 788-792.

94. Kumar, K.R., K.M. Mohanasundaram, G. Arumaikkannu, and R. Subramanian, *Artificial neural networks based prediction of wear and frictional behaviour of aluminium (A380) / fly ash composites.* Tribology, 2012. **6**(1): pp. 15-19.

95. Mahmoud, T.S., *Artificial neural network prediction of the wear rate of powder metallurgy Al/Al_2O_3 metal matrix composites.* Journal of Materials: Design and Applications, 2012. **226**(L1): pp. 3-15.

96. Nirmal, U., *Prediction of friction coefficient of treated betelnut fibre reinforced polyester (T-BFRP) composite using artificial neural networks.* Tribology International, 2010. **43**(8): pp. 1417-1429.

97. Homer, J., S.C. Generalis, and J.H. Robson, *Artificial neural networks for the prediction of liquid viscosity, density, heat of vaporization, boiling point and Pitzer's acentric factor - Part I. Hydrocarbons.* Physical Chemistry Chemical Physics, 1999. **1**(17): pp. 4075-4081.

98. Zhang, Y., J.R.G. Evans, and S. Yang, *The prediction of solid solubility of alloys: developments and applications of Hume-Rothery's rules.* The Journal of Crystallization Physics and Chemistry, 2010. **1**(2): pp. 103-119.

99. Gebhardt, T., D. Music, T. Takahashi, and J.M. Schneider, *Combinatorial thin film materials science: from alloy discovery and optimization to alloy design.* Thin Solid Films, 2012. **520**(17): pp. 5491-5499.

100. Hanak, J.J., *Multiple-sample-concept in materials research : Synthesis, compositional analysis and testing of entire multicomponent systems.* Journal of Materials Science, 1970. **5**(11): pp. 964-971.

101. Kennedy, K., T. Stefansky, G. Davy, V.F. Zackay, and E.R. Parker, *Rapid method for determining ternary-alloy phase diagrams.* Journal of Applied Physics, 1965. **36**(12): pp. 3808-3810.

102. Woo, S.I., K.W. Kim, H.Y. Cho, K.S. Oh, M.K. Jeon, N.H. Tarte, T.S. Kim, and A. Mahmood, *Current status of combinatorial and high-throughput methods for discovering new materials and catalysts.* Qsar & Combinatorial Science, 2005. **24**(1): pp. 138-154.

103. Xiang, X.D., X.D. Sun, G. Briceno, Y.L. Lou, K.A. Wang, H.Y. Chang, W.G. Wallacefreedman, S.W. Chen, and P.G. Schultz, *A combinatorial approach to materials discovery.* Science, 1995. **268**(5218): pp. 1738-1740.

104. Amis, E.J., X.D. Xiang, and J.C. Zhao, *Combinatorial materials science: what's new since Edison?* MRS Bulletin, 2002. **27**(4): pp. 295-297.

105. Barber, Z.H. and M.G. Blamire, *High throughput thin film materials science.* Materials Science and Technology, 2008. **24**(7): pp. 757-770.

106. Xiang, X.D., *Combinatorial materials synthesis and screening: An integrated materials chip approach to discovery and optimization of functional materials.* Annual Review of Materials Science, 1999. **29**: pp. 149-171.

107. Chen, Y., J.R.G. Evans, and S. Yang, *A rapid-doping method for high-throughput discovery applied to thick-film PTCR materials.* Journal of the American Ceramic Society, 2011. **94**(11): pp. 3748-3756.

108. May, R., G. Dandy, and H. Maier, *Review of input variable selection methods for artificial neural networks*, Artificial Neural Networks - Methodological Advances and Biomedical Applications, K. Suzuki, Editor. 2011, InTech, Rijeka, Croatia: pp. 2(1-27).

109. Bassetti, D., Y. Brechet, and M.F. Ashby, *Estimates for material properties. II. The method of multiple correlations.* Proceedings of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences, 1998. **454**(1973): pp. 1323-1336.

110. Wang, R.Y., V.C. Storey, and C.P. Firth, *A framework for analysis of data quality research.* IEEE Transactions on Knowledge and Data Engineering, 1995. **7**(4): pp. 623-640.

111. Strauss, H.J., *Handbook for Chemical Technicians*. 1976: New York : McGraw-Hill, pp.1(52).

112. Lange, N.A., *Lange's Handbook of Chemistry*. 16th Edn. 2005: New York : McGraw-Hill, pp.vii-xii, 1(280-298).

113. Buch, A., *Correlation between the specific heat capacity and the elastic moduli of pure metals.* Zeitschrift Fur Metallkunde, 1999. **90**(9): pp. 744-746.

114. Buch, A., *Relations between the elastic moduli of pure metals - Used for evaluating the accuracy of the values of the moduli.* Materialprufung, 2001. **43**(11-12): pp. 452-455.

115. Swamy, K.M. and K.L. Narayana, *Elastic and acoustic properties of isotropic polycrystalline metals.* Acustica, 1983. **54**(2): pp. 123-125.

116. Starink, M., X. Cheng, and S. Yang, *Hardening of pure metals by high-pressure torsion: a physically based model employing volume-averaged defect evolutions.* Acta Materialia, 2013. **61**(1): pp. 183-192.

117. Buch, A., *The relation between Young's modulus and other physical parameters.* Materialprufung, 2007. **49**(10): pp. 542-548.

118. Siethoff, H. and K. Ahlborn, *Debye temperature-elastic constants relationship for materials with hexagonal and tetragonal symmetry.* Journal of Applied Physics, 1996. **79**(6): pp. 2968-2974.

119. Scott, T.E., *Elastic and mechanical properties*, Handbook on the Physics and Chemistry of Rare Earths. vol. 1 Metals, K.A. Gschneidner, Jr. and L. Eyring, Editors. 1978, pp. 591-705.

120. Pollock, J.J. and A. Zamora, *Automatic spelling correction in scientific and scholarly text.* Communications of the Acm, 1984. **27**(4): pp. 358-368.

121. Hodge, V.J. and J. Austin, *A survey of outlier detection methodologies.* Artificial Intelligence Review, 2004. **22**(2): pp. 85-126.

122. Pfeifer, U., T. Poersch, and N. Fuhr, *Retrieval effectiveness of proper name search methods.* Information Processing & Management, 1996. **32**(6): pp. 667-679.

123. Walczak, B. and D.L. Massart, *Multiple outlier detection revisited.* Chemometrics and Intelligent Laboratory Systems, 1998. **41**(1): pp. 1-15.

124. Yao, J.T., *Knowledge extracted from trained neural networks - What's next?*, Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security 2005, B.V. Dasarathy, Editor. 2005, pp. 151-157.

125. Yao, X., *A review of evolutionary artificial neural networks.* International Journal of Intelligent Systems, 1993. **8**(4): pp. 539-567.

126. Yao, X., *Evolving artificial neural networks.* Proceedings of the IEEE, 1999. **87**(9): pp. 1423-1447.

127. Armentrout, M.M. and A. Kavner, *Incompressibility of osmium metal at ultrahigh pressures and temperatures.* Journal of Applied Physics, 2010. **107**(9): pp. 093528(1-6).

128. Ashida, Y., M. Yamamoto, S. Naito, M. Mabuchi, and T. Hashino, *Measurements of Young's modulus and the modulus of rigidity of the solid solution of hydrogen in zirconium between 300 and 1300 K.* Journal of Applied Physics, 1996. **80**(6): pp. 3254-3258.

129. Greiner, J.D., B.J. Beaudry, and J.F. Smith, *Effect of small amounts of hydrogen on the elastic behavior of lutetium.* Journal of Applied Physics, 1987. **62**(4): pp. 1220-1223.

130. Farraro, R. and R. McLellan, *Temperature dependence of the Young's modulus and shear modulus of pure nickel, platinum, and molybdenum.* Metallurgical and Materials Transactions A, 1977. **8**(10): pp. 1563-1565.

131. MacFarlane, R.E., J.A. Rayne, and C.K. Jones, *Temperature dependence of elastic moduli of iridium.* Physics Letters, 1966. **20**(3): pp. 234-235.

132. Ho, P.S., J.P. Poirier, and A.L. Ruoff, *Pressure dependence of the elastic constants for silver from 77 K to 300 K.* Physica Status Solidi (b), 1969. **35**(2): pp. 1017-1025.

133. Soma, T. and E. Ueda, *Pressure dependence of the elastic stiffness constants of alkali metals.* Journal of Physics F: Metal Physics, 1982. **12**(1): p. 59.

134. Jain, A.L., *Pressure dependence of the elastic shear constants of Li.* Physical Review, 1961. **123**(4): pp. 1234-1238.

135. Guinan, M.W. and Steinber.Dj, *Pressure and temperature derivatives of isotropic polycrystalline shear modulus for 65 elements.* Journal of Physics and Chemistry of Solids, 1974. **35**(11): pp. 1501-1512.

136. Darling, A.S., *The elastic and plastic properties of the platinum metals.* Platinum Metals Review, 1966. **10**(1): pp. 14-19.

137. Sisodia, P. and M.P. Verma, *Polycrystalline elastic moduli of some hexagonal and tetragonal materials.* Physica Status Solidi (a), 1990. **122**(2): pp. 525-534.

138. Mathur, S.S. and Y.P. Sharma, *Third order elastic constants of body centered cubic crystals.* Physica Status Solidi (b), 1970. **39**(1): pp. 39-42.

139. Barsch, G.R., *Adiabatic, isothermal, and intermediate pressure derivatives of the elastic constants for cubic symmetry. I. Basic formulae.* Physica Status Solidi (b), 1967. **19**(1): pp. 129-138.

140. Köster, W., *Die Querkontraktionszahl im periodischen System.* Zeitschrift für Elektrochemie und angewandte physikalische Chemie, 1943. **49**(4-5): pp. 233-237.

141. Köster, W., *Die Temperaturabhangigheit des Elastizitatsmoduls reiner Metall.* Z. Metallkunde, 1948. **39**(1): pp. 1-9.

142. Köster, W. and H. Franz, *Poisson's ratio for metals and alloys.* Metallurgical Reviews, 1961. **6**(21): pp. 1-55.

143. Gale, W.F. and T.C. Totemeier, *Smithells Metals Reference Book*. 2004: Elsevier Butterworth-Heinemann, pp.15(1-8).

144. Pugh, S.F., *Relations between the elastic moduli and the plastic properties of polycrystalline pure metals.* Philosophical Magazine, 1954. **45**(367): pp. 823-843.

145. Hearn, E.J., *An introduction to the mechanics of elastic and plastic deformation of solids and structural materials*. 3rd Edn. Mechanics of Materials. Vol. 1. 1997: Elsevier, pp.3-12, 367-370.

146. Haussühl, S., *Physical Properties of Crystals: An Introduction*. 2008: John Wiley & Sons, pp.214-217.

147. Subramani, V.J., *Atomistic simulations for the prediction of physical properties of cement-based materials*, 2008, University of Arkansas: Fayetteville. p. 58.

148. Nye, J.F., *Physical Properties of Crystals: Their Representation by Tensors and Matrices*. 1985: Oxford University Press, USA, pp.123-125.

149. Rees, D.W.A., *Mechanics of Solids and Structures.* 2000: Imperial College Press, pp.43-48.

150. Singh, A.K., *Mechanics of Solids*. 2007: Prentice-Hall of India Pvt.Ltd, pp.181-192.

151. Mott, P.H. and C.M. Roland, *Limits to Poisson's ratio in isotropic materials—general result for arbitrary deformation.* Physica Scripta, 2013. **87**(5): pp. 055404(1-15).

152. Gschneidner Jr, K.A., *Physical properties and interrelationships of metallic and semimetallic elements*, Solid State Physics, S. Frederick and T. David, Editors. 1964, Academic Press, pp. 275-426.

153. Gere, J.M., *Mechanics of Materials*. 2002: Nelson Thornes, pp.23-26.

154. Ledbetter, H.M., *Ratio of shear and Young's moduli for polycrystalline metallic elements.* Materials Science and Engineering, 1977. **27**(2): pp. 133-136.

155. Brinson, H.F. and L.C. Brinson, *Stress and strain analysis and measurement*, Polymer Engineering Science and Viscoelasticity: An Introduction. 2008, Springer, p. 19.

156. Hill, R., *The elastic behaviour of a crystalline aggregate.* Proceedings of the Physical Society, 1952. **65**(5): p. 349.

157. Man, C.-S. and M. Huang, *A simple explicit formula for the Voigt-Reuss-Hill average of elastic polycrystals with arbitrary crystal and texture symmetries.* Journal of Elasticity, 2012. **106**(1): pp. 105-105.

158. Nadal, M.-H., C. Hubert, and R. Oltra, *High temperature shear modulus determination using a laser-ultrasonic surface acoustic-wave device.* Journal of Applied Physics, 2009. **106**(2): pp. 024906(1-6).

159. Simmons, G., *Single Crystal Elastic Constants and Calculated Aggregate Properties: A Handbook*. 2nd Edn. 1971, Cambridge, Mass.: M.I.T. Press, pp.3-36.

160. Thomsen, L., *Elasticity of polycrystals and rocks.* Journal of Geophysical Research, 1972. **77**(2): pp. 315-327.

161. Hashin, Z. and S. Shtrikman, *On some variational principles in anisotropic and nonhomogeneous elasticity*. Journal of the Mechanics and Physics of Solids, 1962. **10**(4): pp. 335-342.

162. Meister, R. and L. Peselnick, *Variational method of determining effective moduli of polycrystals with tetragonal symmetry*. Journal of Applied Physics, 1966. **37**(11): pp. 4121-4125.

163. Conry, M. *Notes on Wave Propagation in Anisotropic Elastic Solids*. [accessed 05/09/2013] http://www.acronymchile.com/anisotropic_with_lamb_waves.pdf.

164. Pagnotta, L., *Recent progress in identification methods for the elastic characterization of materials.* International Journal of Mechanics, 2008. **2**(4): pp. 129-140.

165. Smith, J.S., M.D. Wyrick, and J.M. Poole, *An evaluation of three techniques for determining the Young's modulus of mechanically alloyed materials*, Dynamic elastic modulus measurements in Materials, A. Wolfenden, Editor 1990, American society for testing and materials. p. 195.

166. Birfel'd, A.A., B.M. Drapkin, and V.V. Potkin, *Experimental study of Young's modulus of glass-plastics.* Mechanics of Composite Materials, 1975. **11**(6): pp. 931-933.

167. *ASTM Standards*, American Society for Testing and Materials.

168. EN, *European Standard*.

169. ISO, *International Organization for Standardization*: ISO Central Secretariat 1, ch. de la Voie-Creuse CP 56 CH-1211 Geneva 20 Switzerland.

170. NPL, *National Physical Laboratory*: Queens Road, Teddington, TW11 0LW, UK.

171. Ashby, M., H. Shercliff, and D. Cebon, *Materials Engineering, Science, Processing and Design*. 2007: Elsevier Ltd, pp.48-55.

172. Callister, W.D., *Materials Science and Engineering: An Introduction*. 2007: John Wiley & Sons, pp.133-146.

173. Beer, F., J.E.R. Johnston, and J. DeWolf, *Mechanics of Materials (in SI Units)*. 4th Edn. 2006, Singapore: McGraw-Hill Education, pp.46-51.

174. Lord, J.D. and R. Morrell, *Measurement Good Practice Guide: Elastic Modulus Measurement*, 2007, National Physical Laboratory. p. 5.

175. Blessing, G.V., *The pulsed ultrasonic velocity method for determining material dynamic elastic moduli*. Dynamic Elastic Modulus Measurements in Materials, Edn. A. Wolfenden. 1990: ASTM, pp.47-57.

176. Szilard, J., *Ultrasonic Testing: Non-conventional Testing Techniques*. 1982: Wiley, pp.2-7.

177. Franco-Villafane, J.A., E. Flores-Olmedo, G. Báez, O. Gandarilla-Carrillo, and R.A. Méndez-Sánchez, *Acoustic resonance spectroscopy for the advanced undergraduate laboratory.* European Journal of Physics, 2012. **33**(6): p. 1761.

178. *ASTM E1875-08: Standard Test Method for Dynamic Young's Modulus, Shear Modulus, and Poisson's Ratio by Sonic Resonance*, [accessed 13/10/2013]. Available from : http://www.astm.org/Standards/E1875.htm.

179. Basheer, I.A. and M. Hajmeer, *Artificial neural networks: fundamentals, computing, design, and application.* Journal of Microbiological Methods, 2000. **43**(1): pp. 3-31.

180. Mackworth, A., *Constraint satisfaction*, Encyclopedia of Artificial Intelligence. 1992, pp. 285-293.

181. Clarke, B., E. Fokoué, and H.H. Zhang, *Principles and theory for data mining and machine learning*. Springer Series in Statistics. 2009: Springer, pp.32-38.

182. Horita, Z. and T.G. Langdon, *Microstructures and microhardness of an aluminum alloy and pure copper after processing by high-pressure torsion.* Materials Science and Engineering A - Structural Materials Properties Microstructure and Processing, 2005. **410**: pp. 422-425.

183. Edalati, K. and Z. Horita, *Universal plot for hardness variation in pure metals processed by high-pressure torsion.* Materials Transactions, 2010. **51**(5): pp. 1051-1054.

184. Edalati, K. and Z. Horita, *Correlations between hardness and atomic bond parameters of pure metals and semi-metals after processing by high-pressure torsion.* Scripta Materialia, 2011. **64**(2): pp. 161-164.

185. Edalati, K. and Z. Horita, *High-pressure torsion of pure metals: influence of atomic bond parameters and stacking fault energy on grain size and correlation with hardness.* Acta Materialia, 2011. **59**(17): pp. 6831-6836.

186. Edalati, K. and Z. Horita, *Significance of homologous temperature in softening behavior and grain size of pure metals processed by high-pressure torsion.*

Materials Science and Engineering A - Structural Materials Properties Microstructure and Processing, 2011. **528**(25-26): pp. 7514-7523.

187. Edalati, K. and Z. Horita, *Correlation of physical parameters with steady-state hardness of pure metals processed by high-pressure torsion.* Materials Science Forum, 2011. **667-669**: pp. 683-688.

188. Edalati, K. and Z. Horita, *Continuous high-pressure torsion.* Journal of Materials Science, 2010. **45**(17): pp. 4578-4582.

189. Kazeminezhad, M. and E. Hosseini, *Modeling of induced empirical constitutive relations on materials with fcc, bcc, and hcp crystalline structures: severe plastic deformation.* International Journal of Advanced Manufacturing Technology, 2010. **47**(9-12): pp. 1033-1039.

190. Weckman, G.R., H.W. Paschold, J.D. Dowler, H.S. Whiting, and W.A. Young, *Using neural networks with limited data to estimate manufacturing cost.* Journal of Industrial and Systems Engineering, 2010. **3**(4): pp. 257-274.

191. Zhilyaev, A.P. and T.G. Langdon, *Using high-pressure torsion for metal processing: Fundamentals and applications.* Progress in Materials Science, 2008. **53**(6): pp. 893-979.

192. Giustolisi, O. and D.A. Savic, *A symbolic data-driven technique based on evolutionary polynomial regression.* Journal of Hydroinformatics, 2006. **8**(3): pp. 207-222.

193. Zhang, G., B. Eddy Patuwo, and M. Y. Hu, *Forecasting with artificial neural networks: the state of the art.* International Journal of Forecasting, 1998. **14**(1): pp. 35-62.

194. Breiman, L., *Random forests.* Machine Learning, 2001. **45**(1): pp. 5-32.

195. Dietterich, T.G., *An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization.* Machine Learning, 2000. **40**(2): pp. 139-157.

196. Brown, M.P.S., W.N. Grundy, D. Lin, N. Cristianini, C.W. Sugnet, T.S. Furey, M. Ares, and D. Haussler, *Knowledge-based analysis of microarray gene expression data by using support vector machines.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(1): pp. 262-267.

197. Koza, J.R., *Genetic programming as a means for programming computers by natural-selection.* Statistics and Computing, 1994. **4**(2): pp. 87-112.

198. Kingston, G.B., H.R. Maier, and M.F. Lambert, *A probabilistic method for assisting knowledge extraction from artificial neural networks used for hydrological prediction.* Mathematical and Computer Modelling, 2006. **44**(5-6): pp. 499-512.

199. Davidson, J.W., D.A. Savic, and G.A. Walters, *Symbolic and numerical regression: experiments and applications.* Information Sciences, 2003. **150**(1-2): pp. 95-117.

200. Brezocnik, M. and L. Gusel, *Predicting stress distribution in cold-formed material with genetic programming.* International Journal of Advanced Manufacturing Technology, 2004. **23**(7-8): pp. 467-474.

201. McKay, B., M. Willis, and G. Barton, *Steady-state modelling of chemical process systems using genetic programming.* Computers & Chemical Engineering, 1997. **21**(9): pp. 981-996.

202. Cai, W., A. Pacheco-Vega, M. Sen, and K.T. Yang, *Heat transfer correlations by symbolic regression.* International Journal of Heat and Mass Transfer, 2006. **49**(23-24): pp. 4352-4359.

203. Savic, D.A., G.A. Walters, and J.W. Davidson, *A genetic programming approach to rainfall-runoff modelling.* Water Resources Management, 1999. **13**(3): pp. 219-231.

204. Giustolisi, O., *Using genetic programming to determine Chezy resistance coefficient in corrugated channels.* Journal of Hydroinformatics, 2004. **6**(3): pp. 157-173.

205. Sharifi, S., M. Sterling, and D.W. Knight, *Prediction of end-depth ratio in open channels using genetic programming.* Journal of Hydroinformatics, 2011. **13**(1): pp. 36-48.

206. Grosman, B. and D.R. Lewin, *Adaptive genetic programming for steady-state process modeling.* Computers & Chemical Engineering, 2004. **28**(12): pp. 2779-2790.

207. Davidson, J.W., D. Savic, and G.A. Walters, *Method for the identification of explicit polynomial formulae for the friction in turbulent pipe flow*. Journal of Hydroinformatics, 1999. **1**(2): pp. 115-126.

208. Fishtine, S.H., *Reliable latent heats of vaporization.* Journal of Industrial & Engineering Chemistry, 1963. **55**(4): pp. 20-28.

209. Liu, Z., *Estimation of heat of vaporization of pure liquid at its normal boiling temperature*. Chemical Engineering Communications, 2001. **184**(1): pp. 221-228.

210. Sivaraman, A., J.W. Magee, and R. Kobayashi, *Generalized correlation of latent heats of vaporization of coal liquid model compounds between their freezing points and critical-points.* Industrial & Engineering Chemistry Fundamentals, 1984. **23**(1): pp. 97-100.

211. Robert C. Reid, J.M.P., Bruce E.Poling, *The Properties of Gases and Liquids*. 4th Edn. 1987, New York: McGraw-Hill Companies, pp.205-240.

212. Koza, J.R., *Genetic Programming: The Programming of Computers by Means of Natural Selection*. 1992, London: The MIT Press, pp.17-62, 237-288.

213. Affenzeller, M., S. Wagner, S. Winkler, and A. Beham, *Genetic Algorithms and Genetic Programming: Modern Concepts and Practical Applications*. 2009: Taylor & Francis, pp.37-38.

214. Elshorbagy, A., G. Corzo, S. Srinivasulu, and D.P. Solomatine, *Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - part 1: Concepts and methodology*. Hydrol. Earth Syst. Sci., 2010. **14**(10): pp. 1931-1941.

215. McKay, B., M.J. Willis, and G.W. Barton. *On the application of genetic programming to chemical process systems*. Evolutionary Computation, IEEE International Conference. 1995: pp. 701-706.

216. O'Reilly, U.M., T. McConaghy, and R. Riolo, *GPTP 2009: an example of evolvability*. Genetic programming theory and practice VII. 2009, New York: Springer, pp.1-18.

217. Zhang, M.J., *Mining small objects in large images using neural networks*, Foundations of Data Mining and Knowledge Discovery. 2005, New York: Springer, pp. 277-303.

218. Olden, J.D., M.K. Joy, and R.G. Death, *An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data.* Ecological Modelling, 2004. **178**(3–4): pp. 389-397.

219. Paliwal, M. and U.A. Kumar, *Assessing the contribution of variables in feed forward neural network.* Applied Soft Computing, 2011. **11**(4): pp. 3690-3696.

220. Dimopoulos, Y., P. Bourret, and S. Lek, *Use of some sensitivity criteria for choosing networks with good generalization ability.* Neural Processing Letters, 1995. **2**(6): pp. 1-4.

221. Gevrey, M., I. Dimopoulos, and S. Lek, *Review and comparison of methods to study the contribution of variables in artificial neural network models.* Ecological Modelling, 2003. **160**(3): pp. 249-264.

222. Goh, A.T.C., *Backpropagation neural networks for modeling complex-systems.* Artificial Intelligence in Engineering, 1995. **9**(3): pp. 143-151.

223. Masters, T., *Practical Neural Network Recipes in C++*. 1993, Boston: Academic Press, pp.174-176.

224. Heaton, J., *Introduction to Neural Networks for Java*. 2nd Edn. 2008: Heaton Research, pp.157-158.

225. Malinov, S. and W. Sha, *Software products for modelling and simulation in materials science.* Computational Materials Science, 2003. **28**(2): pp. 179-198.

226. *Matlab 2010a*, The MathWorks Inc. [accessed 03/12/2010] Available from http://www.mathworks.co.uk.

227. Beale, M.H., M.T. Hagan, and H.B. Demuth, *Mathworks Neural Network Toolbox User's Guide R2010a*. 2010, pp.2-35.

228. *CES EduPack*, 2011, Granta Design. [accessed 15/08/2012] Available from http://www.grantadesign.com/education/.

229. Lide, D.R., *CRC Handbook of Chemistry and Physics*. 90th Edn. 2009, Florida: CRC Press/Taylor and Francis, pp.4(127-135),6(50-69), 6(101-118), 15(13-22).

230. Prechelt, L., *Automatic early stopping using cross validation: quantifying the criteria.* Neural Networks, 1998. **11**: pp. 761-767.

231. Dan Foresee, F. and M.T. Hagan. *Gauss-Newton approximation to Bayesian learning*. International Conference on Neural Networks. 1997. **3**: pp. 1930-1935.

232. Lampinen, J. and A. Vehtari, *Bayesian approach for neural networks - review and case studies.* Neural Networks, 2001. **14**(3): pp. 257-274.

233. Zhang, Y., J. Evans, and S. Yang, *Exploring correlations between properties using artificial neural networks.To be published.*

234. Chapman, A.D., *Principles and methods of data cleaning : primary species and species-occurrence data*. Global Biodiversity Information. 2005, Copenhagen: Global Biodiversity Information Facility, pp.2.

235. Chrisman, N.R., *The error component in spatial data,* Geographical Information Systems. 1991: Wiley, pp.165-174.

236. Kaptay, G., *A unified model for the cohesive enthalpy, critical temperature, surface tension and volume thermal expansion coefficient of liquid metals of bcc, fcc and hcp crystals.* Materials Science and Engineering A-Structural Materials Properties Microstructure and Processing, 2009. **501**(1-2): pp. 255-255.

237. Zhang, Y., J.R.G. Evans, and S. Yang, *Corrected values for boiling points and enthalpies of vaporization of elements in handbooks.* Journal of Chemical & Engineering Data, 2011. **56**(2): pp. 328-337.

238. Kittel, C., *Introduction to Solid State Physics*. 1996: Wiley, pp.48-50.

239. Tsuru, Y., Y. Shinzato, Y. Saito, M. Shimazu, M. Shiono, and M. Morinaga, *Estimation of linear thermal expansion coefficient from cohesive energy obtained by ab-initio calculation of metals and ceramics.* Journal of the Ceramic Society of Japan, 2010. **118**(1375): pp. 241-245.

240. Qi, W.H., M.P. Wang, and G.Y. Xu, *The particle size dependence of cohesive energy of metallic nanoparticles.* Chemical Physics Letters, 2003. **372**(5-6): pp. 632-634.

241. Alymov, M.I. and M.K. Shorshorov, *Surface tension of ultrafine particles.* Nanostructured Materials, 1999. **12**(1-4): pp. 365-368.

242. Buffat, P. and J.P. Borel, *Size effect on melting temperature of gold particles.* Physical Review A, 1976. **13**(6): pp. 2287-2298.

243. Kaptay, G., G. Csicsovszki, and M.S. Yaghmaee, *An absolute scale for the cohesion energy of pure metals,* Materials Science Forum, 2003. **414-415**: pp. 235-240.

244. Brewer, L., *Cohesive Energies of The Elements*, 1975, Lawrence Berkeley Laboratory. Report No. LBL-3720: pp. 1-18.

245. Farid, B. and R.W. Godby, *Cohesive energies of crystals.* Physical Review B, 1991. **43**(17): pp. 14248-14250.

246. Tamura, S., *Empirical correlations between the compressibility and the atomic radius or the cohesive energy for elements.* Journal of Materials Science Letters, 1999. **18**(21): pp. 1753-1754.

247. Gangopadhyay, A.K., J.C. Bendert, N.A. Mauro, and K.F. Kelton, *Inverse correlation between cohesive energy and thermal expansion coefficient in liquid transition metal alloys.* Journal of Physics-Condensed Matter, 2012. **24**(37).

248. Hogan, W.J., A.J.C. Ladd, N.J. Hoffman, and R.R. Peterson, *Condensation of vaporized cascade blanket material*, Laser Program Annual Report. 1986, Lawrence Livermore National Laboratory., p. 8.

249. Hultgren, R., *Selected values of the thermodynamic properties of the elements* 1973: Metals Park, Ohio : American Society for Metals, pp.443-451.

250. Stark, J.G. and H.G. Wallace, *Chemistry Data Book*. 2nd Edn. 1982, London: Murray, pp.50-51.

251. Emsley, J., *The Elements*. 3rd End, Oxford Chemistry Guides. Vol. 2. 1998: Clarendon Press,

252. Lange, N.A. and G.M. Forker, *Lange's Handbook of Chemistry*. 10th Edn. 1967: McGraw-Hill, pp.1(110-111).

253. Sinke, G.C., *Sources and discussion of the data*, Thermodynamic properties of the elements, D.R. Stull and G.C. Sinke, Editors. 1956, American chemical society, pp. 10-35.

254. Wieser, M.E., N. Holden, T.B. Coplen, J.K. Böhlke, M. Berglund, W.A. Brand, P.D. Bièvre, M. Gröning, R.D. Loss, J. Meija, T. Hirata, T. Prohaska, R. Schoenberg, G. O'Connor, T. Walczyk, S. Yoneda, and X.-K. Zhu, *Atomic weights of the elements 2011 (IUPAC Technical Report).* Pure and Applied Chemistry, 2013. **85**(5): pp. 1047-1078.

255. Dulong, P.L. and A.T. Petit, *Recherches sur quelques points importants de la Théorie de la Chaleur.* Annales de Chimie et de Physique, 1819. **10**: pp. 395-413. Research on some important points of the Theory of Heat, contemporary translation from Annals of Philosophy **14**: pp. 189-198.

256. Laing, M. and M. Laing, *Dulong and Petit's law: we should not ignore its importance.* Journal of Chemical Education, 2006. **83**(10): pp. 1499-1504.

257. Khodel, V.A., J.W. Clark, V.R. Shaginyan, and M.V. Zverev, *Second wind of the Dulong-Petit Law at at a quantum critical point.* JETP Letters, 2010. **92**(8): pp. 532-536.

258. Olijnyk, H., *Lattice vibrations and electronic transitions in the rare-earth metals: yttrium, gadolinium and lutetium.* Journal of Physics-Condensed Matter, 2005. **17**(1): pp. 43-52.

259. Schwarz, J.S.P., *Heat capacity - atomic weight relationship*, in: American Chemical Society National Meeting. 1986: New York. [accessed 18/08/2013] Available from http://organicchemist.us/organichem/my_pages/d-ptalk.html.

260. Lu, H.M. and Q. Jiang, *Surface tension and its temperature coefficient for liquid metals.* The Journal of Physical Chemistry B, 2005. **109**(32): pp. 15463-15468.

261. Skapski, A.S., *The surface tension of liquid metals.* The Journal of Chemical Physics, 1948. **16**(4): pp. 389-393.

262. Oriani, R.A., *The surface tension of liquid metals and the excess binding energy of surface atoms.* The Journal of Chemical Physics, 1950. **18**(5): pp. 575-578.

263. Skapski, A.S., *The temperature coefficient of the surface tension of liquid metals.* The Journal of Chemical Physics, 1948. **16**(4): pp. 386-389.

264. Overbury, S.H., P.A. Bertrand, and G.A. Somorjai, *Surface composition of binary systems: prediction of surface phase diagrams of solid solutions.* Chemical Reviews, 1975. **75**(5): pp. 547-560.

265. Fukuyama, H. and Y. Waseda, *High-temperature Measurements of Materials.* 2008: Springer, p.47-49.

266. Becher, P., *The calculation of cohesive energy density from the surface tension of liquids.* Journal of Colloid and Interface Science, 1972. **38**(2): pp. 291-293.

267. Remsburg, R., *Advanced Thermal Design of Electronic Equipment*. 1998: Springer, pp.22.

268. Michaelson, H.B., *Work functions of the elements.* Journal of Applied Physics, 1950. **21**(6): pp. 536-540.

269. Michaelson, H.B., *The work function of the elements and its periodicity.* Journal of Applied Physics, 1977. **48**(11): pp. 4729-4733.

270. Hölzl, J., F.K. Schulte, and H. Wagner, *Solid Surface Physics*. 1979: Springer-Verlag, pp.45-67.

271. Dweydari, A.W. and C.H.B. Mee, *Work function measurements on (100) and (110) surfaces of silver.* Physica Status Solidi (a), 1975. **27**(1): pp. 223-230.

272. Xue, M., W. Wang, F. Wang, J. Ou, C. Li, and W. Li, *Understanding of the correlation between work function and surface morphology of metals and alloys.* Journal of Alloys and Compounds, 2013. **577**: pp. 1-5.

273. Ibragimov, K.I. and V.A. Korolkov, *Temperature dependence of the work function of metals and binary alloys.* Inorganic Materials, 2001. **37**(6): pp. 567-572.

274. Gordy, W. and W.J.O. Thomas, *Electronegativities of the elements.* The Journal of Chemical Physics, 1956. **24**(2): pp. 439-444.

275. Kashetov, A. and N.A. Gorbatyi, *Correlation between the work functions and surface energies of the faces of metal crystals.* Soviet Physics Journal, 1969. **12**(7): pp. 860-863.

276. Chatterjee, B., *An anisotropic relation between work function and melting temperature.* Physics Letters A, 1979. **69**(6): pp. 429-430.

277. Schade, H., *Work functions and sublimation entropies of the elements.* Applied Physics, 1979. **18**(4): pp. 339-344.

278. Guomin, H. and L. Dongyang, *Generic relation between the electron work function and Young's modulus of metals.* Applied Physics Letters, 2011. **99**(4): pp. 041907 (1-3).

279. Samsonov, G.V., *Handbook of The Physicochemical Properties of The Elements*. 1968: Oldbourne, pp.387-398.

280. Buchenau, U., M. Heiroth, H.R. Schober, J. Evers, and G. Oehlinger, *Lattice dynamics of strontium and barium.* Physical Review B, 1984. **30**(6): pp. 3502-3505.

281. Vold, C.L., M.E. Glicksman, E.W. Kammer, and L.C. Cardinal, *The elastic constants for single-crystal lead and indium from room temperature to the melting point.* Journal of Physics and Chemistry of Solids, 1977. **38**(2): pp. 157-160.

282. Neighbours, J.R. and G.A. Alers, *Elastic constants of silver and gold.* Physical Review, 1958. **111**(3): pp. 707-712.

283. Kamm, G.N. and G.A. Alers, *Low-temperature elastic moduli of aluminum.* Journal of Applied Physics, 1964. **35**(2): pp. 327-330.

284. Kammer, E.W., L.C. Cardinal, C.L. Vold, and M.E. Glicksman, *The elastic constants for single-crystal bismuth and tin from room temperature to the melting point.* Journal of Physics and Chemistry of Solids, 1972. **33**(10): pp. 1891-1898.

285. McSkimin, H.J., *Measurement of the elastic constants of single crystal cobalt.* Journal of Applied Physics, 1955. **26**(4): pp. 406-409.

286. Bolef, D.I. and J. de Klerk, *Anomalies in the elastic constants and thermal expansion of chromium single crystals.* Physical Review, 1963. **129**(3): pp. 1063-1067.

287. Ledbetter, H.M. and E.R. Naimon, *Elastic properties of metals and alloys. II. Copper.* Journal of Physical and Chemical Reference Data, 1974. **3**(4): pp. 897-935.

288. Palmer, S.B. and E.W. Lee, *The elastic constants of dysprosium and holmium.* Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 1972. **327**(1571): pp. 519-543.

289. Ledbetter, H.M. and R.P. Reed, *Elastic properties of metals and alloys. I. Iron, nickel, and iron-nickel alloys.* Journal of Physical and Chemical Reference Data, 1973. **2**(3): pp. 531-618.

290. Slutsky, L.J. and C.W. Garland, *Elastic constants of magnesium from 4.2 K to 300 K.* Physical Review, 1957. **107**(4): pp. 972-976.

291. Bolef, D.I. and J. De Klerk, *Elastic constants of single crystal Mo and W between 77 K and 500 K.* Journal of Applied Physics, 1962. **33**(7): pp. 2311-2314.

292. Bolef, D.I., *Elastic constants of single crystals of the bcc transition elements V, Nb, and Ta.* Journal of Applied Physics, 1961. **32**(1): pp. 100-105.

293. Rayne, J.A., *Elastic constants of palladium from 4.2 K to 300 K.* Physical Review, 1960. **118**(6): pp. 1545-1549.

294. Collard, S.M. and R.B. McLellan, *High-temperature elastic constants of platinum single crystals.* Acta Metallurgica et Materialia, 1992. **40**(4): pp. 699-702.

295. Featherston, F.H. and J.R. Neighbours, *Elastic constants of tantalum, tungsten, and molybdenum.* Physical Review, 1963. **130**(4): pp. 1324-1333.

296. Palmer, S.B., E.W. Lee, and M.N. Islam, *The elastic constants of gadolinium, terbium and erbium.* Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 1974. **338**(1614): pp. 341-357.

297. Fisher, E.S. and C.J. Renken, *Single-crystal elastic moduli and the hcp → bcc transformation in Ti, Zr, and Hf.* Physical Review, 1964. **135**(2A): pp. A(482-494).

298. Greiner, J.D., O.N. Carlson, and J.F. Smith, *Single-crystal elastic constants of vanadium and vanadium with oxygen additions.* Journal of Applied Physics, 1979. **50**(6): pp. 4394-4398.

299. Ledbetter, H.M., *Elastic properties of zinc: a compilation and a review.* Journal of Physical and Chemical Reference Data, 1977. **6**(4): pp. 1181-1203.

300. Smith, J.F., F.H. Spedding, and C.E. Carlson, *Elastic properties of yttrium and eleven of the rare earth elements.* Journal of Metals, 1957. **9**: pp. 1212-1213.

301. Rosen, M., *Elastic moduli and Debye temperatures of the polycrystalline rare-earth metals at 4.2 K and 300 K.* Physical Review Letters, 1967. **19**(12): pp. 695-696.

302. Rosen, M., *Elastic moduli and ultrasonic attenuation of gadolinium, terbium, dysprosium, holmium, and erbium from 4.2 K to 300 K.* Physical Review, 1968. **174**(2): pp. 504-514.

303. Gust, W.H. and E.B. Royce, *New electronic interactions in rare-earth metals at high pressure.* Physical Review B, 1973. **8**(8): pp. 3595-3609.

304. Andreeva, L.P., G.P. Zinov'eva, and P.V. Gel'd, *Anomaly in the velocity of ultrasound in gadolinium near the Curie point.* Sov. Phys. Solid State (Engl. Transl.), 1972. **13**(11): pp. 2896-2898.

305. Greiner, J.D., D.M. Schlader, O.D. McMasters, J.K.A. Gschneidner, and J.F. Smith, *Elastic constants of neodymium single crystals in the temperature range 4.2 K to 300 K.* Journal of Applied Physics, 1976. **47**(8): pp. 3427-3431.

306. Rosen, M., *Elastic moduli and ultrasonic attenuation of praseodymium, neodymium, and samarium from 4.2 K to 300 K.* Physical Review, 1969. **180**(2): pp. 540-544.

307. Bridgman, P.W., *Certain effects of pressure on seven rare earth metals.* Proceedings of the American Academy of Arts and Sciences. Vol. 83. 1954: American Academy of Arts & Sciences, pp.3-21.

308. Fisher, E.S., M.H. Manghnani, and R. Kikuta, *Hydrostatic pressure derivatives of the single crystal elastic moduli of Gd, Dy and Er.* Journal of Physics and Chemistry of Solids, 1973. **34**(4): pp. 687-703.

309. Rosen, M., D. Kalir, and H. Klimker, *Single-crystal elastic constants and magnetoelasticity of erbium from 4.2 K to 300 K.* Physical Review B, 1973. **8**(9): pp. 4399-4404.

310. Plessis, P.V.d., *Elastic constants of single crystal erbium.* Journal of Physics F: Metal Physics, 1976. **6**(5): pp. 873-888.

311. Brookes, C.A., Greenwoo.Jh, and J.L. Routbort, *High-temperature tensile properties of iridium single crystals.* Journal of the Institute of Metals, 1970. **98**: pp. 27-31.

312. Merker, J., D. Lupton, M. Toepfer, and H. Knake, *High temperature mechanical properties of the platinum group metals elastic properties of platinum, rhodium and iridium and their alloys at high temperatures.* Platinum Metals Review, 2001. **45**(2): pp. 74-82.

313. Bridgman, P.W., *Viscosities to 30,000 kg/cm^2*. Daedalus: Proceedings of the American Academy of Arts and Sciences, 1949. **77**(4): pp. 117-128.

314. Robertson, W.M. and D.J. Montgomery, *Elastic modulus of isotopically-concentrated lithium.* Physical Review, 1960. **117**(2): pp. 440-442.

315. Nash, H.C. and C.S. Smith, *Single-crystal elastic constants of lithium.* Journal of Physics and Chemistry of Solids, 1959. **9**(2): pp. 113-118.

316. Felice, R.A., J. Trivisonno, and D.E. Schuele, *Temperature and pressure dependence of the single-crystal elastic constants of ^{6}Li and natural lithium.* Physical Review B, 1977. **16**(12): pp. 5173-5184.

317. Bridgman, P.W., *The effect of tension on the electrical resistance of certain abnormal metals.* Proceedings of the American Academy of Arts and Sciences, 1922. **57**(3): pp. 41-66.

318. Greiner, J.D., J.R.J. Schiltz, J.J. Tonnies, F.H. Spedding, and J.F. Smith, *Elastic constants of praseodymium single crystals in the temperature range 4.2 K to 300 K.* Journal of Applied Physics, 1973. **44**(9): pp. 3862-3867.

319. Guillermet, A.F. and G. Grimvall, *Thermodynamic properties of technetium.* Journal of the Less Common Metals, 1989. **147**(2): pp. 195-211.

320. Love, G.R., C.C. Koch, H.L. Whaley, and Z.R. McNutt, *Elastic moduli and Debye temperature of polycrystalline technetium by ultrasonic velocity measurements.* Journal of the Less Common Metals, 1970. **20**(1): pp. 73-75.

321. Maurer, D., R. Heichele, N. Lingg, V. Müller, and K. Rieder, H., *Ultrasonic study of the elastic behaviour of zone refined single crystalline rhodium.* J. Phys. IV France, 1996. **06**(C8): pp. 535-538.

322. Webelements. *Bulk Modulus: Periodicity*. [accessed 15/01/2012]; Available from: http://www.webelements.com/periodicity/bulk_modulus/.

323. Walker, E., J. Ashkenazi, and M. Dacorogna, *Elastic moduli of rhodium: correct prediction by a new theoretical method.* Physical Review B, 1981. **24**(4): pp. 2254-2256.

324. Voronov, F., O. Stal'gorova, and E. Gromnitskaya, *Anomalies in the variation of elastic properties of cesium during phase transformations under a pressure up to 5 GPa.* Journal of Experimental and Theoretical Physics, 2002. **95**(1): pp. 77-82.

325. Kollarits, F.J. and J. Trivisonno, *Single-crystal elastic constants of cesium.* Journal of Physics and Chemistry of Solids, 1968. **29**(12): pp. 2133-2139.

326. Anderson, M.S. and C.A. Swenson, *Experimental equations of state for cesium and lithium metals to 20 kbar and the high-pressure behavior of the alkali metals.* Physical Review B, 1985. **31**(2): pp. 668-680.

327. Bridgman, P.W., *Compression of 39 substances to 100,000 kg/cm^2*. Daedalus: Proceedings of the American Academy of Arts and Sciences, 1948. **76**: pp. 55-70.

328. Wiki. *Periodic Table of Elements*. [accessed 12/08/2011]; Available from: http://en.wikipedia.org/wiki/Periodic_table_of_elements.

329. Narayana, K.L. and K.M. Swamy, *Debye temperature in noble-metals.* Materials Science and Engineering, 1975. **18**(1): pp. 157-158.

330. Pantea, C., I. Stroe, H. Ledbetter, J.B. Betts, Y. Zhao, L.L. Daemen, H. Cynn, and A. Migliori, *Elastic constants of osmium between 5 K and 300 K.* Physical Review B, 2009. **80**(2): pp. 024112(1-10).

331. Cynn, H., J.E. Klepeis, C.-S. Yoo, and D.A. Young, *Osmium has the lowest experimentally determined compressibility.* Physical Review Letters, 2002. **88**(13): pp. 135701(1-4).

332. Pandey, D.K., D. Singh, and P.K. Yadawa, *Ultrasonic study of osmium and ruthenium.* Platinum Metals Review, 2009. **53**(2): pp. 91-97.

333. Occelli, F., D.L. Farber, J. Badro, C.M. Aracne, D.M. Teter, M. Hanfland, B. Canny, and B. Couzinet, *Experimental evidence for a high-pressure isostructural phase transition in osmium.* Physical Review Letters, 2004. **93**(9): pp. 095502(1-4).

334. Pantea, C., I. Mihut, H. Ledbetter, J.B. Betts, Y. Zhao, L.L. Daemen, H. Cynn, and A. Migliori, *Bulk modulus of osmium, 4 K to 300 K.* Acta Materialia, 2009. **57**(2): pp. 544-548.

335. Kenichi, T., *Bulk modulus of osmium: high-pressure powder X-ray diffraction experiments under quasihydrostatic conditions.* Physical Review B, 2004. **70**(1): pp. 012101(1-4).

336. Brown, H.L., P.E. Armstrong, and C.P. Kempter, *Elastic properties of polycrystalline Sc, Re, Ru and Pt-21 Ir*. Journal of the Less Common Metals, 1966. **11**(2): pp. 135-140.

337. Fisher, E.S. and D. Dever, *Temperature dependence of elastic moduli of ruthenium, rhenium, cobalt, dysprosium, and erbium: a study of the elastic anisotropy-phase transformation relationship.* Trans. Met. Soc. AIME, 1967. **239**: pp. 48-57

338. Manghnani, M.H., K. Katahara, and E.S. Fisher, *Ultrasonic equation of state of rhenium.* Physical Review B, 1974. **9**(4): pp. 1421-1431.

339. Shepard, M.L. and J.F. Smith, *Elastic constants of rhenium single crystals in the temperature range 4.2 K to 298 K.* Journal of Applied Physics, 1965. **36**(4): pp. 1447-1450.

340. EnvironmentalChemistry. *Periodic Table of Elements*. [accessed 15/01/2012]; Available from: http://environmentalchemistry.com/yogi/periodic/.

341. Fisher, E.S. and D. Dever, *Anomalous effects of temperature and impurities on the elastic moduli of scandium single crystals.* 7th Rare Earth Research Conference 1968: Coronado, California.

342. Leisure, R.G., R.B. Schwarz, A. Migliori, and M. Lei, *Room-temperature elastic-constants of Sc and $ScD_{0.18}$.* Physical Review B, 1993. **48**(2): pp. 1276-1279.

343. Rao, R.R. and C.S. Menon, *Lattice dynamics and elastic constants of scandium.* Solid State Communications, 1973. **12**(6): pp. 527-528.

344. Rosen, M., *Elastic moduli of thulium and ytterbium from 4.2 K to 300 K.* Journal of Physics and Chemistry of Solids, 1971. **32**(10): pp. 2351-2356.

345. Lim, C.M., C. Edwards, S. Dixon, and S.B. Palmer, *Ultrasound studies of single crystal thulium in an applied magnetic field.* Journal of Magnetism and Magnetic Materials, 2001. **234**(3): pp. 387-394.

346. Smith, J.F. and J.A. Gjevre, *Elastic constants of yttrium single crystals in the temperature range 4.2 K to 400 K.* Journal of Applied Physics, 1960. **31**(4): pp. 645-647.

347. Savage, S.J., S.B. Palmer, D. Fort, R.G. Jordan, and D.W. Jones, *The elastic constants of high-purity yttrium.* Journal of Physics F: Metal Physics, 1980. **10**(3): pp. 347-352.

348. Kreith, F., D.Y. Goswami, and B.I. Sandor, *The CRC Handbook of Mechanical Engineering*. 2nd Edn. 2004: Taylor & Francis, pp.2435-2536.

349. Reynolds, M.B., *New Elastic Constants of Some Metals*. 1951, Oak Ridge, Tenn.: U.S. Atomic Energy Commission, pp.1-35.

350. Carlson, O.N., P. Chiotti, G. Murphy, D. Peterson, and B.A. Rogers. *The metallurgy of thorium and its alloys*. Proceedings of the International Conference on the Peaceful Uses of Atomic Energy. 1956. United Nations Publications.

351. Armstrong, P.E., O.N. Carlson, and J.F. Smith, *Elastic constants of thorium single crystals in the range 77 K to 400 K.* Journal of Applied Physics, 1959. **30**(1): pp. 36-41.

352. Greiner, J.D., D.T. Peterson, and J.F. Smith, *Comparison of the single-crystal elastic constants of Th and a $ThC^{0.063}$ alloy*. Journal of Applied Physics, 1977. **48**(8): pp. 3357-3361.

353. Salama, K., F.R. Brotzen, and P.L. Donoho, *Elastic constants of holmium between 78 K and 300 K.* Journal of Applied Physics, 1973. **44**(1): pp. 180-183.

354. Rosen, M., D. Kalir, and H. Klimker, *Single-crystal elastic-constants and magnetoelasticity of holmium from 4.2 K to 300 K.* Journal of Physics and Chemistry of Solids, 1974. **35**(9): pp. 1333-1338.

355. Myers, A., *Some elastic and thermal properties of zirconium and tungsten.* Philosophical Magazine, 1960. **5**(57): pp. 927-938.

356. Fisher, E.S. and C.J. Renken, *Adiabatic elastic moduli of single crystal alpha zirconium.* Journal of Nuclear Materials, 1961. **4**(3): pp. 311-315.

357. Fisher, E.S., M.H. Manghnani, and T.J. Sokolowski, *Hydrostatic pressure derivatives of the single-crystal elastic moduli of zirconium.* Journal of Applied Physics, 1970. **41**(7): pp. 2991-2998.

358. Quimby, S.L. and S. Siegel, *The elastic constants of crystalline sodium at 80 K.* Physical Review, 1937. **52**(6): p. 665.

359. Daniels, W.B., *Pressure variation of the elastic constants of sodium.* Physical Review, 1960. **119**(4): pp. 1246-1252.

360. Martinson, R.H., *Variation of the elastic constants of sodium with temperature and pressure.* Physical Review, 1969. **178**(3): pp. 902-913.

361. Migliori, A., C. Pantea, H. Ledbetter, I. Stroe, J.B. Betts, J.N. Mitchell, M. Ramos, F. Freibert, D. Dooley, S. Harrington, and C.H. Mielke, *Alpha-plutonium's polycrystalline elastic moduli over its full temperature range.* The Journal of the Acoustical Society of America, 2007. **122**(4): pp. 1994-2001.

362. Beonisz, S.E., *The mechanical properties of alpha plutonium in compression.* Journal of Nuclear Materials, 1963. **9**(1): pp. 101-106.

363. Nadal, M.-H. and L. Bourgeois, *Elastic moduli of Pu and Ga stabilized delta Pu: experimental data and phenomenological behavior at high temperature.* Journal of Applied Physics, 2010. **108**(7): pp. 073532(1-6).

364. DeCadenet, J., *Plutonium*. 1960, London: Cleaver-Hume, p.107.

365. Oetting, F.L. and R.O. Adams, *The chemical thermodynamics of nuclear materials VIII. The high-temperature heat capacity of unalloyed plutonium metal.* The Journal of Chemical Thermodynamics, 1983. **15**(6): pp. 537-554.

366. Lacquer, H., *The Metal Plutonium*. 1961, Metals Park: ASM,

367. Rosen, M., G. Erez, and S. Shtrikman, *Evidence for a cooperative electron transition in plutonium at low temperatures.* Physical Review Letters, 1968. **21**(7): pp. 430-431.

368. Merz, M., J. Hammer, and H. Kjarmo. in *The 15th International Conference on Plutonium and other Actinides*. 1975. Baden, Germany.

369. Calder, C.A., E.C. Draney, and W.W. Wilcox, *Noncontact measurement of the elastic constants of plutonium at elevated temperatures.* Journal of Nuclear Materials, 1981. **97**(1–2): pp. 126-136.

370. Greiner, J.D., O.D. McMasters, and J.F. Smith, *Single-crystal elastic constants of gamma-cerium.* Scripta Metallurgica, 1980. **14**(9): pp. 989-991.

371. Winder, D.R. and C.S. Smith, *Single-crystal elastic constants of indium.* Journal of Physics and Chemistry of Solids, 1958. **4**(1-2): pp. 128-134.

372. Chandrasekhar, B.S. and J.A. Rayne, *Elastic constants of indium from 1.4 K to 300 K.* Physical Review, 1961. **124**(4): pp. 1011-1014.

373. Kim, S. and H. Ledbetter, *Low-temperature elastic coefficients of polycrystalline indium.* Materials Science and Engineering: A, 1998. **252**(1): pp. 139-143.

374. Bridgman, P.W., *The compression of 46 substances to 50,000 kg/cm^2*. Daedalus: Proceedings of the American Academy of Arts and Sciences, 1940. **74**(3): pp. 21-51.

375. Bender, O., *Elastizitätsmessungen an Alkalimetall-Einkristallen in tiefer Temperatur (Elascitity measurments on alkali metals-single crystals in deep temperatures).* Annalen der Physik, 1939. **426**(4): pp. 359-376.

376. Smith, P.A. and C.S. Smith, *Pressure derivatives of the elastic constants of potassium.* Journal of Physics and Chemistry of Solids, 1965. **26**(2): pp. 279-289.

377. Marquardt, W.R. and J. Trivisonno, *Low temperature elastic constants of potassium.* Journal of Physics and Chemistry of Solids, 1965. **26**(2): pp. 273-278.

378. Ferris, R.W., M.L. Shepard, and J.F. Smith, *Elastic constants of thallium single crystals in the temperature range 4.2 K to 300 K.* Journal of Applied Physics, 1963. **34**(4): pp. 768-770.

379. Gold, L., *Evaluation of the stiffness coefficients for beryllium from ultrasonic measurements in polycrystalline and single crystal specimens.* Physical Review, 1950. **77**(3): pp. 390-395.

380. Smith, J.F. and C.L. Arbogast, *Elastic constants of single crystal beryllium.* Journal of Applied Physics, 1960. **31**(1): pp. 99-102.

381. Silversmith, D.J. and B.L. Averbach, *Pressure dependence of the elastic constants of beryllium and beryllium-copper alloys.* Physical Review B, 1970. **1**(2): pp. 567-571.

382. Rowlands, W.D. and J.S. White, *The determination of the elastic constants of beryllium in the temperature range 25 °C to 300 °C.* Journal of Physics F: Metal Physics, 1972. **2**(2): pp. 231-236.

383. Migliori, A., H. Ledbetter, D.J. Thoma, and T.W. Darling, *Beryllium's monocrystal and polycrystal elastic constants.* Journal of Applied Physics, 2004. **95**(5): pp. 2436-2440.

384. Walsh, K.A. and E.E. Vidal, *Beryllium Chemistry and Processing*. 2009: ASM International, p.28.

385. Garland, C.W. and J. Silverman, *Elastic constants of cadmium from 4.2 K to 300 K.* Physical Review, 1960. **119**(4): pp. 1218-1222.

386. Bridgman, P.W., *Certain physical properties of single crystals of tungsten, antimony, bismuth, tellurium, cadmium, zinc, and tin.* Proceedings of the American Academy of Arts and Sciences, 1925. **60**(6): pp. 305-383.

387. Grüneisen, E. and E. Goens, *Untersuchungen an Metallkristallen.* Zeitschrift für Physik, 1924. **26**(1): pp. 235-249.

388. Chang, Y.A. and L. Himmel, *Elastic constants of cadmium from 300 K to 575 K.* Journal of Applied Physics, 1966. **37**(10): pp. 3787-3790.

389. Rosen, M., *Elastic moduli and ultrasonic attenuation of polycrystalline europium from 4.2 K to 300 K.* Physical Review, 1968. **166**(2): pp. 561-564.

390. Burkhanov, A.M., N.P. Grazhdankina, and I.G. Fakidov, *Investigation of the moduli of elasticity and compressibility of europium in the temperature range 70 K to 295 K.* Sov. Phys. Solid State (Engl. Transl.), 1967: pp. 586-587.

391. Bodryakov, V.Y. and S.A. Nikitin, *Young's modulus and internal friction of europium.* Journal of Alloys and Compounds, 1998. **269**(1–2): pp. 224-232.

392. Prestonthomas, H., *The international temperature scale of 1990 (ITS-90).* Metrologia, 1990. **27**(1): pp. 3-10.

393. Lyall, K.R. and J.F. Cochran, *Velocity of sound and acoustic attenuation in pure gallium single crystals.* Canadian Journal of Physics, 1971. **49**(9): pp. 1075-1097.

394. Langill, T.J. and J. Trivisonno, *Temperature-dependence of elastic-constants of gallium.* Canadian Journal of Physics, 1975. **53**(6): pp. 581-582.

395. Lyapin, A., E. Gromnitskaya, O. Yagafarov, O. Stal'gorova, and V. Brazhkin, *Elastic properties of crystalline and liquid gallium at high pressures.* Journal of Experimental and Theoretical Physics, 2008. **107**(5): pp. 818-827.

396. Nixon, L.W., D.A. Papaconstantopoulos, and M.J. Mehl, *Electronic structure and superconducting properties of lanthanum.* Physical Review B, 2008. **78**(21): pp. 214510(1-13).

397. Tonnies, J.J., J.K.A. Gschneidner, and F.H. Spedding, *Elastic moduli and thermal expansion of lutetium single crystals from 4.2 K to 300 K.* Journal of Applied Physics, 1971. **42**(9): pp. 3275-3283.

398. Roberts, C.A. and R. Meister, *The elastic constants of rubidium.* Journal of Physics and Chemistry of Solids, 1966. **27**(9): pp. 1401-1407.

399. Gutman, E.J. and J. Trivisonno, *Temperature dependence of the elastic constants of rubidium.* Journal of Physics and Chemistry of Solids, 1967. **28**(5): pp. 805-809.

400. Anderson, M.S. and C.A. Swenson, *Experimental compressions for sodium, potassium, and rubidium metals to 20 kbar from 4.2 K to 300 K.* Physical Review B, 1983. **28**(10): pp. 5395-5418.

401. Ellinger, F.H. and W.H. Zachariasen, *The crystal structure of samarium metal and of samarium monoxide.* Journal of the American Chemical Society, 1953. **75**(22): pp. 5650-5652.

402. Fisher, E.S., *Temperature dependence of the elastic moduli in alpha uranium single crystals, part IV (298 K to 923 K).* Journal of Nuclear Materials, 1966. **18**(1): pp. 39-54.

403. Rosen, M., *Elastic moduli and ultrasonic attenuation of polycrystalline uranium from 4.2 K to 300 K.* Physics Letters A, 1968. **28**(6): pp. 438-439.

404. Armstrong, P.E., D.T. Eash, and J.E. Hockett, *Elastic moduli of alpha, beta and gamma polycrystalline uranium.* Journal of Nuclear Materials, 1972. **45**(3): pp. 211-216.

405. Abey, A.E. and B.P. Bonner, *Elastic constants of polycrystalline alpha-uranium.* Journal of Applied Physics, 1975. **46**(4): pp. 1427-1428.

406. Fisher, E.S. and H.J. McSkimin, *Adiabatic elastic moduli of single crystal alpha-uranium.* Journal of Applied Physics, 1958. **29**(10): pp. 1473-1484.

407. McSkimin, H.J. and E.S. Fisher, *Temperature dependence of the adiabatic elastic moduli of single-crystal alpha uranium.* Journal of Applied Physics, 1960. **31**(9): pp. 1627-1639.

408. Fisher, E.S. and D. Dever, *Elastic moduli and phase transition in uranium at T<43 K.* Physical Review, 1968. **170**(3): pp. 607-613.

409. Boguslavskii, Y.Y., V.A. Goncharova, and G.G. Il'ina, *Elastic behavior of ytterbium at a semimetal-semiconductor-semimetal transition under pressure.* Journal of Experimental and Theoretical Physics, 1995. **80**(2): pp. 248-53.

410. Wicks, L.R. and W.R. Tyson, *Elastic and plastic anisotropy of indium.* Canadian Journal of Physics, 1975. **53**(14): pp. 1338-1348.

411. Mclellan, R.B. and Y.C. Angel, *The elastic behavior of sodium.* Journal of Physics and Chemistry of Solids, 1996. **57**(11): pp. 1709-1711.

412. Rosen, M., *Effect of the low-temperature phase transformations on the elastic behavior of cerium.* Physical Review, 1969. **181**(2): pp. 932-935.

413. Kammer, E.W., *Method for evaluating transducer-loading effects on ultrasonic transit-time measurements.* The Journal of the Acoustical Society of America, 1965. **37**(5): pp. 929-931.

414. Edalati, K., A. Yamamoto, Z. Horita, and T. Ishihara, *High-pressure torsion of pure magnesium: evolution of mechanical properties, microstructures and hydrogen storage capacity with equivalent strain.* Scripta Materialia, 2011. **64**(9): pp. 880-883.

415. Kolobov, Y.R., B. Kieback, K.V. Ivanov, T. Weissgaerber, N.V. Girsova, Y.I. Pochivalov, G.P. Grabovetskaya, M.B. Ivanov, V.U. Kazyhanov, and I.V. Alexandrov, *The structure and microhardness evolution in submicrocrystalline molybdenum processed by severe plastic deformation followed by annealing.* International Journal of Refractory Metals & Hard Materials, 2003. **21**(1-2): pp. 69-73.

416. Wadsack, R., R. Pippan, and B. Schedler, *Structural refinement of chromium by severe plastic deformation.* Fusion Engineering and Design, 2003. **66-68**: pp. 265-269.

417. Wei, Q., H.T. Zhang, B.E. Schuster, K.T. Ramesh, R.Z. Valiev, L.J. Kecskes, R.J. Dowding, L. Magness, and K. Cho, *Microstructure and mechanical properties of super-strong nanocrystalline tungsten processed by high-pressure torsion.* Acta Materialia, 2006. **54**(15): pp. 4079-4089.

418. Myneni, G.R., *Physical and mechanical properties of niobium for SRF science and technology*, Single Crystal - Large Grain Niobium Technology, G.C.T.H.A. Myneni, Editor. 2007, pp. 41-47.

419. Cullity, B.D., *Elements of X-ray Diffraction*. 1956: Addison-Wesley Publishing Company, Inc., pp.482-484.

420. Evans, J.R.G., M.J. Edirisinghe, P.V. Coveney, and J. Eames, *Combinatorial searches of inorganic materials using the ink jet printer: science, philosophy*

and technology. Journal of the European Ceramic Society, 2001. **21**(13): pp. 2291-2299.

421. Etherincton, H., *Conclusions concerning latent heat of fusion and melting point of the metals.* Journal of the Society of Chemical Industry, 1926. **45**(26): pp. 430-431.

422. Kanno, H., *Alkali halides: relationship between heat of fusion, melting point, electronic polarizability and the ionic radius ratio.* Nature, 1968. **218**(5143): pp. 765-766.

423. Donoho, D.L., *High-dimensional data analysis: the curses and blessings of dimensionality* Aide-Memoire of a Lecture at AMS Conference on Math Challenges of the 21st Century2000. pp. 1-33.

424. Giustolisi, O. and D. Laucelli, *Improving generalization of artificial neural networks in rainfall-runoff modelling.* Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 2005. **50**(3): pp. 439-457.

425. Sha, W., *Comment on "Prediction of the flow stress of 0.4C–1.9Cr–1.5Mn–1.0Ni–0.2Mo steel during hot deformation"by R.H. Wu et al. [J. Mater. Process. Technol. 116 (2001) 211]*. Journal of Materials Processing Technology, 2006. **171**(2): pp. 283-284.

426. Sha, W., *Comment on "Flow forecasting for a Hawaii stream using rating curves and neural networks" by G.B. Sahoo and C. Ray [Journal of Hydrology 317 (2006) 63–80]*. Journal of Hydrology, 2007. **340**(1-2): pp. 119-121.

427. Sha, W., *Comments on "Water quality retrievals from combined Landsat TM Data and ERS-2 SAR Data in the Gulf of Finland"*. IEEE Transactions on Geoscience and Remote Sensing, 2007. **45**(6): pp. 1896-1897.

428. Sha, W. and K.L. Edwards, *The use of artificial neural networks in materials science based research.* Materials & Design, 2007. **28**(6): pp. 1747-1752.

429. Horita, Z. and T.G. Langdon, *Microstructural evolution in the processing of bulk samples using high-pressure torsion*, Materials Science Forum, 2007. **539-543**: pp. 80-85.

430. Gludovatz, B., S. Wurster, T. Weingaertner, A. Hoffmann, and R. Pippan, *Influence of impurities on the fracture behaviour of tungsten.* Philosophical Magazine, 2011. **91**(22): pp. 3006-3020.

431. Matsunaga, H. and Z. Horita, *Softening and microstructural coarsening without twin formation in fcc metals with low stacking fault energy after processing by high-pressure torsion.* Materials Transactions, 2009. **50**(7): pp. 1633-1637.

432. An, X.H., S.D. Wu, Z.F. Zhang, R.B. Figueiredo, N. Gao, and T.G. Langdon, *Evolution of microstructural homogeneity in copper processed by high-pressure torsion.* Scripta Materialia, 2010. **63**(5): pp. 560-563.

433. Chen, Y.J., Y.J. Li, J.C. Walmsley, N. Gao, H.J. Roven, M.J. Starink, and T.G. Langdon, *Microstructural heterogeneity in hexagonal close-packed pure Ti processed by high-pressure torsion.* Journal of Materials Science, 2012. **47**(12): pp. 4838-4844.

434. Popova, E.N., V.V. Popov, E.P. Romanov, and V.P. Pilyugin, *Thermal stability of nanocrystalline Nb produced by severe plastic deformation.* Physics of Metals and Metallography, 2006. **101**(1): pp. 52-57.

435. Schuster, B.E., J.P. Ligda, Z.L. Pan, and Q. Wei, *Nanocrystalline refractory metals for extreme condition applications.* JOM, 2011. **63**(12): pp. 27-31.

436. Valiev, R.Z., Y.V. Ivanisenko, E.F. Rauch, and B. Baudelet, *Structure and deformaton behaviour of Armco iron subjected to severe plastic deformation.* Acta Materialia, 1996. **44**(12): pp. 4705-4712.

437. Xu, C., Z. Horita, and T.G. Langdon, *The evolution of homogeneity in processing by high-pressure torsion.* Acta Materialia, 2007. **55**(1): pp. 203-212.

438. Zhilyaev, A.P., F. Gálvez, A. Sharafutdinov, and M.T. Pérez-Prado, *Influence of the high pressure torsion die geometry on the allotropic phase transformations in pure Zr.* Materials Science and Engineering: A, 2010. **527**(16–17): pp. 3918-3928.

439. Zhilyaev, A.P., S. Lee, G.V. Nurislamova, R.Z. Valiev, and T.G. Langdon, *Microhardness and microstructural evolution in pure nickel during high-pressure torsion.* Scripta Materialia, 2001. **44**(12): pp. 2753-2758.

440. Todaka, Y., Y. Miki, M. Umemoto, C. Wang, and K. Tsuchiya, *Tensile property of submicrocrystalline pure Fe produced by HPT-straining*. Materials Science Forum, 2008. **584-586**: pp. 597-602.

441. Bachmaier, A., M. Hafok, R. Schuster, and R. Pippan., *Limitations in the refinement by severe plastic deformation: the effect of processing.* Reviews on Advanced Materials Science, 2010. **25**: pp. 16-22.

442. Harai, Y., Y. Ito, and Z. Horita, *High-pressure torsion using ring specimens.* Scripta Materialia, 2008. **58**(6): pp. 469-472.

443. Edalati, K., T. Fujioka, and Z. Horita, *Evolution of mechanical properties and microstructures with equivalent strain in pure Fe processed by high pressure torsion.* Materials Transactions, 2009. **50**(1): pp. 44-50.

444. Rusz, S., K. Malanik, and M. Klos, *Increasing SPD effectiveness by changing deformation process in the first pass through the ECAP die: numerical modelling.* International Journal of Computational Materials Science and Surface Engineering, 2007. **1**(5): pp. 620-632.

445. Garson, G.D., *Interpreting neural-network connection weights.* AI Expert, 1991. **6**(4): pp. 46-51.

446. Francone, F.D., *Discipulus with Notitia and Solution Analytics*, Machine Learning Technologies, Inc. Available from http://www.rmltech.com/.

447. Francone, F.D., *Discipulus Owner's Manual*. 2010, Littleton, Colorado: Register Machine Learning Technologies, Inc., pp.2, 32, 94-96, 144-147, 151-153.

448. Yu, T. and J. Rutherford, *Modeling sparse engine test data using genetic programming*, The 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2001: San Francisco, California, USA. pp. 42-48.

449. Giacalone, A., *L'equazione di Kirchhoff modificata, dedotta dalla equazione di Gibbs-Helmholtz.* Gazetta Chimica Italiana, 1951. **81**(180): pp. 183-190.

450. Chen, N.H., *Generalized correlation for latent heat of vaporization.* Journal of Chemical and Engineering Data, 1965. **10**(2): pp. 207-210.

451. Ravisankar, P., V. Ravi, and I. Bose, *Failure prediction of dotcom companies using neural network-genetic programming hybrids.* Information Sciences, 2010. **180**(8): pp. 1257-1267.

452. Brezocnik, M., J. Balic, and K. Kuzman, *Genetic programming approach to determining of metal materials properties.* Journal of Intelligent Manufacturing, 2002. **13**(1): pp. 5-17.

453. Silva, S., *GPLAB - A Genetic Programming Toolbox for MATLAB,* 2009. Available from http://gplab.sourceforge.net/index.html.

454. Webelements. *Poisson's Ratio: Periodicity*. [accessed 15/01/2012]; Available from: http://www.webelements.com/periodicity/poissons_ratio/.

455. Kaye, G.W.C. and T.H. Laby, *Tables of Physical and Chemical Constants* 15th Edn. 1986: Harlow : Longman, pp.31-32.

456. Rao, C.N., *A Handbook of Chemistry and Physics*. 2nd Edn. 1970: Affiliated East-West Press, pp.154-171.

457. Webelements. *Young's Modulus: Periodicity*. [accessed 15/01/2012]; Available from: http://www.webelements.com/periodicity/youngs_modulus/.

458. Webelements. *Rigidity Modulus: Periodicity*. [accessed 15/01/2012]; Available from: http://www.webelements.com/periodicity/rigidity_modulus/.

459. Lahteenkorva, E.E. and J.T. Lenkkeri, *Effects of magnetic order on elastic moduli of chromium.* Journal of Physics F: Metal Physics, 1981. **11**(4): pp. 767-773.

460. Carroll, K.J., *Elastic constants of niobium from 4.2 K to 300 K.* Journal of Applied Physics, 1965. **36**(11): pp. 3689-3690.

461. Hsu, D.K. and R.G. Leisure, *Elastic constants of palladium and β-phase palladium hydride between 4 K and 300 K.* Physical Review B, 1979. **20**(4): pp. 1339-1344.

462. Konti, A., *Lattice dynamics and thermodynamic properties of platinum.* Journal of Chemical Physics, 1971. **55**(8): pp. 3997-4000.

463. Naimon, E.R., *Third-order elastic constants of magnesium. I. Experimental.* Physical Review B, 1971. **4**(12): pp. 4291-4296.

464. Chung, D.H. and W.R. Buessem, *The Voigt-Reuss-Hill approximation and elastic moduli of polycrystalline MgO, CaF_2, β-ZnS, ZnSe, and CdTe.* Journal of Applied Physics. **38**(6): pp. 2535-2540.

465. Antonangeli, D., M. Krisch, G. Fiquet, J. Badro, D.L. Farber, A. Bossak, and S. Merkel, *Aggregate and single-crystalline elasticity of hcp cobalt at high pressure.* Physical Review B, 2005. **72**(13): pp. 134303(1-7).

Appendix

Appendix I

Table 1 The dataset of 37 elements with 24 properties used in the ANN combinatorial search

Property	**Melting point**	**Specific heat capacity**	**Heat of fusion**	**Thermal conductivity at 300K**	**Thermal expansion coefficient at 300K**	**Bulk modulus at 300K**
Abbreviation	**MP**	**SHC**	**HF**	**TC**	**TEC**	**BM**
Symbol	T_m	C_p	H_{fus}	λ	α_L	K
Unit	**°C**	**J/kg·°C**	**kJ/mol**	**W/m·°C**	**μstrain/°C**	**GPa**
Li	181	3600	3	84.7	56	11.6
Be	1280	1820	12.2	202	11.5	100
Na	97.8	1230	2.6	141	70.6	6.81
Mg	649	1020	8.95	156	26.1	35.4
Al	660	900	10.8	237	23	72.2
K	63.7	750	2.33	102	83	3.18
Ca	839	630	8.54	200	22	15.2
Ti	1660	520	15.5	21.9	8.35	105
Cr	1860	450	16.9	93.7	6.2	190
Cu	1080	380	13.1	401	16.5	131
Zn	420	390	7.32	116	25	59.8
Rb	39.1	363	2.19	58.2	90	3.14
Zr	1850	270	16.9	22.7	5.78	83.3
Nb	2470	260	26.4	53.7	7.07	170
Mo	2620	250	32	138	5.43	273
Ru	2310	238	24	117	9.1	348
Rh	1970	242	21.5	150	8.4	271
Pd	1550	240	17.6	71.8	11.2	181
Ag	962	235	11.3	429	19.2	101
Cd	321	230	6.19	96.8	29.8	46.7
In	156	230	3.26	81.6	33	41.1
Sb	631	210	19.9	24.3	8.5	38.3
Cs	28.4	240	2.09	35.9	97	2.03
Ba	729	204	7.75	18.4	18.1	10.3
La	921	190	6.2	13.5	4.9	24.3

Property	Melting point	Specific heat capacity	Heat of fusion	Thermal conductivity at 300K	Thermal expansion coefficient at 300K	Bulk modulus at 300K
Abbreviation	**MP**	**SHC**	**HF**	**TC**	**TEC**	**BM**
Symbol	T_m	C_p	H_{fus}	λ	α_L	K
Unit	**°C**	**J/kg·°C**	**kJ/mol**	**W/m·°C**	**μstrain/°C**	**GPa**
Hf	2230	140	24.1	23	5.9	109
Ta	3000	140	31.6	57.5	6.6	200
W	3410	130	35.4	174	4.59	323
Re	3180	130	33.2	47.9	6.63	372
Os	3050	130	31.8	87.6	5.55	462
Ir	2410	130	26.1	147	6.4	383
Pt	1770	130	19.6	71.6	9	278
Au	1060	128	12.6	317	14.2	173
Hg	-38.9	139	2.3	8.34	49	38.2
Tl	304	130	4.14	46.1	28	35.9
Pb	328	130	4.8	35.5	29.1	42.9
Bi	271	120	11.3	7.87	13.4	31.5

Property	Electrical resistivity at 300K	Young's modulus at 300K	Atomic radius	Shear modulus at 300K	Cohesive energy	Lattice parameter a
Abbreviation	**ER**	**YM**	**AR**	**SM**	**CE**	**LPA**
Symbol	ρ_e	E	r	G	E_{coh}	a
Unit	**μohm·cm**	**GPa**	**nm**	**GPa**	**kJ/mol**	**nm**
Li	8.55	10	0.152	4.31	160	0.349
Be	4	301	0.113	146	322	0.227
Na	4.2	5	0.154	3.5	109	0.423
Mg	4.38	44.4	0.16	17.7	148	0.321
Al	2.65	70.5	0.143	27.1	322	0.405
K	6.15	2.4	0.227	1.3	90.9	0.523
Ca	3.43	21	0.197	7.5	176	0.558
Ti	42	110	0.145	40.1	470	0.295
Cr	12.7	259	0.125	119	396	0.288
Cu	1.67	124	0.128	46	339	0.361
Zn	5.92	95	0.133	37.9	130	0.266
Rb	12.5	2	0.248	1.02	84.6	0.559
Zr	42.1	94	0.16	34.8	610	0.323

Property	Electrical resistivity at 300K	Young's modulus at 300K	Atomic radius	Shear modulus at 300K	Cohesive energy	Lattice parameter a
Abbreviation	**ER**	**YM**	**AR**	**SM**	**CE**	**LPA**
Symbol	ρ_e	E	r	G	E_{coh}	a
Unit	**μohm·cm**	**GPa**	**nm**	**GPa**	**kJ/mol**	**nm**
Nb	12.5	104	0.143	38.2	730	0.33
Mo	5.2	322	0.136	118	658	0.315
Ru	7.6	430	0.134	163	645	0.271
Rh	4.51	330	0.134	150	557	0.38
Pd	10.8	127	0.138	52.1	377	0.389
Ag	1.59	80	0.144	29.2	286	0.409
Cd	6.83	62	0.149	24.6	112	0.298
In	8.37	14	0.163	3.8	241	0.325
Sb	39	67	0.182	20.4	261	0.431
Cs	20	1.8	0.265	0.66	80.5	0.605
Ba	50	13	0.217	5	179	0.502
La	57	50	0.188	15.2	427	0.377
Hf	35.1	139	0.156	54	610	0.319
Ta	12.5	183	0.143	70	782	0.33
W	5.65	401	0.137	156	837	0.316
Re	19.3	461	0.137	182	780	0.276
Os	8.12	550	0.135	214	784	0.274
Ir	5.3	533	0.136	214	666	0.384
Pt	10.6	175	0.138	62.2	565	0.392
Au	2.35	78.3	0.144	28.1	367	0.408
Hg	94.1	22	0.16	10.2	64.6	0.299
Tl	18	12	0.17	2.8	182	0.346
Pb	20.6	20	0.175	5.5	197	0.495
Bi	107	34	0.155	13.1	210	0.455

Property	Poisson's ratio	Magnetic susceptibility	Debye temperature	T dependence of resistivity	Surface energy (liquid)	Work function
Abbreviation	**PR**	**MS**	**DT**	**TDR**	**SE**	**WF**
Symbol	ν	χ_m	θ_D	T_r	Υ	W
Unit			**°C**	**Ω/°C**	**J/m^2**	**eV**
Li	0.362	0.0000137	70.9	0.00437	0.395	2.9
Be	0.08	-0.0000232	1170	0.0075	1.39	4.98
Na	0.315	0.0000085	-115	0.0055	0.195	2.75
Mg	0.3	0.0000118	127	0.0042	0.559	3.7
Al	0.34	0.0000208	155	0.00429	0.914	4.28
K	0.35	0.00000574	-182	0.0054	0.111	2.3
Ca	0.31	0.0000193	-43.2	0.004	0.361	2.87
Ti	0.34	0.000181	147	0.0055	1.65	4.33
Cr	0.209	0.000313	357	0.003	1.7	4.5
Cu	0.35	-9.63E-06	69.9	0.00433	1.29	4.65
Zn	0.35	-0.0000156	53.9	0.0042	0.782	4.33
Rb	0.356	0.0000038	-217	0.0053	0.083	2.16
Zr	0.34	0.000109	17.9	0.004	1.48	4.05
Nb	0.38	0.000226	1.85	0.00228	1.9	4.3
Mo	0.3	0.000119	177	0.0047	2.25	4.6
Ru	0.286	0.0000661	327	0.0045	2.25	4.71
Rh	0.27	0.000168	207	0.00457	2	4.98
Pd	0.375	0.000802	0.85	0.0038	1.5	5.12
Ag	0.37	-0.0000238	-48.2	0.0041	0.903	4.26
Cd	0.3	-0.0000191	-64.2	0.00426	0.57	4.22
In	0.46	-0.0000511	-165	0.0051	0.556	4.12
Sb	0.31	-0.0000683	-62.2	0.0051	0.367	4.55
Cs	0.356	0.00000515	-235	0.005	0.069	2.14
Ba	0.28	0.00000662	-163	0.0061	0.224	2.7
La	0.288	0.0000661	-131	0.0022	0.72	3.5
Hf	0.3	0.0000703	-21.2	0.0038	1.63	3.9
Ta	0.35	0.000178	-33.2	0.0036	2.15	4.25
W	0.284	0.000078	127	0.00483	2.5	4.55
Re	0.293	0.0000937	157	0.0031	2.7	4.96
Os	0.285	0.0000147	227	0.0042	2.5	4.83
Ir	0.26	0.0000375	147	0.00433	2.25	5.27
Pt	0.39	0.000279	-33.2	0.00392	1.8	5.65

Property	Poisson's ratio	Magnetic susceptibility	Debye temperature	T dependence of resistivity	Surface energy (liquid)	Work function
Abbreviation	**PR**	**MS**	**DT**	**TDR**	**SE**	**WF**
Symbol	ν	χ_m	θ_D	T_r	Υ	W
Unit			**°C**	**Ω/°C**	**J/m^2**	**eV**
Au	0.425	-0.0000345	-108	0.00398	1.14	5.1
Hg	0.364	-0.0000285	-201	0.00099	0.498	4.49
Tl	0.46	-0.000371	-195	0.0052	0.464	3.84
Pb	0.44	-0.0000158	-168	0.00422	0.468	4.25
Bi	0.33	-0.0000165	-154	0.00445	0.378	4.22

Property	Neutron absorption cross section (0.025 eV)	Molar volume	Neutron scattering cross section (0.025 eV)	Atomic weight	Heat of vaporization at the normal boiling point	Boiling point
Abbreviation	**NACS**	**MV**	**NSCS**	**AW**	**HV**	**BP**
Symbol	σ_A	V_m	σ_S	A_r	ΔH_{vb}	T_b
Unit	**Barns**	**m^3/kmol**	**Barns**	**kg/kmol**	**kJ/mol**	**°C**
Li	70.7	0.013	1.4	6.94	146	1350
Be	0.0092	0.00488	6.14	9.01	292	2970
Na	0.53	0.0237	3.2	23	97	883
Mg	0.063	0.014	3.42	24.3	127	1090
Al	0.23	0.01	1.49	27	293	2470
K	2.1	0.0454	1.5	39.1	79.9	774
Ca	0.43	0.0259	3	40.1	154	1480
Ti	6.1	0.0106	4	47.9	421	3290
Cr	3.1	0.00723	3.8	52	344	2670
Cu	3.79	0.00709	7.9	63.5	300	2570
Zn	1.1	0.00917	4.2	65.4	115	907
Rb	0.37	0.0558	6.2	85.5	72.2	688
Zr	0.185	0.014	6.4	91.2	611	4380
Nb	1.15	0.0108	5	92.9	682	4740
Mo	2.65	0.00939	5.8	95.9	598	4610
Ru	2.56	0.00814	6	101	595	3900
Rh	150	0.00829	5	103	493	3730

Property	Neutron absorption cross section (0.025 eV)	Molar volume	Neutron scattering cross section (0.025 eV)	Atomic weight	Heat of vaporization at the normal boiling point	Boiling point
Abbreviation	**NACS**	**MV**	**NSCS**	**AW**	**HV**	**BP**
Symbol	σ_A	V_m	σ_S	A_r	ΔH_{vb}	T_b
Unit	**Barns**	**m^3/kmol**	**Barns**	**kg/kmol**	**kJ/mol**	**°C**
Pd	6.9	0.00885	5	106	357	3140
Ag	63.6	0.0103	6	108	251	2210
Cd	2450	0.013	5.6	112	99.6	765
In	194	0.0157	2	115	232	2080
Sb	5.4	0.0182	4.2	122	262	1630
Cs	29	0.071	20	133	67.7	678
Ba	1.2	0.0382	8	137	142	1640
La	9	0.0226	9.3	139	414	3460
Hf	102	0.0134	8	178	575	5200
Ta	21	0.0109	6.2	181	743	5420
W	18.5	0.00953	5	184	824	5660
Re	88	0.00886	11.3	186	715	5630
Os	15.3	0.00843	11	190	746	5030
Ir	426	0.00857	14	192	604	4130
Pt	10	0.0091	11.2	195	510	3830
Au	98.8	0.0102	9.3	197	334	2810
Hg	375	0.0148	20	201	59.2	357
Tl	3.4	0.0172	9.7	204	164	1460
Pb	0.17	0.0183	11.4	207	178	1740
Bi	0.033	0.0214	9	209	209	1610

Appendix II

Table 2 The top 50 binary order correlations obtained from the ANN combinatorial search. Abbreviations are listed in Appendix I.

Rank	ANNs evaluation parameters			Conditions	
	ψ	δ	E_c	Predicted property	Input property
1	0.000	0.001	2.2%	HV	CE
2	0.001	0.003	2.5%	CE	HV
3	0.002	0.04	1.6%	AR	MV
4	0.002	0.013	4.3%	AW	SHC
5	0.004	0.024	5.6%	MV	AR
6	0.005	0.024	6.9%	BP	HV
7	0.005	0.024	6.9%	CE	BP
8	0.005	0.026	6.9%	BP	CE
9	0.006	0.024	7.5%	HV	BP
10	0.009	0.009	9.3%	SM	YM
11	0.010	0.028	9.4%	TC	ER
12	0.010	0.004	9.9%	YM	SM
13	0.013	0.104	4.7%	SHC	AW
14	0.019	0.007	13.8%	ER	TC
15	0.019	0.044	13.2%	SE	MP
16	0.021	0.058	13.3%	CE	MP
17	0.025	0.148	5.1%	WF	BM
18	0.025	0.083	13.3%	BP	MP
19	0.027	0.081	14.3%	MV	SE
20	0.029	0.049	16.4%	TEC	HF
21	0.031	0.079	15.8%	HV	MP
22	0.032	0.078	16.1%	HF	MP
23	0.033	0.175	5.2%	AR	WF
24	0.033	0.109	14.7%	CE	HF
25	0.034	0.085	16.5%	HV	HF
26	0.035	0.134	13.1%	MV	BM
27	0.041	0.078	18.6%	TEC	MP
28	0.042	0.116	16.9%	CE	SE
29	0.044	0.087	19.1%	BM	SE
30	0.044	0.088	19.1%	SE	BM

Rank	ANNs evaluation parameters			Conditions	
	ψ	δ	E_c	Predicted property	Input property
31	0.045	0.141	15.9%	MV	SM
32	0.046	0.092	19.3%	MV	WF
33	0.053	0.139	18.3%	MV	LPA
34	0.053	0.169	15.6%	BP	SE
35	0.053	0.138	18.4%	MV	YM
36	0.053	0.225	5.0%	WF	MV
37	0.055	0.227	6.3%	AR	LPA
38	0.056	0.063	22.9%	MP	HF
39	0.060	0.048	24.1%	MP	CE
40	0.061	0.081	23.3%	SE	YM
41	0.061	0.24	5.8%	AR	AW
42	0.065	0.106	23.1%	SE	SM
43	0.066	0.097	23.8%	SE	AW
44	0.067	0.185	18.1%	BP	HF
45	0.067	0.254	5.2%	AR	SE
46	0.069	0.12	23.4%	CE	AW
47	0.073	0.097	25.2%	SE	CE
48	0.074	0.186	19.8%	HV	SE
49	0.077	0.145	23.7%	HF	HV
50	0.078	0.145	23.9%	HF	CE

Appendix III

Table 3 The top 25 binary order correlation-groups obtained from the ANN combinatorial search. Abbreviations are listed in Appendix I.

Rank /Property group	ANNs evaluation parameters				Conditions	
	ψ_2	ψ	δ	E_c	Predicted property	Input property
1	0.001	0.000	0.001	2.2%	HV	CE
HV,CE	0.001	0.001	0.003	2.5%	CE	HV
2	0.003	0.002	0.040	1.6%	AR	MV
AR,MV	0.003	0.004	0.024	5.6%	MV	AR
3	0.005	0.005	0.024	6.9%	CE	BP
CE,BP	0.005	0.005	0.026	6.9%	BP	CE
4	0.006	0.005	0.024	6.9%	BP	HV
BP,HV	0.006	0.006	0.024	7.5%	HV	BP
5	0.009	0.009	0.009	9.3%	SM	YM
SM,YM	0.009	0.010	0.004	9.9%	YM	SM
6	0.011	0.002	0.013	4.3%	AW	SHC
AW,SHC	0.011	0.013	0.104	4.7%	SHC	AW
7	0.016	0.010	0.028	9.4%	TC	ER
TC,ER	0.016	0.019	0.007	13.8%	ER	TC
8	0.044	0.044	0.087	19.1%	BM	SE
BM,SE	0.044	0.044	0.088	19.1%	SE	BM
9	0.047	0.032	0.078	16.1%	HF	MP
HF,MP	0.047	0.056	0.063	22.9%	MP	HF
10	0.049	0.021	0.058	13.3%	CE	MP
CE,MP	0.049	0.060	0.048	24.1%	MP	CE
11	0.050	0.053	0.225	5.0%	WF	MV
WF,MV	0.050	0.046	0.092	19.3%	MV	WF
12	0.061	0.042	0.116	16.9%	CE	SE
CE,SE	0.061	0.073	0.097	25.2%	SE	CE
13	0.063	0.034	0.085	16.5%	HV	HF
HV,HF	0.063	0.077	0.145	23.7%	HF	HV
14	0.064	0.033	0.109	14.7%	CE	HF
CE,HF	0.064	0.078	0.145	23.9%	HF	CE
15	0.077	0.019	0.044	13.2%	SE	MP

Rank /Property group	ANNs evaluation parameters				Conditions	
	ψ_2	ψ	δ	E_c	Predicted property	Input property
SE,MP	0.077	0.094	0.088	29.3%	MP	SE
16	0.078	0.033	0.175	5.2%	AR	WF
AR,WF	0.078	0.095	0.303	5.7%	WF	AR
17	0.078	0.074	0.186	19.8%	HV	SE
HV,SE	0.078	0.081	0.160	23.6%	SE	HV
18	0.089	0.031	0.079	15.8%	HV	MP
HV,MP	0.089	0.109	0.077	32.1%	MP	HV
19	0.094	0.053	0.169	15.6%	BP	SE
BP,SE	0.094	0.114	0.131	31.1%	SE	BP
20	0.096	0.080	0.160	23.4%	HF	SE
HF,SE	0.096	0.107	0.148	29.2%	SE	HF
21	0.099	0.084	0.222	18.7%	BP	TEC
BP,TEC	0.099	0.110	0.229	24.0%	TEC	BP
22	0.101	0.053	0.139	18.3%	MV	LPA
MV,LPA	0.101	0.123	0.336	10.1%	LPA	MV
23	0.111	0.108	0.249	21.5%	HV	TEC
HV,TEC	0.111	0.114	0.202	27.0%	TEC	HV
24	0.117	0.067	0.185	18.1%	BP	HF
BP,HF	0.117	0.142	0.186	32.7%	HF	BP
25	0.119	0.055	0.227	6.3%	AR	LPA
AR,LPA	0.119	0.145	0.361	12.2%	LPA	AR

Appendix IV

Table 4 The top 50 ternary order correlations obtained from the ANN combinatorial search. Abbreviations are listed in Appendix I.

Rank	ANNs evaluation parameters			Conditions	
	ψ	δ	E_c	Predicted property	Input property
1	0.000	0.001	1.1%	SHC	CE,AW
2	0.000	0.003	1.2%	AW	SHC,DT
3	0.000	0.000	1.3%	SHC	SE,AW
4	0.000	0.007	1.3%	AR	MP,MV
5	0.000	0.000	1.5%	SHC	SM,AW
6	0.000	0.002	1.7%	CE	ER,HV
7	0.000	0.000	1.8%	SHC	MV,AW
8	0.000	0.003	1.9%	AW	SHC,YM
9	0.000	0.003	1.9%	CE	HF,HV
10	0.000	0.003	1.9%	HV	CE,MV
11	0.000	0.003	1.9%	HV	HF,CE
12	0.000	0.003	2.0%	AW	SHC,HF
13	0.000	0.001	2.1%	AW	SHC,SE
14	0.000	0.002	2.1%	CE	MP,HV
15	0.000	0.001	2.2%	CE	MV,HV
16	0.000	0.002	2.2%	CE	DT,HV
17	0.000	0.002	2.2%	CE	SE,HV
18	0.000	0.002	2.2%	SHC	AR,AW
19	0.000	0.003	2.2%	AW	SHC,TEC
20	0.001	0.000	2.3%	SHC	HF,AW
21	0.001	0.003	2.3%	AW	MP,SHC
22	0.001	0.003	2.3%	CE	BM,HV
23	0.001	0.004	2.3%	HV	MP,CE
24	0.001	0.004	2.3%	HV	YM,CE
25	0.001	0.005	2.3%	AW	SHC,CE
26	0.001	0.006	2.3%	HV	CE,SE
27	0.001	0.006	2.4%	CE	SM,HV
28	0.001	0.003	2.5%	HV	AR,CE
29	0.001	0.004	2.5%	CE	NACS,HV
30	0.001	0.004	2.5%	HV	BM,CE

Rank	ANNs evaluation parameters			Conditions	
	ψ	δ	E_c	Predicted property	Input property
31	0.001	0.005	2.5%	SHC	NSCS,AW
32	0.001	0.006	2.5%	HV	CE,TDR
33	0.001	0.010	2.4%	HV	SM,CE
34	0.001	0.003	2.6%	CE	NSCS,HV
35	0.001	0.003	2.6%	HV	ER,CE
36	0.001	0.003	2.7%	CE	WF,HV
37	0.001	0.004	2.7%	CE	HV,BP
38	0.001	0.005	2.7%	AW	SHC,SM
39	0.001	0.009	2.6%	CE	TC,HV
40	0.001	0.002	2.8%	CE	AR,HV
41	0.001	0.002	2.8%	CE	YM,HV
42	0.001	0.003	2.8%	CE	LPA,HV
43	0.001	0.013	2.5%	AW	SHC,LPA
44	0.001	0.005	2.8%	CE	TDR,HV
45	0.001	0.006	2.8%	HV	CE,NSCS
46	0.001	0.002	2.9%	SHC	AW,HV
47	0.001	0.008	2.8%	HV	CE,BP
48	0.001	0.005	2.9%	HV	CE,NACS
49	0.001	0.006	2.9%	CE	TEC,HV
50	0.001	0.001	3.0%	HV	CE,WF

Appendix V

Table 5 The top 25 ternary order correlation-groups obtained by ANN combinatorial search. Abbreviations are listed in Appendix I.

Rank /Property group	ANNs evaluation parameters				Conditions	
	ψ_3	ψ	δ	E_c	Predicted property	Input property
1	0.004	0.001	0.004	2.7%	CE	HV,BP
CE,HV,BP	0.004	0.001	0.008	2.8%	HV	CE,BP
	0.004	0.006	0.037	6.9%	BP	CE,HV
2	0.009	0.007	0.040	7.1%	BP	TEC,CE
BP,TEC,CE	0.009	0.013	0.006	11.3%	TEC	CE,BP
	0.009	0.005	0.040	5.8%	CE	TEC,BP
3	0.014	0.005	0.016	6.8%	SE	CE,MV
SE,CE,MV	0.014	0.020	0.053	13.2%	CE	SE,MV
	0.014	0.007	0.027	7.8%	MV	CE,SE
4	0.015	0.018	0.051	12.4%	HV	SE,MV
HV,SE,MV	0.015	0.014	0.017	11.6%	SE	MV,HV
	0.015	0.014	0.039	11.1%	MV	SE,HV
5	0.016	0.006	0.072	3.6%	WF	BM,SE
WF,BM,SE	0.016	0.024	0.042	14.9%	BM	SE,WF
	0.016	0.009	0.020	9.4%	SE	BM,WF
6	0.016	0.011	0.035	10.0%	MV	SE,BP
MV,SE,BP	0.016	0.011	0.049	9.4%	SE	MV,BP
	0.016	0.023	0.085	12.4%	BP	SE,MV
7	0.019	0.015	0.029	11.9%	SE	AR,HV
SE,AR,HV	0.019	0.021	0.140	3.0%	AR	SE,HV
	0.019	0.020	0.066	12.7%	HV	AR,SE
8	0.021	0.006	0.009	7.5%	BM	SM,PR
BM,SM,PR	0.021	0.032	0.008	18.0%	SM	BM,PR
	0.021	0.007	0.078	2.1%	PR	BM,SM
9	0.027	0.008	0.021	8.8%	SE	MP,MV
SE,MP,MV	0.027	0.040	0.047	19.5%	MP	SE,MV
	0.027	0.011	0.025	10.0%	MV	MP,SE
10	0.028	0.003	0.014	5.1%	CE	HF,BP
CE,HF,BP	0.028	0.042	0.096	18.1%	HF	CE,BP
	0.028	0.005	0.043	6.0%	BP	HF,CE

Rank /Property group	ANNs evaluation parameters				Conditions	
	ψ_3	ψ	δ	E_c	Predicted property	Input property
11	0.028	0.006	0.029	7.0%	HV	TEC,BP
HV,TEC,BP	0.028	0.043	0.078	19.3%	TEC	HV,BP
	0.028	0.007	0.040	7.4%	BP	TEC,HV
12	0.032	0.016	0.057	11.2%	MV	TEC,SE
MV,TEC,SE	0.032	0.045	0.120	17.6%	TEC	SE,MV
	0.032	0.023	0.068	13.4%	SE	TEC,MV
13	0.035	0.011	0.003	10.5%	SM	YM,SE
SM,YM,SE	0.035	0.012	0.006	11.1%	YM	SM,SE
	0.035	0.053	0.104	20.5%	SE	YM,SM
14	0.040	0.009	0.033	8.9%	SE	TEC,AR
SE,TEC,AR	0.040	0.060	0.136	20.3%	TEC	AR,SE
	0.040	0.019	0.132	3.9%	AR	TEC,SE
15	0.040	0.008	0.007	8.8%	YM	TEC,SM
YM,TEC,SM	0.040	0.061	0.089	23.0%	TEC	YM,SM
	0.040	0.013	0.017	11.4%	SM	TEC,YM
16	0.044	0.012	0.007	10.8%	BM	YM,PR
BM,YM,PR	0.044	0.034	0.009	18.3%	YM	BM,PR
	0.044	0.062	0.243	5.4%	PR	BM,YM
17	0.044	0.022	0.083	12.2%	HF	MP,TEC
HF,MP,TEC	0.044	0.065	0.088	24.0%	MP	HF,TEC
	0.044	0.005	0.004	6.7%	TEC	MP,HF
18	0.044	0.025	0.051	14.8%	MP	HF,BP
MP,HF,BP	0.044	0.039	0.113	16.3%	HF	MP,BP
	0.044	0.058	0.159	18.2%	BP	MP,HF
19	0.044	0.046	0.012	21.4%	DT	YM,AW
DT,YM,AW	0.044	0.047	0.015	21.6%	YM	DT,AW
	0.044	0.039	0.159	11.9%	AW	YM,DT
20	0.045	0.005	0.032	6.1%	CE	MV,BP
CE,MV,BP	0.045	0.068	0.128	22.7%	MV	CE,BP
	0.045	0.007	0.042	7.1%	BP	CE,MV
21	0.046	0.014	0.109	4.8%	WF	BM,CE
WF,BM,CE	0.046	0.041	0.041	19.9%	BM	CE,WF
	0.046	0.063	0.166	18.8%	CE	BM,WF
22	0.047	0.040	0.118	16.3%	TEC	HF,HV
TEC,HF,HV	0.047	0.050	0.137	17.7%	HF	TEC,HV

Rank /Property group	ANNs evaluation parameters				Conditions	
	ψ_3	ψ	δ	E_c	Predicted property	Input property
	0.047	0.049	0.125	18.3%	HV	HF,TEC
23	0.047	0.030	0.089	15.0%	HF	TEC,CE
HF,TEC,CE	0.047	0.055	0.145	18.4%	TEC	HF,CE
	0.047	0.050	0.140	17.5%	CE	HF,TEC
24	0.047	0.022	0.073	13.0%	BP	AR,SE
BP,AR,SE	0.047	0.071	0.264	3.5%	AR	SE,BP
	0.047	0.018	0.028	13.2%	SE	AR,BP
25	0.048	0.000	0.002	2.1%	CE	MP,HV
CE,MP,HV	0.048	0.072	0.038	26.6%	MP	CE,HV
	0.048	0.001	0.004	2.3%	HV	MP,CE

Appendix VI

Table 6 The top 50 quaternary order correlations obtained from the ANN combinatorial search. Abbreviations are listed in Appendix I.

Rank	ANNs evaluation parameters			Conditions	
	ψ	δ	E_c	Predicted property	Input property
1	0.000	0.000	0.6%	SHC	HF,YM,AW
2	0.000	0.001	0.6%	SHC	ER,CE,AW
3	0.000	0.000	0.7%	SHC	BM,MV,AW
4	0.000	0.000	0.7%	SHC	CE,DT,AW
5	0.000	0.000	0.7%	SHC	DT,AW,HV
6	0.000	0.000	0.7%	SHC	SM,CE,AW
7	0.000	0.000	0.7%	SHC	YM,AW,BP
8	0.000	0.005	0.6%	AR	MP,LPA,MV
9	0.000	0.000	0.8%	SHC	BM,AW,HV
10	0.000	0.000	0.8%	SHC	BM,LPA,AW
11	0.000	0.000	0.8%	SHC	CE,MV,AW
12	0.000	0.000	0.8%	SHC	SE,WF,AW
13	0.000	0.000	0.8%	SHC	TEC,ER,AW
14	0.000	0.000	0.8%	SHC	YM,LPA,AW
15	0.000	0.001	0.8%	SHC	YM,SE,AW
16	0.000	0.007	0.5%	AR	HF,LPA,MV
17	0.000	0.000	0.9%	SHC	DT,MV,AW
18	0.000	0.000	0.9%	SHC	HF,AW,BP
19	0.000	0.000	0.9%	SHC	TEC,SE,AW
20	0.000	0.006	0.8%	AR	MP,MV,AW
21	0.000	0.000	1.0%	SHC	HF,CE,AW
22	0.000	0.000	1.0%	SHC	MP,TEC,AW
23	0.000	0.000	1.0%	SHC	MV,AW,HV
24	0.000	0.000	1.0%	SHC	SM,WF,AW
25	0.000	0.000	1.0%	SHC	TDR,NSCS,AW
26	0.000	0.000	1.0%	SHC	TEC,DT,AW
27	0.000	0.000	1.0%	SHC	YM,CE,AW
28	0.000	0.000	1.0%	SHC	YM,NSCS,AW
29	0.000	0.001	1.0%	CE	TDR,SE,HV
30	0.000	0.001	1.0%	SHC	NSCS,AW,BP

Rank	ANNs evaluation parameters			Conditions	
	ψ	δ	E_c	Predicted property	Input property
31	0.000	0.001	1.0%	SHC	PR,WF,AW
32	0.000	0.002	1.0%	AW	SHC,YM,PR
33	0.000	0.009	0.5%	AR	ER,LPA,MV
34	0.000	0.009	0.5%	AR	HF,TEC,MV
35	0.000	0.000	1.1%	SHC	AR,AW,HV
36	0.000	0.000	1.1%	SHC	MS,WF,AW
37	0.000	0.000	1.1%	SHC	MV,NSCS,AW
38	0.000	0.000	1.1%	SHC	SE,MV,AW
39	0.000	0.000	1.1%	SHC	TEC,SM,AW
40	0.000	0.001	1.1%	CE	ER,MS,HV
41	0.000	0.001	1.1%	SHC	AR,CE,AW
42	0.000	0.001	1.1%	SHC	SM,DT,AW
43	0.000	0.005	1.0%	MV	TEC,ER,AR
44	0.000	0.003	1.1%	AW	SHC,SE,HV
45	0.000	0.007	0.9%	AR	HF,ER,MV
46	0.000	0.006	1.0%	AR	ER,LPA,SE
47	0.000	0.000	1.2%	HV	CE,PR,SE
48	0.000	0.000	1.2%	SHC	BM,SE,AW
49	0.000	0.000	1.2%	SHC	TEC,YM,AW
50	0.000	0.000	1.2%	SHC	YM,DT,AW

Appendix VII

Table 7 The top 25 quaternary order correlation-groups obtained from the ANN combinatorial search. Abbreviations are listed in Appendix I.

Rank/Property group	ANNs evaluation parameters				Conditions	
	ψ_4	ψ	δ	E_c	Predicted property	Input property
1	0.001	0.000	0.006	1.4%	AW	SHC,CE,HV
AW,SHC,CE,HV	0.001	0.001	0.004	2.9%	HV	SHC,CE,AW
	0.001	0.001	0.005	3.1%	CE	SHC,AW,HV
	0.001	0.002	0.001	4.6%	SHC	CE,AW,HV
2	0.002	0.000	0.000	1.7%	SHC	AR,MV,AW
SHC,AR,MV,AW	0.002	0.001	0.005	2.4%	AW	SHC,AR,MV
	0.002	0.001	0.034	1.0%	AR	SHC,MV,AW
	0.002	0.004	0.029	5.4%	MV	SHC,AR,AW
3	0.004	0.001	0.003	2.6%	HV	AR,CE,MV
HV,AR,CE,MV	0.004	0.001	0.002	3.3%	CE	AR,MV,HV
	0.004	0.002	0.046	1.7%	AR	CE,MV,HV
	0.004	0.007	0.040	7.3%	MV	AR,CE,HV
4	0.004	0.000	0.002	2.1%	HV	YM,SM,CE
HV,YM,SM,CE	0.004	0.001	0.004	2.9%	CE	YM,SM,HV
	0.004	0.005	0.002	6.8%	YM	SM,CE,HV
	0.004	0.006	0.008	8.0%	SM	YM,CE,HV
5	0.004	0.001	0.002	3.5%	SHC	CE,AW,BP
SHC,CE,AW,BP	0.004	0.001	0.003	3.5%	AW	SHC,CE,BP
	0.004	0.006	0.020	7.5%	CE	SHC,AW,BP
	0.004	0.006	0.038	6.6%	BP	SHC,CE,AW
6	0.006	0.003	0.051	1.9%	AR	CE,MV,BP
AR,CE,MV,BP	0.006	0.005	0.024	6.4%	CE	AR,MV,BP
	0.006	0.005	0.027	6.8%	BP	AR,CE,MV
	0.006	0.008	0.024	8.5%	MV	AR,CE,BP
7	0.006	0.001	0.003	2.4%	AW	SHC,HV,BP
AW,SHC,HV,BP	0.006	0.001	0.002	3.8%	SHC	AW,HV,BP
	0.006	0.006	0.033	6.9%	BP	SHC,AW,HV
	0.006	0.009	0.043	8.4%	HV	SHC,AW,BP
8	0.006	0.001	0.003	2.4%	HV	CE,SE,MV
HV,CE,SE,MV	0.006	0.001	0.004	2.4%	CE	SE,MV,HV

Rank/Property group	ANNs evaluation parameters				Conditions	
	ψ_4	ψ	δ	E_c	Predicted property	Input property
	0.006	0.007	0.027	7.6%	MV	CE,SE,HV
	0.006	0.009	0.014	9.3%	SE	CE,MV,HV
9	0.006	0.003	0.019	5.2%	BP	YM,SM,CE
BP,YM,SM,CE	0.006	0.003	0.027	5.2%	CE	YM,SM,BP
	0.006	0.006	0.006	7.8%	SM	YM,CE,BP
	0.006	0.010	0.004	9.9%	YM	SM,CE,BP
10	0.006	0.003	0.050	1.7%	AR	MV,HV,BP
AR,MV,HV,BP	0.006	0.006	0.041	6.7%	HV	AR,MV,BP
	0.006	0.008	0.031	8.3%	BP	AR,MV,HV
	0.006	0.007	0.042	7.5%	MV	AR,HV,BP
11	0.007	0.006	0.005	7.9%	SM	YM,HV,BP
SM,YM,HV,BP	0.007	0.004	0.024	6.1%	BP	YM,SM,HV
	0.007	0.005	0.026	6.6%	HV	YM,SM,BP
	0.007	0.009	0.007	9.6%	YM	SM,HV,BP
12	0.007	0.003	0.014	5.0%	CE	SE,MV,BP
CE,SE,MV,BP	0.007	0.006	0.032	6.7%	BP	CE,SE,MV
	0.007	0.007	0.019	8.0%	SE	CE,MV,BP
	0.007	0.010	0.026	9.4%	MV	CE,SE,BP
13	0.007	0.005	0.026	6.5%	HV	SE,MV,BP
HV,SE,MV,BP	0.007	0.007	0.019	8.2%	SE	MV,HV,BP
	0.007	0.008	0.028	8.5%	MV	SE,HV,BP
	0.007	0.009	0.041	8.4%	BP	SE,MV,HV
14	0.008	0.006	0.024	7.1%	CE	TC,ER,BP
CE,TC,ER,BP	0.008	0.006	0.038	6.8%	BP	TC,ER,CE
	0.008	0.009	0.018	9.5%	ER	TC,CE,BP
	0.008	0.009	0.020	9.5%	TC	ER,CE,BP
15	0.008	0.002	0.037	1.5%	PR	BM,YM,SM
PR,BM,YM,SM	0.008	0.004	0.005	6.6%	SM	BM,YM,PR
	0.008	0.010	0.004	9.8%	YM	BM,SM,PR
	0.008	0.011	0.012	10.6%	BM	YM,SM,PR
16	0.008	0.002	0.016	4.6%	SE	TEC,CE,MV
SE,TEC,CE,MV	0.008	0.005	0.010	6.8%	MV	TEC,CE,SE
	0.008	0.006	0.010	7.9%	TEC	CE,SE,MV
	0.008	0.014	0.034	11.2%	CE	TEC,SE,MV
17	0.011	0.002	0.016	4.6%	MV	AR,CE,SE

Rank/Property group	ANNs evaluation parameters				Conditions	
	ψ_4	ψ	δ	E_c	Predicted property	Input property
MV,AR,CE,SE	0.011	0.002	0.045	1.8%	AR	CE,SE,MV
	0.011	0.012	0.030	10.5%	SE	AR,CE,MV
	0.011	0.018	0.064	11.7%	CE	AR,SE,MV
18	0.012	0.004	0.031	5.6%	CE	TEC,MV,BP
CE,TEC,MV,BP	0.012	0.005	0.034	6.1%	BP	TEC,CE,MV
	0.012	0.011	0.007	10.4%	TEC	CE,MV,BP
	0.012	0.019	0.040	13.0%	MV	TEC,CE,BP
19	0.012	0.005	0.019	6.9%	HV	TC,ER,BP
HV,TC,ER,BP	0.012	0.007	0.035	7.3%	BP	TC,ER,HV
	0.012	0.010	0.012	9.9%	TC	ER,HV,BP
	0.012	0.020	0.009	14.2%	ER	TC,HV,BP
20	0.013	0.002	0.009	3.9%	MV	TEC,SE,HV
MV,TEC,SE,HV	0.013	0.007	0.002	8.4%	TEC	SE,MV,HV
	0.013	0.006	0.021	7.6%	SE	TEC,MV,HV
	0.013	0.023	0.082	12.6%	HV	TEC,SE,MV
21	0.013	0.003	0.018	5.2%	CE	HF,TEC,BP
CE,HF,TEC,BP	0.013	0.005	0.014	7.0%	BP	HF,TEC,CE
	0.013	0.011	0.010	10.2%	TEC	HF,CE,BP
	0.013	0.022	0.055	13.8%	HF	TEC,CE,BP
22	0.013	0.010	0.008	10.0%	TEC	HF,SE,MV
TEC,HF,SE,MV	0.013	0.010	0.016	9.7%	MV	HF,TEC,SE
	0.013	0.016	0.033	12.1%	SE	HF,TEC,MV
	0.013	0.016	0.049	11.7%	HF	TEC,SE,MV
23	0.014	0.000	0.002	2.2%	CE	SE,WF,HV
CE,SE,WF,HV	0.014	0.001	0.004	2.5%	HV	CE,SE,WF
	0.014	0.007	0.036	7.4%	SE	CE,WF,HV
	0.014	0.024	0.146	5.0%	WF	CE,SE,HV
24	0.014	0.004	0.055	3.4%	AR	TEC,SE,HV
AR,TEC,SE,HV	0.014	0.012	0.015	10.9%	TEC	AR,SE,HV
	0.014	0.013	0.043	10.6%	SE	TEC,AR,HV
	0.014	0.020	0.057	12.9%	HV	TEC,AR,SE
25	0.014	0.000	0.003	2.0%	CE	TEC,SE,HV
CE,TEC,SE,HV	0.014	0.001	0.004	2.4%	HV	TEC,CE,SE
	0.014	0.020	0.031	13.9%	TEC	CE,SE,HV
	0.014	0.017	0.067	11.2%	SE	TEC,CE,HV

Appendix VIII

Table 8 Poisson's ratio of the 68 elements in the source pool.

Symbol	[328]	[229]	[454]	[455]	[251]	[456]	[228]	[348]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Ac							0.269					0.269	0.269	0.269	0.269	0.0000	0.269	0.00%
Ag	0.37		0.37	0.367	0.367	0.37	0.37	0.37	0.37	0.367	0.37	0.370	0.367	0.369	0.370	0.0000	0.370	0.82%
Al	0.35		0.35	0.345	0.345		0.34	0.33	0.34	0.345	0.346	0.350	0.330	0.343	0.345	0.0000	0.345	6.06%
Au	0.44		0.44	0.44	0.42	0.42	0.425	0.42	0.42	0.42	0.42	0.440	0.420	0.427	0.420	0.0001	0.420	4.76%
Ba					0.28		0.28			0.28	0.229	0.280	0.229	0.267	0.280	0.0007	0.280	22.27%
Be	0.032		0.032		0.02		0.08		0.032	0.02	0.118	0.118	0.020	0.048	0.032	0.0014	0.032	490.00%
Bi	0.33		0.33	0.33	0.33		0.33		0.33	0.33	0.335	0.335	0.330	0.331	0.330	0.0000	0.330	1.52%
Ca	0.31		0.31		0.31		0.31		0.31	0.31	0.307	0.310	0.307	0.310	0.310	0.0000	0.310	0.98%
Cd	0.3		0.3	0.3	0.3		0.3		0.39	0.3	0.303	0.390	0.300	0.312	0.300	0.0010	0.300	30.00%
Ce	0.24	0.24	0.24		0.248		0.248		0.248	0.248	0.242	0.248	0.240	0.244	0.245	0.0000	0.248	3.33%
Co	0.31		0.31		0.32		0.31		0.31	0.32	0.311	0.320	0.310	0.313	0.310	0.0000	0.310	3.23%
Cr	0.21		0.21	0.21	0.21		0.209			0.21	0.213	0.213	0.209	0.210	0.210	0.0000	0.210	1.91%
Cs					0.295		0.356			0.295	0.295	0.356	0.295	0.310	0.295	0.0009	0.295	20.68%
Cu	0.34		0.34	0.343	0.343		0.35	0.36	0.35	0.343	0.344	0.360	0.340	0.346	0.343	0.0000	0.343	5.88%
Dy	0.247	0.247	0.25		0.247		0.243		0.243		0.23	0.250	0.230	0.244	0.247	0.0000	0.247	8.70%
Er	0.237	0.237	0.24		0.237		0.238		0.238		0.233	0.240	0.233	0.237	0.237	0.0000	0.237	3.00%
Eu	0.152	0.152	0.15		0.152		0.286				0.152	0.286	0.150	0.174	0.152	0.0030	0.152	90.67%
Fe	0.29		0.29	0.293	0.27		0.28		0.28	0.293	0.288	0.293	0.270	0.286	0.289	0.0001	0.280	8.52%
Fr							0.356					0.356	0.356	0.356	0.356	0.0000	0.356	0.00%

Symbol	[328]	[229]	[454]	[455]	[251]	[456]	[228]	[348]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Ga	0.47				0.47		0.235			0.47	0.233	0.470	0.233	0.376	0.470	0.0167	0.470	101.72%
Gd	0.259	0.259	0.26		0.259		0.259		0.259		0.256	0.260	0.256	0.259	0.259	0.0000	0.259	1.56%
Hf	0.37		0.37		0.26		0.3		0.37	0.26	0.282	0.370	0.260	0.316	0.300	0.0027	0.370	42.31%
Ho	0.231	0.231	0.23		0.231		0.255		0.255		0.272	0.272	0.230	0.244	0.231	0.0003	0.231	18.26%
In					0.45		0.46			0.45	0.445	0.460	0.445	0.451	0.450	0.0000	0.450	3.37%
Ir	0.26		0.26		0.26		0.26		0.26	0.26	0.26	0.260	0.260	0.260	0.260	0.0000	0.260	0.00%
K					0.35		0.35				0.375	0.375	0.350	0.358	0.350	0.0002	0.350	7.14%
La	0.28	0.28	0.28		0.28		0.288		0.288	0.28	0.173	0.288	0.173	0.269	0.280	0.0015	0.280	66.47%
Li					0.36		0.362			0.36	0.356	0.362	0.356	0.360	0.360	0.0000	0.360	1.69%
Lu	0.261	0.261	0.26		0.261		0.233				0.26	0.261	0.233	0.256	0.261	0.0001	0.261	12.02%
Mg	0.29		0.29	0.291	0.291	0.35	0.3	0.35	0.28	0.291	0.29	0.350	0.280	0.302	0.291	0.0007	0.290	25.00%
Mn			0.24		0.24		0.24			0.24	0.156	0.240	0.156	0.223	0.240	0.0014	0.240	53.85%
Mo	0.31		0.31		0.293		0.3	0.32	0.31	0.293	0.295	0.320	0.293	0.304	0.305	0.0001	0.310	9.22%
Na					0.34		0.315			0.34	0.366	0.366	0.315	0.340	0.340	0.0004	0.340	16.19%
Nb	0.4		0.4	0.397	0.397		0.38		0.39	0.397	0.397	0.400	0.380	0.395	0.397	0.0000	0.397	5.26%
Nd	0.281	0.281	0.28		0.281		0.306		0.306		0.288	0.306	0.280	0.289	0.281	0.0001	0.281	9.29%
Ni	0.31		0.31	0.312	0.312		0.31	0.31	0.32	0.312	0.297	0.320	0.297	0.310	0.310	0.0000	0.310	7.74%
Np							0.255					0.255	0.255	0.255	0.255	0.0000	0.255	0.00%
Os	0.25		0.25		0.25		0.285		0.25	0.25	0.255	0.285	0.250	0.256	0.250	0.0002	0.250	14.00%
Pa							0.282					0.282	0.282	0.282	0.282	0.0000	0.282	0.00%
Pb	0.44		0.44	0.44	0.44		0.44		0.44	0.44	0.409	0.440	0.409	0.436	0.440	0.0001	0.440	7.58%
Pd	0.39		0.39		0.39		0.375		0.39	0.39	0.39	0.390	0.375	0.388	0.390	0.0000	0.390	4.00%
Pm	0.28	0.28	0.28		0.28		0.278					0.280	0.278	0.280	0.280	0.0000	0.280	0.72%
Pr	0.281	0.281	0.28		0.281		0.305		0.305		0.289	0.305	0.280	0.289	0.281	0.0001	0.281	8.93%

Symbol	[328]	[229]	[454]	[455]	[251]	[456]	[228]	[348]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Pt	0.38		0.38	0.377	0.39	0.39	0.39	0.39	0.44	0.39	0.39	0.440	0.377	0.392	0.390	0.0003	0.390	16.71%
Pu	0.21		0.21		0.18		0.15		0.21	0.18		0.210	0.150	0.190	0.195	0.0006	0.210	40.00%
Ra							0.304					0.304	0.304	0.304	0.304	0.0000	0.304	0.00%
Rb					0.3		0.356			0.3	0.374	0.374	0.300	0.333	0.328	0.0015	0.300	24.67%
Re	0.3		0.3		0.26		0.293		0.296	0.26	0.289	0.300	0.260	0.285	0.293	0.0003	0.260	15.38%
Rh	0.26		0.26		0.26		0.27		0.26	0.26	0.26	0.270	0.260	0.261	0.260	0.0000	0.260	3.85%
Ru	0.3		0.3		0.25		0.286		0.296	0.25	0.269	0.300	0.250	0.279	0.286	0.0005	0.250	20.00%
Sc	0.279	0.279	0.28		0.279		0.269		0.309		0.258	0.309	0.258	0.279	0.279	0.0002	0.279	19.77%
Sm	0.274	0.274	0.27		0.274		0.352		0.352		0.268	0.352	0.268	0.295	0.274	0.0015	0.274	31.34%
Sn	0.36		0.36	0.357	0.357	0.33	0.33	0.33	0.33	0.357	0.354	0.360	0.330	0.347	0.356	0.0002	0.330	9.09%
Sr	0.28		0.28		0.28		0.304		0.28	0.28	0.305	0.305	0.280	0.287	0.280	0.0001	0.280	8.93%
Ta	0.34		0.34	0.342	0.342		0.35	0.35	0.35	0.342	0.34	0.350	0.340	0.344	0.342	0.0000	0.340	2.94%
Tb	0.261	0.261	0.26		0.261		0.261		0.261		0.263	0.263	0.260	0.261	0.261	0.0000	0.261	1.15%
Tc							0.293				0.309	0.309	0.293	0.301	0.301	0.0001	0.293	5.46%
Th	0.27		0.27		0.26		0.3	0.27	0.265	0.26	0.286	0.300	0.260	0.273	0.270	0.0002	0.270	15.38%
Ti	0.32		0.32	0.321	0.361		0.34	0.3	0.36	0.361	0.321	0.361	0.300	0.334	0.321	0.0005	0.320	20.33%
Tl	0.45		0.45		0.45		0.46		0.45	0.45	0.428	0.460	0.428	0.448	0.450	0.0001	0.450	7.48%
Tm	0.213	0.213	0.21		0.213		0.235				0.217	0.235	0.210	0.217	0.213	0.0001	0.213	11.90%
U	0.23		0.23		0.2		0.24	0.21	0.23	0.2	0.201	0.240	0.200	0.218	0.220	0.0003	0.230	20.00%
V	0.37		0.37	0.365	0.365		0.36		0.35	0.365	0.361	0.370	0.350	0.363	0.365	0.0000	0.365	5.71%
W	0.28		0.28	0.28	0.28		0.284	0.28	0.3	0.28	0.28	0.300	0.280	0.283	0.280	0.0000	0.280	7.14%
Y	0.243	0.243	0.24		0.265		0.258		0.265	0.265	0.246	0.265	0.240	0.253	0.252	0.0001	0.265	10.42%
Yb	0.207	0.207	0.21		0.207		0.284		0.284		0.199	0.284	0.199	0.228	0.207	0.0015	0.207	42.71%
Zn	0.25		0.25	0.249	0.249		0.35	0.25	0.27	0.249	0.257	0.350	0.249	0.264	0.250	0.0011	0.249	40.56%

Symbol	[328]	[229]	[454]	[455]	[251]	[456]	[228]	[348]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Zr	0.34		0.34		0.38		0.34			0.38	0.332	0.380	0.332	0.352	0.340	0.0005	0.340	14.46%

Appendix IX

Table 9 Bulk moduli (GPa) of the 68 elements in the source pool.

Symbol	[328]	[229]	[340]	[322]	[455]	[251]	[228]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Ac							25				25.00	25.00	25.00	25.00	0.00	25.00	0.00%
Ag	100		103.6	100	103.6	103.6	101	96.2	103.6	95.4	103.60	95.40	100.78	101.00	10.37	103.60	8.60%
Al	76		76	76	75.5	75.2	72.2	65.8	75.2	76	76.00	65.80	74.21	75.50	11.40	76.00	15.50%
Au	180		171	220	217	171	173	142.9	171	16.08	220.00	16.08	162.44	171.00	3591.59	171.00	1268.16%
Ba	9.6		9.6	9.6			10.3	9.6		9.57	10.30	9.57	9.71	9.60	0.08	9.60	7.63%
Be	130		110	130		110	100	126.6	110	125.57	130.00	100.00	117.77	117.79	133.55	110.00	30.00%
Bi	31		31	31	31.3		31.5	35		33.6	35.00	31.00	32.06	31.30	2.55	31.00	12.90%
Ca	17		17	17		17.2	15.2	16.9	17.2	16.7	17.20	15.20	16.78	17.00	0.43	17.00	13.16%
Cd	42		42	42	41.6	51	46.7	46.7	51	52.9	52.90	41.60	46.21	46.70	20.69	42.00	27.16%
Ce	21.5	21.5	22	22			26.2	19.8		21.7	26.20	19.80	22.10	21.70	3.83	21.50	32.32%
Co	180		181.5	180		181.5	191	181.8	181.5	190	191.00	180.00	183.41	181.50	19.69	181.50	6.11%
Cr	160		160.2	160	160.1	160.2	190	108.7	160.2	162	190.00	108.70	157.93	160.20	437.17	160.20	74.79%
Cs	1.6		1.6	1.6			2.03	2.7		1.4	2.70	1.40	1.82	1.60	0.23	1.60	92.86%
Cu	140		137.8	140	137.8	137.8	131	129.9	137.8	137	140.00	129.90	136.57	137.80	13.16	137.80	7.78%
Dy	40.5	40.5	40.5	41		40.5	38.4	41		38.9	41.00	38.40	40.16	40.50	0.94	40.50	6.77%
Er	44.4	44.4	44.4	44		44.4	41.1	46.5		45.5	46.50	41.10	44.34	44.40	2.38	44.40	13.14%
Eu	8.3	8.3	8.3	8.3		8.3	14.7	13.1		8.7	14.70	8.30	9.75	8.30	6.76	8.30	77.11%
Fe	170		170	170	169.8	160	168	161.3	169.8	166	170.00	160.00	167.21	169.80	15.72	170.00	6.25%
Fr							2				2.00	2.00	2.00	2.00	0.00	2.00	0.00%
Ga							35	42.6		58.2	58.20	35.00	45.27	42.60	139.89	35.00	66.29%
Gd	37.9	37.9	38	38		37.9	38.3	38.9		38	38.90	37.90	38.11	38.00	0.12	37.90	2.64%

Symbol	[328]	[229]	[340]	[322]	[455]	[251]	[228]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Hf	110		109	110		109	109	108.7	109	109	110.00	108.70	109.21	109.00	0.25	109.00	1.20%
Ho	40.2	40.2	40.2	40		40.2	39.7	45.9		48.9	48.90	39.70	41.91	40.20	12.14	40.20	23.17%
In							41.1	36.4		42	42.00	36.40	39.83	41.10	9.04	36.40	15.38%
Ir	320		371	320		371	383	322.6	371	359	383.00	320.00	352.20	365.00	715.01	371.00	19.69%
K	3.1		3.1	3.1			3.18	4.2		3.3	4.20	3.10	3.33	3.14	0.19	3.10	35.48%
La	27.9	27.9	28	28			24.3	30.3		16.2	30.30	16.20	26.09	27.90	22.10	27.90	87.04%
Li	11		11	11			11.6	11.4		12.1	12.10	11.00	11.35	11.20	0.20	11.00	10.00%
Lu	47.6	47.6	47.6	48		47.6	41.1	42.6		42.7	48.00	41.10	45.60	47.60	8.49	47.60	16.79%
Mg	45		35.6	45	35.6	35.6	35.4	33.8	35.6	35.4	45.00	33.80	37.44	35.60	18.68	35.60	33.14%
Mn	120		120	120			59.6	116.3		92.6	120.00	59.60	104.75	118.15	603.45	120.00	101.34%
Mo	230		261.2	230		261.2	273	217.4	261.2	263	273.00	217.40	249.63	261.20	419.42	261.20	25.57%
Na	6.3		6.3	6.3			6.81	6.9		6.74	6.90	6.30	6.56	6.52	0.08	6.30	9.52%
Nb	170		170.3	170	170.3	170.3	170	172.4	170.3	171	172.40	170.00	170.51	170.30	0.60	170.30	1.41%
Nd	31.8	31.8	31.8	32		31.8	32.7	32.5		32	32.70	31.80	32.05	31.90	0.13	31.80	2.83%
Ni	180		180	180	177.3	177.3	186	185.2	177.3	183	186.00	177.30	180.68	180.00	11.22	177.30	4.91%
Np							68				68.00	68.00	68.00	68.00	0.00	68.00	0.00%
Os	462		373			373	462		373	379.9	462.00	373.00	403.82	376.45	2038.32	373.00	23.86%
Pa							76				76.00	76.00	76.00	76.00	0.00	76.00	0.00%
Pb	46		46	46	45.8	45.8	42.9	41.5	45.8	44.7	46.00	41.50	44.94	45.80	2.71	45.80	10.84%
Pd	180		187	180		187	181	185.2	187	174.5	187.00	174.50	182.71	183.10	20.92	187.00	7.16%
Pm	33	33	33	33		33	35				35.00	33.00	33.33	33.00	0.67	33.00	6.06%
Pr	28.8	28.8	28.8	29		28.8	30.6	29.9		28.4	30.60	28.40	29.14	28.80	0.53	28.80	7.75%
Pt	230		276	230	228	276	278	263.2	276	259.1	278.00	228.00	257.37	263.20	483.08	276.00	21.93%
Pu							54				54.00	54.00	54.00	54.00	0.00	54.00	0.00%

Symbol	[328]	[229]	[340]	[322]	[455]	[251]	[228]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Ra							13.2				13.20	13.20	13.20	13.20	0.00	13.20	0.00%
Rb	2.5		2.5	2.5			3.14	3		2.3	3.14	2.30	2.66	2.50	0.11	2.50	36.52%
Re	370		334	370		334	372	370.4	334	365	372.00	334.00	356.18	367.50	341.16	334.00	11.38%
Rh	275		276	380		276	271	270.3	276	221.2	380.00	221.20	280.69	275.50	1957.39	276.00	71.79%
Ru	220		286	220		286	348	384.6	286	298.5	384.60	220.00	291.14	286.00	3174.72	286.00	74.82%
Sc	56.6	56.6	56.6	57		56.6	57.3	67.1		55.1	67.10	55.10	57.86	56.60	14.34	56.60	21.78%
Sm	37.8	37.8	37.8	38		37.8	29.4	38.3		35.7	38.30	29.40	36.58	37.80	9.04	37.80	30.27%
Sn	58		58.2	58	58.2	58.2	111	51.5	58.2	55.4	111.00	51.50	62.97	58.20	329.52	58.20	115.53%
Sr			12			12	11.6	12	12	11.6	12.00	11.60	11.87	12.00	0.04	12.00	3.45%
Ta	200		196.3	200	196.3	196.3	200	188.7	196.3	193	200.00	188.70	196.32	196.30	13.71	196.30	5.99%
Tb	38.7	38.7	38.7	38.7		38.7	39.9	40		40.7	40.70	38.70	39.26	38.70	0.66	38.70	5.17%
Tc							297			281	297.00	281.00	289.00	289.00	128.00	281.00	5.69%
Th	54		54	54		54	54.3	53.8	54	57.7	57.70	53.80	54.48	54.00	1.72	54.00	7.25%
Ti	110		108.4	110	107.7	108.4	105	123.5	108.4	107	123.50	105.00	109.82	108.40	28.62	108.40	17.62%
Tl	43		28.5	43		28.5	35.9	44.6	28.5	35.7	44.60	28.50	35.96	35.80	48.69	28.50	56.49%
Tm	44.5	44.5	44.5	45		44.5	36.7	37.7		43.6	45.00	36.70	42.63	44.50	11.43	44.50	22.62%
U	100		100	100		97.9	98.7	101	97.9	113	113.00	97.90	101.06	100.00	24.50	100.00	15.42%
V	160		158	160	158	158	162	161.3	158	157	162.00	157.00	159.14	158.00	3.01	158.00	3.18%
W	310		311	310	311	311	323	333.3	311	310	333.30	310.00	314.48	311.00	66.93	311.00	7.52%
Y	41.2	41.2	41	41			36.6	46.9		41.5	46.90	36.60	41.34	41.20	8.95	41.00	28.14%
Yb	30.5	30.5	30.5	31		30.5	13.3	13.8		13.2	31.00	13.20	24.16	30.50	78.99	30.50	134.85%
Zn	70		69.4	70	72	69.4	59.8	57.8	69.4	68.3	72.00	57.80	67.34	69.40	24.67	69.40	24.57%
Zr	91.1		89.8			89.8	83.3	89.3	89.8	95.3	95.30	83.30	89.77	89.80	12.41	89.80	14.41%

Appendix X

Table 10 Young's moduli (GPa) of the 68 elements in the source pool.

Symbol	[328]	[229]	[340]	[457]	[455]	[251]	[456]	[228]	[348]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Ac								34					34.00	34.00	34.00	34.00	0.00	34.00	0.00%
Ag	83		82.7	83	82.7	82.7		80	72.4	80	82.7	74.4	83.00	72.40	80.36	82.70	15.00	82.70	14.64%
Al	70		70	70	70.3	70.6		70.5	68.9	70.8	70.6	70.3	70.80	68.90	70.20	70.30	0.29	70.00	2.76%
Au	79		78.5	78	78	78.5	80	78.3	74.5	77.5	78.5	77.2	80.00	74.50	78.00	78.30	1.90	78.50	7.38%
Ba	13		12.8	13		12.8		13		12.6	12.8	15.58	15.58	12.60	13.20	12.90	0.95	12.80	23.65%
Be	287		318	287		318	294	301	289.6	287	318	287.25	318.00	287.00	298.69	291.80	196.60	287.00	10.80%
Bi	32		32	32	31.9	34		34	31.7	34.1	34	33.64	34.10	31.70	32.93	32.82	1.16	32.00	7.57%
Ca	20		20	20		19.6		21		19.6	19.6	19.34	21.00	19.34	19.89	19.80	0.26	19.60	8.58%
Cd	50		50	50	49.9	62.6		62	55.2	62.3	62.6	62.54	62.60	49.90	56.71	58.60	38.48	50.00	25.45%
Ce	33.6	33.6	34	34		33.5		30		30	33.5	33.6	34.00	30.00	32.87	33.60	2.68	33.60	13.33%
Co	209		211	209		211		208	206.8	209	211	215.2	215.20	206.80	210.00	209.00	5.91	209.00	4.06%
Cr	279		279	279	279.1	279		259	248.2	245	279	279	279.10	245.00	270.53	279.00	198.58	279.00	13.92%
Cs	1.7		1.7	1.7		1.7		1.8		1.7	1.7	1.72	1.80	1.70	1.72	1.70	0.00	1.70	5.88%
Cu			129.8	130	129.8	129.8		124	117.2	129	129.8	128.2	130.00	117.20	127.51	129.80	18.56	129.80	10.92%
Dy	61.4	61.4	61.4	61		61.4		63		63.1		63	63.10	61.00	61.96	61.40	0.81	61.40	3.44%
Er	69.9	69.9	69.9	70		69.9		73		73.3		73	73.30	69.90	71.11	69.95	2.72	69.90	4.86%
Eu	18.2	18.2	18	18		18.2		15				18.2	18.20	15.00	17.69	18.20	1.41	18.20	21.33%
Fe	211		211	211	211.4	208		211	196.5	211	211.4	211	211.40	196.50	209.33	211.00	21.30	211.00	7.58%
Fr								2					2.00	2.00	2.00	2.00	0.00	2.00	0.00%
Ga	9.8		9.81			9.81		9.81			9.81	93.2	93.20	9.80	23.71	9.81	1159.04	9.81	851.02%
Gd	54.8	54.8	55	55		54.8		55		56.2		55.68	56.20	54.80	55.16	55.00	0.26	54.80	2.55%
Hf	78		141	78		141		139		78.3	141	143	143.00	78.00	117.41	140.00	1060.90	141.00	83.33%

Symbol	[328]	[229]	[340]	[457]	[455]	[251]	[456]	[228]	[348]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Ho	64.8	64.8	64.8	65		64.8		72				66.9	72.00	64.80	66.16	64.80	7.23	64.80	11.11%
In	11		10.6	11		10.6		14		10.5	10.6	13.8	14.00	10.50	11.51	10.80	2.21	10.60	33.33%
Ir	528		528	528		528		533	517.1	528	528	517.1	533.00	517.10	526.13	528.00	28.91	528.00	3.07%
K	3.53		3.53			3.53		2.4				2.47	3.53	2.40	3.09	3.53	0.36	3.53	47.08%
La	36.6	36.6	37.9	37		37.9		50		37.5	37.9	31.9	50.00	31.90	38.14	37.50	23.28	37.90	56.74%
Li	4.9		4.91	4.9		4.91		10		4.9	4.91	10.44	10.44	4.90	6.23	4.91	6.07	4.90	113.06%
Lu	68.6	68.6	68.6	69		68.6		84				61.5	84.00	61.50	69.84	68.60	46.15	68.60	36.59%
Mg	45		44.7	45	44.7	44.7	44.8	44.4	44.1	44.3	44.7	44.63	45.00	44.10	44.64	44.70	0.08	44.70	2.04%
Mn	198		191	198		191		198	158.6	198	191	191	198.00	158.60	190.51	191.00	155.45	191.00	24.84%
Mo	329		324.8	329		324.8	340	322	275.8	329	324.8	323.7	340.00	275.80	322.29	324.80	292.33	324.80	23.28%
Na	10		6.8	10		6.8		5		10	6.8	5.41	10.00	5.00	7.60	6.80	4.39	6.80	100.00%
Nb	105		104.9	105	104.9	104.9		104	103.4	100	104.9	105	105.00	100.00	104.20	104.90	2.47	104.90	5.00%
Nd	41.4	41.4	41.4	41		41.4		38		37.9		40.7	41.40	37.90	40.40	41.20	2.35	41.40	9.23%
Ni	200		200	200	199.5	199.5		208	213.7	196	199.5	222.5	222.50	196.00	203.87	200.00	69.44	199.50	13.52%
Np								90					90.00	90.00	90.00	90.00	0.00	90.00	0.00%
Os			559			559		550	551.6		559	558.5	559.00	550.00	556.18	558.75	17.68	559.00	1.64%
Pa								100					100.00	100.00	100.00	100.00	0.00	100.00	0.00%
Pb	16		16	16	16.1	16.1		20	13.8	15.7	16.1	24.23	24.23	13.80	17.00	16.05	8.74	16.00	75.58%
Pd	121		121	121		121		127		121	121	115.2	127.00	115.20	121.03	121.00	9.95	121.00	10.24%
Pm	46	46	46	46		46		42					46.00	42.00	45.33	46.00	2.67	46.00	9.52%
Pr	37.3	37.3	37.3	37		37.3		50		35.2		36	50.00	35.20	38.43	37.30	22.48	37.30	42.05%
Pt	168		170	168	168	170	166.7	175	146.9	170	170	171	175.00	146.90	167.60	170.00	51.91	170.00	19.13%
Pu	96		96	96		87.5		97	96.5	96.1	87.5		97.00	87.50	94.08	96.00	16.59	96.00	10.86%
Ra								16					16.00	16.00	16.00	16.00	0.00	16.00	0.00%
Rb	2.4		2.35	2.4		2.35		2		2.35	2.35	1.73	2.40	1.73	2.24	2.35	0.06	2.35	38.73%
Re	463		466	463		466		461		463	466	461.4	466.00	461.00	463.68	463.00	4.27	463.00	1.08%

Symbol	[328]	[229]	[340]	[457]	[455]	[251]	[456]	[228]	[348]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Rh	380		379	275		379		330	289.6	275	379	318.5	380.00	275.00	333.90	330.00	2172.57	379.00	38.18%
Ru	447		432	447		432		430		447	432	413.7	447.00	413.70	435.09	432.00	133.95	432.00	8.05%
Sc	74.4	74.4	74.4	74		74.4		80		77		80	80.00	74.00	76.08	74.40	6.74	74.40	8.11%
Sm	49.7	49.7	49.7	50		49.7		45		34.1		49.7	50.00	34.10	47.20	49.70	30.79	49.70	46.63%
Sn	50		49.9	50	49.9	49.9		50	41.4	54	49.9	48.47	54.00	41.40	49.35	49.90	9.79	49.90	30.43%
Sr			15.7			15.7		15			15.7	13.57	15.70	13.57	15.13	15.70	0.86	15.70	15.70%
Ta	186		185.7	186	185.7	185.7		183	186.2	177	185.7	184.9	186.20	177.00	184.59	185.70	7.97	185.70	5.20%
Tb	55.7	55.7	55.7	56		55.7		57		57.5		57.98	57.98	55.70	56.41	55.85	0.88	55.70	4.09%
Tc								380				322	380.00	322.00	351.00	351.00	1682.00	322.00	18.01%
Th	79		78.3	79		78.3		73	58.6	79.2	78.3	73.81	79.20	58.60	75.28	78.30	44.50	78.30	35.15%
Ti	116		120.2	116	115.7	120.2	116	110	110.3	104	120.2	114.66	120.20	104.00	114.84	116.00	25.25	116.00	15.58%
Tl	8		7.9	8		7.9		12		7.95	7.9	15.42	15.42	7.90	9.38	7.98	7.97	7.90	95.19%
Tm	74	74	74	74		74		76				74	76.00	74.00	74.29	74.00	0.57	74.00	2.70%
U	208		208	208		175.8		186	165.5	208	175.8	220.7	220.70	165.50	195.09	208.00	378.12	208.00	33.35%
V	128		127.6	128	127.6	127.6		129	131	147	127.6	130.9	147.00	127.60	130.43	128.00	35.67	127.60	15.20%
W	411		411	411	411	411	340	401	344.7	407	411	409.6	411.00	340.00	397.12	411.00	743.49	411.00	20.88%
Y	63.5	63.5	66.3	64		66.3		64.4		66.4	66.3	63.29	66.40	63.29	64.89	64.40	1.96	66.30	4.91%
Yb	23.9	23.9	23.9	24		23.9		18				23.86	24.00	18.00	23.07	23.90	4.99	23.90	33.33%
Zn	108		104.5	108	108.4	104.5		95	82.7	92.2	104.5	99.3	108.40	82.70	100.71	104.50	71.10	104.50	31.08%
Zr	88		98	68		98		94		68.4	98	95.9	98.00	68.00	88.54	94.95	168.61	98.00	44.12%

Appendix XI

Table 11 Shear moduli (GPa) of the 68 elements in the source pool.

Symbol	[328]	[229]	[340]	[458]	[455]	[251]	[456]	[228]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Ac								13.8				13.80	13.80	13.80	13.80	0.00	13.80	0.00%
Ag	30		30.3	30	30.3	30.3	26	29.2	28.8	30.3	27.1	30.30	26.00	29.23	30.00	2.32	30.30	16.54%
Al	26		26	26	26.1	26.2		27.1	26.7	26.2	26.1	27.10	26.00	26.27	26.10	0.15	26.00	4.23%
Au	27		26	27	27	26		28.1	27.7	26	27.1	28.10	26.00	26.88	27.00	0.57	26.00	8.08%
Ba	4.9		4.86	4.9		4.86		5	4.9	4.86	6.34	6.34	4.86	5.08	4.90	0.26	4.86	30.45%
Be	132		156	132		156		146	132	156	128.4	156.00	128.40	142.30	139.00	155.46	132.00	21.50%
Bi	12		12	12	12	12.8		13.1	12.4	12.8	12.6	13.10	12.00	12.41	12.40	0.19	12.00	9.17%
Ca	7.4		7.4	7.4		7.9		7.5	7.36	7.9	7.4	7.90	7.36	7.53	7.40	0.05	7.40	7.34%
Cd	19		19	19	19.2	24		24.6	24	24	24	24.60	19.00	21.87	24.00	7.18	24.00	29.47%
Ce	13.5	13.5	14	14		13.5		12.2	12	13.5	13.53	14.00	12.00	13.30	13.50	0.51	13.50	16.67%
Co	75		82	75		82		77.9	74.8	82	82.1	82.10	74.80	78.85	79.95	12.47	82.00	9.76%
Cr	115		115.3	115	115.4	115.3		119	71.6	115.3	115	119.00	71.60	110.77	115.30	217.34	115.00	66.20%
Cs			0.65			0.65		0.66		0.65	0.664	0.66	0.65	0.65	0.65	0.00	0.65	2.15%
Cu	48		48.3	48	48.3	48.3		46	45.5	48.3	47.7	48.30	45.50	47.60	48.00	1.16	48.30	6.15%
Dy	24.7	24.7	24.7	25		24.7		25.9	25.4		25.6	25.90	24.70	25.09	24.85	0.23	24.70	4.86%
Er	28.3	28.3	28.3	28		28.3		30.2	29.6		29.6	30.20	28.00	28.83	28.30	0.70	28.30	7.86%
Eu	7.9	7.9	7.9	7.9		7.9		6			7.9	7.90	6.00	7.63	7.90	0.52	7.90	31.67%
Fe	82		82	82	81.6	81		83.1	81.2	81.6	81.9	83.10	81.00	81.82	81.90	0.36	82.00	2.59%
Fr								0.63				0.63	0.63	0.63	0.63	0.00	0.63	0.00%
Ga			6.67			6.67		6.67		6.67	37.8	37.80	6.67	12.90	6.67	193.82	6.67	466.72%

Symbol	[328]	[229]	[340]	[458]	[455]	[251]	[456]	[228]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Gd	21.8	21.8	22	22		21.8		22.7	12.5		22.17	22.70	12.50	20.85	21.90	11.46	21.80	81.60%
Hf	30		56	30		56		54	30.4	56	55.8	56.00	30.00	46.03	54.90	173.63	56.00	86.67%
Ho	26.3	26.3	26.3	26		26.3		27.2	26.7		26.31	27.20	26.00	26.43	26.30	0.13	26.30	4.62%
In			3.68			3.68		3.8		3.68	4.78	4.78	3.68	3.92	3.68	0.23	3.68	29.89%
Ir	210		209	210		209		214	210	209	209.6	214.00	209.00	210.08	209.80	2.73	209.00	2.39%
K	1.3		1.3	1.3		1.3		1.3	1.27		0.9	1.30	0.90	1.24	1.30	0.02	1.30	44.44%
La	14.3	14.3	14.9	14		14.9		15.2	14.9	14.9	13.6	15.20	13.60	14.56	14.90	0.28	14.90	11.76%
Li	4.2		4.24	4.2		4.24		4.31	4.22	4.24	3.85	4.31	3.85	4.19	4.23	0.02	4.24	11.95%
Lu	27.2	27.2	27.2	27		27.2		34.5			24.41	34.50	24.41	27.82	27.20	9.74	27.20	41.34%
Mg	17		17.3	17	17.3	17.3	16.7	17.7	34.2	17.3	17.3	34.20	16.70	18.91	17.30	28.93	17.30	104.79%
Mn			79.5			79.5		78		79.5	82.6	82.60	78.00	79.82	79.50	2.84	79.50	5.90%
Mo	126		125.6	20		125.6	147	118	19.7	125.6	125	147.00	19.70	103.61	125.60	2316.35	125.60	646.19%
Na	3.3		2.53	3.3		2.53		3.5	3.34	2.53	1.98	3.50	1.98	2.88	2.92	0.30	2.53	76.77%
Nb	38		37.5	38	37.5	37.5		38.2	58.9	37.5	37.6	58.90	37.50	40.08	37.60	49.89	37.50	57.07%
Nd	16.3	16.3	16.3	16		16.3		14.8	14.5		15.8	16.30	14.50	15.79	16.15	0.53	16.30	12.41%
Ni	76		76	76	76	76		76.5	73.6	76	85.8	85.80	73.60	76.88	76.00	11.89	76.00	16.58%
Np								40.6				40.60	40.60	40.60	40.60	0.00	40.60	0.00%
Os	222		223	222		223		214	222	223	222.5	223.00	214.00	221.44	222.25	9.25	222.00	4.21%
Pa								39.8				39.80	39.80	39.80	39.80	0.00	39.80	0.00%
Pb	5.6		5.6	5.6	5.59	5.59	5.4	5.5	5.59	5.59	8.6	8.60	5.40	5.87	5.59	0.93	5.59	59.26%
Pd	44		43.6	44		43.6		52.1	43.7	43.6	44.8	52.10	43.60	44.93	43.85	8.57	43.60	19.50%
Pm	18	18	18	18		18		17				18.00	17.00	17.83	18.00	0.17	18.00	5.88%
Pr	14.8	14.8	14.8	15		14.8		13.8	13.5		13.96	15.00	13.50	14.43	14.80	0.34	14.80	11.11%
Pt	61		60.9	61	61	60.9	64.2	62.2	52	60.9	60.6	64.20	52.00	60.47	60.95	10.01	60.90	23.46%

Symbol	[328]	[229]	[340]	[458]	[455]	[251]	[456]	[228]	[279]	[143]	[115]	Vmax	Vmin	Vmean	Vmedian	Vvar	Vmode	Dminmaxp
Pu	43		43	43		34.5		44.6	43.4	34.5		44.60	34.50	40.86	43.00	19.18	43.00	29.28%
Ra								6.1				6.10	6.10	6.10	6.10	0.00	6.10	0.00%
Rb			0.91			0.91		1.02		0.91	0.63	1.02	0.63	0.88	0.91	0.02	0.91	61.90%
Re	178		181	178		181		182	178	181	179	182.00	178.00	179.75	180.00	2.79	178.00	2.25%
Rh	150		147	150		147		150	150	147	148.9	150.00	147.00	148.74	149.45	2.21	150.00	2.04%
Ru	173		173	173		173		163	173	173	163	173.00	163.00	170.50	173.00	21.43	173.00	6.13%
Sc	29.1	29.1	29.1	29		29.1		31.9	29.4		31.8	31.90	29.00	29.81	29.10	1.60	29.10	10.00%
Sm	19.5	19.5	19.5	20		19.5		12.9	12.6		19.6	20.00	12.60	17.89	19.50	10.09	19.50	58.73%
Sn	18		18.4	18	18.4	18.4	16.7	20.8	16.7	18.4	17.9	20.80	16.70	18.17	18.20	1.29	18.40	24.55%
Sr	6.1		6.03	6.1		6.03		5.33	6.08	6.03	5.2	6.10	5.20	5.86	6.03	0.14	6.03	17.31%
Ta	69		69.2	69	69.2	69.2		70	68.7	69.2	69	70.00	68.70	69.17	69.20	0.13	69.20	1.89%
Tb	22.1	22.1	22.1	22		22.1		23.3	22.8		22.96	23.30	22.00	22.43	22.10	0.26	22.10	5.91%
Tc								145			123	145.00	123.00	134.00	134.00	242.00	123.00	17.89%
Th	31		30.8	31		30.8		28.4	31	30.8	28.7	31.00	28.40	30.31	30.80	1.20	30.80	9.15%
Ti	44		45.6	44	43.8	45.6		40.1	38	45.6	43.4	45.60	38.00	43.34	44.00	6.93	45.60	20.00%
Tl	2.8		2.71	2.8		2.71		2.8	2.75	2.71	5.4	5.40	2.71	3.09	2.78	0.88	2.71	99.26%
Tm	30.5	30.5	30.5	31		30.5		31			30.4	31.00	30.40	30.63	30.50	0.07	30.50	1.97%
U	111		111	111		73.1		75	111	73.1	84.4	111.00	73.10	93.70	97.70	354.58	111.00	51.85%
V	47		46.7	47	46.7	46.7		47.4	54	46.7	48.1	54.00	46.70	47.81	47.00	5.60	46.70	15.63%
W	161		160.5	161	160.6	160.6	135	156	149	160.6	160	161.00	135.00	156.43	160.55	70.88	160.60	19.26%
Y	25.6	25.6	25.5	26		25.5		26.3	26.2	25.5	25.4	26.30	25.40	25.73	25.60	0.12	25.50	3.54%
Yb	9.9	9.9	9.9	9.9		9.9		7.1	7.03		9.95	9.95	7.03	9.20	9.90	1.73	9.90	41.54%
Zn	43		41.9	43	43.4	41.9		37.9	37.2	41.9	39.5	43.40	37.20	41.08	41.90	5.30	41.90	16.67%
Zr	33		35	33		35		34.8		35	36	36.00	33.00	34.54	35.00	1.26	35.00	9.09%

Appendix XII

Table 12 The database to predict ΔH_V from 13 different properties.

Symbol	Hardness increment after HPT	Burgers vector	Atomic number	Binding energy per nucleon	Cohesive energy	Density at 300K	Heat of fusion
Unit	GPa	nm		keV	kJ/mol	kg/m^3	kJ/mol
Abb.	ΔH_V	b	A_n	BE/A	E_{coh}	ρ	H_{fus}
Mg	0.057	0.3197	12	8260	148	1740	8.95
Al	0.146	0.2864	13	8330	322	2700	10.8
Ti	1.628	0.2896	22	8720	470	4540	15.5
V	1.726	0.2618	23	8740	511	6110	16
Cr	3.696	0.2498	24	8780	396	7190	16.9
Fe	2.412	0.2482	26	8790	416	7870	13.8
Co	2.501	0.2506	27	8770	426	8900	16.2
Ni	2.383	0.2492	28	8730	429	8900	17.5
Cu	0.929	0.2556	29	8750	339	8960	13.1
Zn	0.009	0.2665	30	8740	130	7130	7.32
Zr	1.629	0.3179	40	8710	610	6510	16.9
Nb	2.000	0.2864	41	8660	730	8570	26.4
Pd	1.666	0.2751	46	8580	377	12000	17.6
Ag	0.690	0.2889	47	8550	286	10500	11.3
Hf	1.721	0.3127	72	8020	610	13300	24.1
Ta	3.259	0.2856	73	8020	782	16700	31.6
Pt	1.976	0.2775	78	7930	565	21500	19.6
Au	0.588	0.2884	79	7920	367	19300	12.6

Symbol	Lattice parameter, a	Melting temperature	Molar volume	Shear modulus at 300K	Specific heat capacity	Thermal expansion coefficient at 300K	Work function
Unit	nm	K	m^3/kmol	GPa	J/kg°C	µstrain/°C	eV
Abb.	a	T_m	V_m	G	C_p	α_L	W
Mg	0.321	922.15	0.014	17.7	1020	26.1	3.7
Al	0.405	933.15	0.01	27.1	900	23	4.28
Ti	0.295	1933.15	0.0106	40.1	520	8.35	4.33
V	0.303	2163.15	0.00834	47.4	490	8.3	4.3
Cr	0.288	2133.15	0.00723	119	450	6.2	4.5
Fe	0.287	1803.15	0.00709	83.1	440	12.3	4.5
Co	0.251	1763.15	0.00662	77.9	420	13.4	5
Ni	0.352	1723.15	0.00659	76.5	440	13.3	5.15

Symbol	Lattice parameter, a	Melting temperature	Molar volume	Shear modulus at 300K	Specific heat capacity	Thermal expansion coefficient at 300K	Work function
Unit	**nm**	**K**	**m^3/kmol**	**GPa**	**J/kg°C**	**µstrain/°C**	**eV**
Abb.	***a***	***T_m***	***V_m***	***G***	***C_p***	***α_L***	***W***
Cu	0.361	1353.15	0.00709	46	380	16.5	4.65
Zn	0.266	693.15	0.00917	37.9	390	25	4.33
Zr	0.323	2123.15	0.014	34.8	270	5.78	4.05
Nb	0.33	2743.15	0.0108	38.2	260	7.07	4.3
Pd	0.389	1823.15	0.00885	52.1	240	11.2	5.12
Ag	0.409	1235.15	0.0103	29.2	235	19.2	4.26
Hf	0.319	2503.15	0.0134	54	140	5.9	3.9
Ta	0.33	3273.15	0.0109	70	140	6.6	4.25
Pt	0.392	2043.15	0.0091	62.2	130	9	5.65
Au	0.408	1333.15	0.0102	28.1	128	14.2	5.1

Appendix XIII

Discussion on the shear modulus data of 17 elements used in Chapter 5

Table 13 Comparison of shear modulus data (GPa) of 17 elements from different sources (indicated data used in the ANN analysis).*

Symbol	Database				Handbook				Journal
	*[228]	[328]	[340]	[458]	[455]	[251]	[456]	[279]	[185]
Ag	29.2	30	30.3	30	30.3	30.3	26	28.8	27
Al	27.1	26	26	26	26.1	26.2		26.7	26.2
Au	28.1	27	26	27	27	26		27.7	27.7
Co	77.9	75	82	75		82		74.8	
Cr	119	115	115.3	115	115.4	115.3		71.6	115
Cu	46	48	48.3	48	48.3	48.3		45.5	48.3
Fe	83.1	82	82	82	81.6(soft); 60(cast)	60(cast iron); 81 (steel)		81.2	81.6
Mg	17.7	17	17.3	17	17.3	17.3	16.7	34.2	17.3
Nb	38.2	38	37.5	38	37.5	37.5		58.9	37.5
Ni	76.5	76	76	76	76 (soft); 83.9 (hard)	76		73.6	75
Pd	52.1	44	43.6	44		43.6		43.7	43.6
Pt	62.2	61	60.9	61	61	60.9	64.2	52	
Ta	70	69	69.2	69	69.2	69.2		68.7	69
Ti	40.1	44	45.6	44	43.8	45.6		38	45.6
V	47.4	47	46.7	47	46.7	46.7		54	46.7

Symbol	Database				Handbook				Journal
	*[228]	[328]	[340]	[458]	[455]	[251]	[456]	[279]	[185]
Zn	37.9	43	41.9	43	43.4	41.9		37.2	41.9
Zr	34.8	33	35	33		35		36 (From magnesium thermal Reduction); 32.7(Iodide, bar)	35

Table 14 The elastic constants C_{ij} (GPa) in GPa of Cr, Nb, V, Pd, Pt, Co, Mg and Ti with Voigt-Reuss-Hill modulus.

Symbol	Structure	C_{11}	C_{12}	C_{44}	C_{13}	C_{33}	C_{66}	G_V	G_R	G_H	G in CES	Ref for C_{ij}
Cr	bcc	355.37	75.37	100.00				116.00	112.90	114.45	119	[459]
Nb	bcc	245.60	138.70	29.30				38.96	35.76	37.36	38.2	[460]
V	bcc	230.06	119.71	42.81				47.76	47.03	47.39	47.4	[298]
Pd	bcc	223.78	173.12	71.25				52.88	41.30	47.09	52.1	[461]
Pt	fcc	346.70	250.70	76.50				65.10	61.82	63.46	62.2	[462]
Co	hcp	307.10	165.00	75.50	102.70	358.10	71.05	84.54	80.18	82.36	77.9	[285]
Mg	hcp	59.18	25.68	16.34	21.52	61.47	16.75	17.29	17.18	17.23	17.7	[463]
Ti	hcp	162.40	92.00	46.70	69.00	180.70	35.20	44.09	42.62	43.36	40.1	[297]

The shear moduli of the 17 elements gathered from various sources are shown in Table 13. Although most data are reasonably consistent, for Cr, Mg, Nb, V, Pd, Pt and Ti, their variations are larger than 15%. Most discrepancies can be found in the Handbook of the physicochemical properties of the elements [279]. CES data agree well with other sources except for Pd, which is about 20% larger.

For a cubic single crystal [464]:

$$G_V = \frac{C_{11} - C_{12} + 3C_{44}}{5}$$

$$G_R = 5\left[\frac{4}{C_{11} - C_{12}} + \frac{3}{C_{44}}\right]^{-1}$$

$$G_H = \frac{G_V + G_R}{2}$$

For hexagonal symmetry [465]:

$$G_V = \frac{1}{15}\left(2C_{11} + C_{33} - 2C_{13} - C_{12} + 6C_{44} + 3C_{66}\right)$$

$$G_R = 15\left[4A\left(C_{11} + C_{12}\right) + 8AC_{13} + \frac{6}{C_{44}} + \frac{6}{C_{66}} + 2AC_{33}\right]^{-1}$$

$$G_H = \frac{G_V + G_R}{2}$$

with

$$A = \left[C_{33}\left(C_{11} + C_{12}\right) - 2C_{13}^2\right]^{-1}$$

$$C_{66} = \frac{C_{11} - C_{12}}{2}$$

Most shear modulus measurements were carried out on anisotropic single crystals. The shear modulus of isotropic polycrystalline materials can be derived from elastic constants of the single crystal, but the relation between elastic constants and shear modulus is not unique. One of the widely used averaging methods is the Voigt-Reuss-Hill approximation (VRH) [157], which is obtained by calculate the arithmetic mean of the Voigt average (***G_V***) and Reuss average (***G_R***) of the elasticity based on the

assumption of a uniform strain field and uniform stress field respectively [158]. In literatures, the VRH average (G_H) for cubic and hexagonal polycrystals can be determined from the single crystal elastic tensor C_{ij} by explicit formulas as below, where the subscript letters H, V and R denote the Hill, Voigt and Reuss average values:

Table 14 lists the literature values of single crystal elastic constants of 8 elements (Cr, Nb, V, Pd, Pt, Co, Mg and Ti), which are considered to have inconsistent shear modulus out of the 17 metals. The elastic property of a cubic crystal is characterized by three independent elastic constants C_{11}, C_{12} and C_{44}. While for hexagonal phase, there are two more independent elastic constants, i.e. C_{13} and C_{33}. The VRH average (G_H) of the 8 elements calculated from the corresponding single crystal elastic constants, are compared with data recorded in CES. The elastic moduli determined from VRH method are always in the general agreement with the ones collected from CES, which are used in our ANNs and physical models. The difference between G_V and G_R is about 5% for Pt, and even less for other elements, except for Pd, where G_V and G_R differs by more than 20% (10 GPa). Because $G_V - G_R$ approximation provides the lowest upper bounds and highest lower bounds, and G_H is an empirical estimation, the true shear modulus of isotropic crystals may lie anywhere between G_V and G_R [160]. Thus, the large discrepancy in the VRH values for Pd is understandable. In this case, the data recorded in CES for Pd is closer to the calculated G_V. The impact of changing this value is within the overall experimental error.

Appendix XIV

Table 15 Property data of 175 compounds used in ANNs and SR.

Name	Molecular Form	M_w	T_b / K	T_c / K	P_c / Pa	D_m /D	ΔH_{vb} / J mol^{-1}
Hydrogen sulfide	H2S	34.081	213.6	373.1	9000000	0.97833	18670
Sulfur dioxide	O2S	64.064	263.1	430.64	7884000	1.63305	24940
Ethylbenzene	C8H10	106.165	409.34	617.15	3609000	0.59	35570
Styrene	C8H8	104.15	418	635.2	3870000	0.123	38700
Benzonitrile	C7H5N	103.122	464.3	700	4200000	4.18	45900
Benzyl alcohol	C7H8O	108.138	478.46	715	4300000	1.71	50480
Benzaldehyde	C7H6O	106.122	452	695	4700000	3	42500
Anisole	C7H8O	108.138	426.9	646.5	4240000	1.38	38970
Nitrous oxide	N2O	44.012	184.67	309.52	7245000	0.16083	16530
Hydrogen iodide	HI	127.912	237.6	424	8310000	0.448	19760
Tetrafluorohydrazine	F4N2	104.007	199	309	3750000	0.257	13270
Nitric oxide	NO	30.006	121.41	180	6480000	0.15872	13830
2-Ethylhexyl acetate	C10H20O2	172.265	472	642	2090000	1.8	43500
2-Ethyl-1-hexanol	C8H18O	130.228	457.8	640.6	2800000	1.74	54200
Ethyl propanoate	C5H10O2	102.132	372.3	546.7	3450000	1.74	33880
Ethyl butanoate	C6H12O2	116.158	394.5	568.8	3100000	1.74	35470
1,1-Diethoxyethane	C6H14O2	118.174	375.4	540	3220000	1.38	36280
p-Cresol	C7H8O	108.138	475.13	704.6	4070000	1.48	47450
4-Methylaniline	C7H9N	107.153	473.6	667	2400000	1.52	44300
1,2-Dibromoethane	C2H4Br2	187.861	404.8	583	7200000	1.19	34770
1-Butene	C4H8	56.107	266.89	419.29	4005000	0.359	22070
1-Butyne	C4H6	54.091	281.23	440	4600000	0.782	24520
1-Propanethiol	C3H8S	76.161	341	537	4600000	1.6	29540
1,2-Dichloroethane	C2H4Cl2	98.959	356.7	561	5400000	1.83	31980
Propylamine	C3H9N	59.11	320.37	499.2	4740000	1.17	29550
Propanenitrile	C3H5N	55.079	370.29	561.3	4260000	4.05	31810
Acrylonitrile	C3H3N	53.063	350.5	540	4660000	3.92	32600
1,2-Ethanediamine	C2H8N2	60.098	390	613.1	6710000	1.99	37980
Allyl alcohol	C3H6O	58.079	370.5	547.1	5640000	1.6	40000
1,2-Ethanediol	C2H6O2	62.068	470.5	720	8000000	2.36	50500
Methyl formate	C2H4O2	60.052	304.9	487.2	6000000	1.77	27920
2-Pentanone	C5H10O	86.132	375.41	561.1	3683000	2.7	33440
Vinyl acetate	C4H6O2	86.09	346	519.2	4185000	1.79	34600
Diisopropylamine	C6H15N	101.19	357.1	523.1	3020000	1.15	30400
Diisopropyl ether	C6H14O	102.174	341.6	500.3	2832000	1.13	29100

Acetic anhydride	C4H6O3	102.089	412.7	606	4000000	2.8	38200
m-Cresol	C7H8O	108.138	475.42	705.8	4360000	1.48	47400
3-Methylaniline	C7H9N	107.153	476.5	709	4200000	1.45	44900
2,4-Dimethylpyridine	C7H9N	107.153	431.53	647	3950000	2.3	38530
2,6-Dimethylpyridine	C7H9N	107.153	417.16	624	3850000	1.66	37460
Toluene	C7H8	92.139	383.78	591.8	4110000	0.375	33180
4-Methylpyridine	C6H7N	93.127	418.51	646	4670000	2.7	37510
Chlorobenzene	C6H5Cl	112.557	404.87	633.4	4520000	1.69	35190
Phenol	C6H6O	94.111	455.02	694.2	5930000	1.224	45690
3-Methylpyridine	C6H7N	93.127	417.29	644.6	4650000	2.4	37350
2-Methylpyridine	C6H7N	93.127	402.53	621	4620000	1.85	36170
Pentanoic acid	C5H10O2	102.132	459.3	637.2	3630000	1.61	44100
Propyl acetate	C5H10O2	102.132	374.69	549.7	3360000	1.78	33920
1-Pentene	C5H10	70.133	303.11	464.8	3560000	0.5	25200
Butylamine	C4H11N	73.137	350.15	531.9	4200000	1	31810
Butanenitrile	C4H7N	69.106	390.8	585.4	3880000	3.73	33680
1-Butanethiol	C4H10S	90.187	371.7	570	4000000	1.53	32230
2-Methoxyethanol	C3H8O2	76.095	397.3	597.6	5285000	2.36	37540
Diethylamine	C4H11N	73.137	328.7	499.7	3754000	0.92	29060
Ethyl vinyl ether	C4H8O	72.106	308.7	475	4070000	1.26	26200
Ethyl formate	C3H6O2	74.079	327.6	508.54	4740000	1.93	29910
Pyrrole	C4H5N	67.09	402.94	640	5700000	1.767	38750
Tetrahydrofuran	C4H8O	72.106	338	540.5	5190000	1.75	29810
Furan	C4H4O	68.074	304.7	490.2	5300000	0.66	27100
Tetrahydrothiophene	C4H8S	88.172	394.3	632	5400000	1.9	34660
Thiophene	C4H4S	84.14	357.2	580	5700000	0.55	31480
Isobutyl acetate	C6H12O2	116.158	389.7	561	2990000	1.86	35900
Pentanenitrile	C5H9N	83.132	414.5	610.3	3580000	4.12	36090
Propyl formate	C4H8O2	88.106	354.1	538	4060000	1.89	33610
Pyridine	C5H5N	79.101	388.38	620	5650000	2.215	35090
2-Ethoxyethyl acetate	C6H12O3	132.157	429.6	608	3170000	2.25	40760
Butyl vinyl ether	C6H12O	100.158	367	540	3200000	1.25	31580
Dipropyl ether	C6H14O	102.174	363.23	530.6	3028000	1.21	31310
Diethylene glycol	C4H10O3	106.12	519	750	4700000	2.31	52300
Diethylene glycol monomethyl ether	C5H12O3	120.147	466	672	3670000	1.6	46600
Diethylene glycol monoethyl ether	C6H14O3	134.173	469	670	3167000	1.6	47500
Dibutylamine	C8H19N	129.244	432.8	607.5	3110000	0.98	38440
Propene	C3H6	42.08	225.46	364.9	4600000	0.366	18420
Dimethyl ether	C2H6O	46.068	248.4	400.378	5356000	1.3	21510
Cyclopentanone	C5H8O	84.117	403.72	624	4600000	3.3	36350

Triethylamine	C6H15N	101.19	362	535.6	3030000	0.66	31010
Propanal	C3H6O	58.079	321	505	5260000	2.72	28310
Butanal	C4H8O	72.106	348	537	4320000	2.72	31500
Pyrrolidine	C4H9N	71.121	359.71	569	5700000	1.57	33010
Butyl acetate	C6H12O2	116.158	399.3	575.6	3140000	1.87	36280
Isopentyl acetate	C7H14O2	130.185	415.7	586.1	2760000	1.86	37500
Octanoic acid	C8H16O2	144.212	512	693	2870000	1.15	58500
Dimethylamine	C2H7N	45.084	280.03	437.2	5340000	1.01	26400
Chlorotrifluorosilane	ClF3Si	120.534	203.2	307.7	3460000	0.636	18700
Ethyl acetate	C4H8O2	88.106	350.26	523.3	3870000	1.78	31940
Mesityl oxide	C6H10O	98.142	403	605	4000000	2.79	36100
Tetrahydropyran	C5H10O	86.132	361	572	4770000	1.58	31170
Dipropylamine	C6H15N	101.19	382.5	555.8	3630000	1.03	33470
Dibutyl ether	C8H18O	130.228	413.43	584	3000000	1.17	36490
Methyloxirane	C3H6O	58.079	308	485	5200000	2.01	27350
Phosphorothioc trifluoride	F3PS	120.034	220.9	346	3820000	0.64	19600
Hydrazine	H4N2	32.045	386.7	653	14700000	1.75	41800
sec-Butylamine	C4H11N	73.137	335.88	514	5000000	1.28	29920
Diethyl sulfide	C4H10S	90.187	365.3	557.8	3900000	1.645	31770
2-Octanol	C8H18O	130.228	452.5	629.6	2754000	1.71	44400
1,1,1-Trifluoroethane	C2H3F3	84.04	225.9	345.86	3764000	2.347	18990
Fluorobenzene	C6H5F	96.102	357.88	560.09	4551000	1.6	31190
1,3-Propanediol	C3H8O2	76.095	487.6	718.2	6550000	2.55	57900
1-Chloropropane	C3H7Cl	78.541	319.7	503	4580000	2.05	27180
Isobutyl formate	C5H10O2	102.132	371.4	551	3880000	1.88	33600
Methyl propyl ether	C4H10O	74.121	312.3	476.2	3801000	1.107	26750
Glycerol	C3H8O3	92.094	563	850	7500000	2.56	61000
1,2-Propanediol	C3H8O2	76.095	460.8	676.4	5941000	2.25	52400
cis-2-Butene	C4H8	56.107	276.86	435.75	4226000	0.253	23340
Iodobenzene	C6H5I	204.008	461.6	721	4520000	1.7	39500
2-Hexanone	C6H12O	100.158	400.8	587.1	3300000	2.66	36350
Isopropyl methyl ether	C4H10O	74.121	303.92	464.4	3762000	1.247	26050
Diethyl ether	C4H10O	74.121	307.7	466.7	3644000	1.098	26520
2-Pentanol	C5H12O	88.148	392.5	560.3	3675000	1.66	41400
Aniline	C6H7N	93.127	457.32	705	5630000	1.13	42440
Ethyl methyl sulfide	C3H8S	76.161	339.9	533	4250000	1.56	29530
Dimethyl disulfide	C2H6S2	94.199	382.89	607.8	5070000	1.85	33780
Pentyl acetate	C7H14O2	130.185	422.4	599	2730000	1.75	38420
Carbon monoxide	CO	28.01	81.7	132.86	3494000	0.1098	6040
Ethanol	C2H6O	46.068	351.44	514	6137000	1.69	38560
Acetic acid	C2H4O2	60.052	391.1	590.7	5780000	1.7	23700

Methanol	CH4O	32.042	337.8	512.5	8084000	1.7	35210
2-Propanol	C3H8O	60.095	355.5	508.3	4764000	1.58	39850
Acetone	C3H6O	58.079	329.2	508.1	4700000	2.88	29100
Trichloromethane	CHCl3	119.378	334.32	536.4	5470000	1.04	29240
1-Propanol	C3H8O	60.095	370.4	536.8	5169000	1.55	41440
1-Butanol	C4H10O	74.121	390.88	563	4414000	1.66	43290
1-Pentanol	C5H12O	88.148	411.13	588.1	3897000	1.7	44360
1,1,1-Trichloroethane	C2H3Cl3	133.404	347.24	545	4300000	1.755	29860
Chloromethane	CH3Cl	50.488	249.06	416.25	6679000	1.8963	21400
Methylamine	CH5N	31.058	266.83	430.8	7620000	1.31	25600
Bromoethane	C2H5Br	108.965	311.7	503.9	6230000	2.04	27040
Propane	C3H8	44.096	231.1	369.83	4248000	0.084	19040
Chloroethane	C2H5Cl	64.514	285.5	460.4	5300000	2.05	24650
Chloroethene	C2H3Cl	62.498	259.4	432	5670000	1.45	20800
Acetonitrile	C2H3N	41.052	354.8	545.5	4850000	3.92519	29750
Ethanethiol	C2H6S	62.134	308.2	499	5490000	1.58	26790
Dichloromethane	CH2Cl2	84.933	313	510	6100000	1.6	28060
Dimethyl sulfide	C2H6S	62.134	310.48	503	5530000	1.554	27000
Oxirane	C2H4O	44.052	283.8	469	7200000	1.89	25540
Isobutane	C4H10	58.122	261.42	407.885	3639000	0.132	21300
Isopropylamine	C3H9N	59.11	304.91	472.2	4550000	1.19	27830
1,1-Dichloroethane	C2H4Cl2	98.959	330.5	523	5070000	2.06	28850
1,1-Difluoroethane	C2H4F2	66.05	249.1	386.7	4500000	2.27	21560
Dichlorofluoromethane	CHCl2F	102.923	282.1	451.58	5180000	1.29	25200
Chlorodifluoromethane	CHClF2	86.469	232.5	369.5	5035000	1.42	20200
Trimethylamine	C3H9N	59.11	276.02	433	4080000	0.612	22940
Nitromethane	CH3NO2	61.041	374.34	588	5870000	3.46	33990
tert-Butylamine	C4H11N	73.137	317.19	483.7	3850000	1.29	28270
2-Methyl-2-propanol	C4H10O	74.121	355.6	506.2	3972000	1.66	39070
Trichlorofluoromethane	CCl3F	137.368	296.9	471.1	4470000	0.46	25100
Dichlorodifluoromethane	CCl2F2	120.914	243.4	384.95	4136000	0.51	20100
Chlorotrifluoromethane	CClF3	104.459	191.8	302	3870000	0.5	15800
2-Methyl-2-butanol	C5H12O	88.148	375.6	543.7	3710000	1.82	39040
Trifluoroacetic acid	C2HF3O2	114.023	346	491.3	3258000	2.28	33300
1,2-Dichloro-1,1,2,2-tetrafluoroethane	C2Cl2F4	170.921	276.7	418.78	3252000	0.5	23300
Chloropentafluoroethane	C2ClF5	154.466	234.1	353.2	3229000	0.52	19410
Perchloryl fluoride	ClFO3	102.449	226.4	368.4	5370000	0.023	19330
Hydrogen chloride	ClH	36.461	188	324.7	8310000	1.1086	16150
Ammonia	H3N	17.031	239.82	405.56	11357000	1.4718	23330
Water	H2O	18.015	373.2	647.14	22060000	1.8546	40650

Nitrogen trifluoride	F3N	71.002	144.4	234	4460000	0.235	11560
Phosphorus(III) fluoride	F3P	87.969	171.4	271.2	4330000	1.03	16500
Isopentane	C5H12	72.149	301.03	460.4	3380000	0.13	24690
Isobutylamine	C4H11N	73.137	340.9	519	4070000	1.27	30610
2-Methyl-1-propanol	C4H10O	74.121	381.04	547.8	4295000	1.64	41820
2-Butanone	C4H8O	72.106	352.74	536.7	4207000	2.779	31300
Phosphine	H3P	33.998	185.4	324.5	6540000	0.574	14600
1,1,2-Trichloroethane	C2H3Cl3	133.404	387	602	4480000	1.4	34820
Trichloroethene	C2HCl3	131.388	360.36	544.2	5020000	0.8	31400
Methyl acetate	C3H6O2	74.079	330.02	506.5	4750000	1.72	30320
Dibutyl phthalate	C16H22O4	278.344	613	797	1660000	2.82	79200
o-Xylene	C8H10	106.165	417.7	630.3	3732000	0.64	36240
o-Cresol	C7H8O	108.138	464.19	697.6	4170000	1.45	45190
2-Methylaniline	C7H9N	107.153	473.5	717	4700000	1.6	44600
Diethyl oxalate	C6H10O4	146.141	458.9	645.8	3060000	2.49	42000
3-Pentanone	C5H10O	86.132	374.9	561.4	3729000	2.82	33450
g-Butyrolactone	C4H6O2	86.09	477	732.5	5100000	4.27	52200
Furfural	C5H4O2	96.085	434.9	670	5890000	3.54	43200
Acetophenone	C8H8O	120.149	475	709.6	4010000	3.02	43980

Appendix XV

Table 16 The contribution of each input variable in ANN models, data are generated through the 'CW' method and the 'PaD' method.

The 'CW' method	M_w	T_b	T_c	P_c	D_m
Experiment 1	15.73%	31.10%	24.02%	22.84%	6.32%
Experiment 2	12.67%	29.42%	15.52%	25.90%	16.49%
Experiment 3	17.81%	32.89%	29.46%	6.35%	13.48%
Mean	15.41%	31.13%	23.00%	18.36%	12.10%

The 'PaD' method	M_w	T_b	T_c	P_c	D_m
Experiment 1	4.79%	52.08%	40.28%	0.65%	2.21%
Experiment 2	2.11%	58.10%	17.96%	21.77%	0.06%
Experiment 3	10.17%	74.38%	15.03%	0.41%	0.00%
Mean	5.69%	61.52%	24.43%	7.61%	0.76%

Appendix XVI

Table 17 The contribution of each input variable in SR models, data are generated through Discipulus.

Frequency	M_w	T_b	T_c	P_c	D_m
Experiment 1	97%	100%	93%	63%	77%
Experiment 2	100%	100%	90%	47%	90%
Experiment 3	100%	100%	90%	30%	97%
Mean	99%	100%	91%	47%	88%

Average impact	M_w	T_b	T_c	P_c	D_m
Experiment 1	2.20%	60.23%	15.34%	0.83%	0.30%
Experiment 2	1.00%	70.93%	15.97%	2.05%	4.87%
Experiment 3	4.67%	63.16%	7.36%	2.17%	0.51%
Mean	2.62%	64.77%	12.89%	1.68%	1.90%

Maximum impact	M_w	T_b	T_c	P_c	D_m
Experiment 1	8.97%	94.44%	62.22%	1.47%	1.30%
Experiment 2	2.83%	92.04%	34.21%	4.01%	89.63%
Experiment 3	61.21%	93.54%	26.93%	3.23%	5.22%
Mean	24.34%	93.34%	41.12%	2.91%	32.05%

Appendix XVII

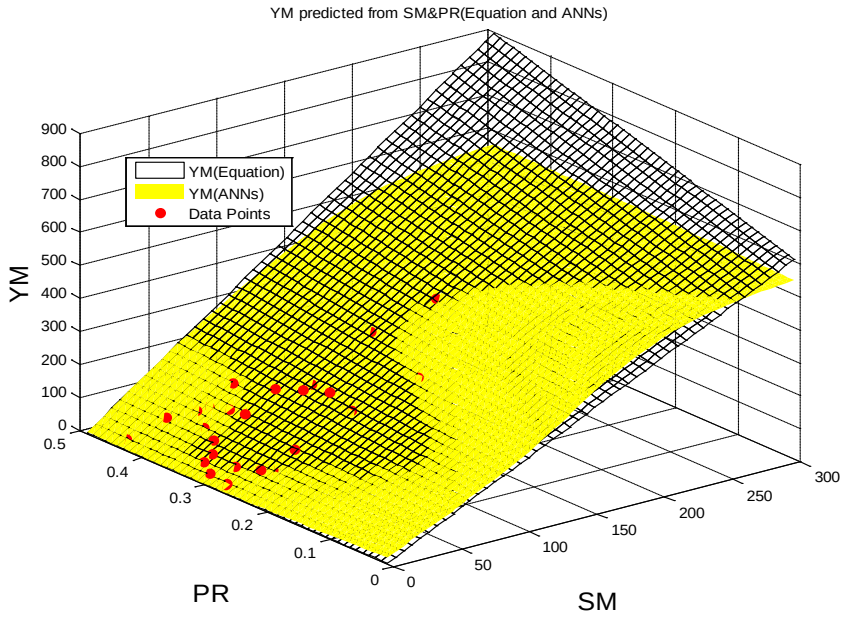
Table 18 The 'best' experimental elastic property data for 36 elements. Underlined data indicate the property have inconsistent values in the source pool.

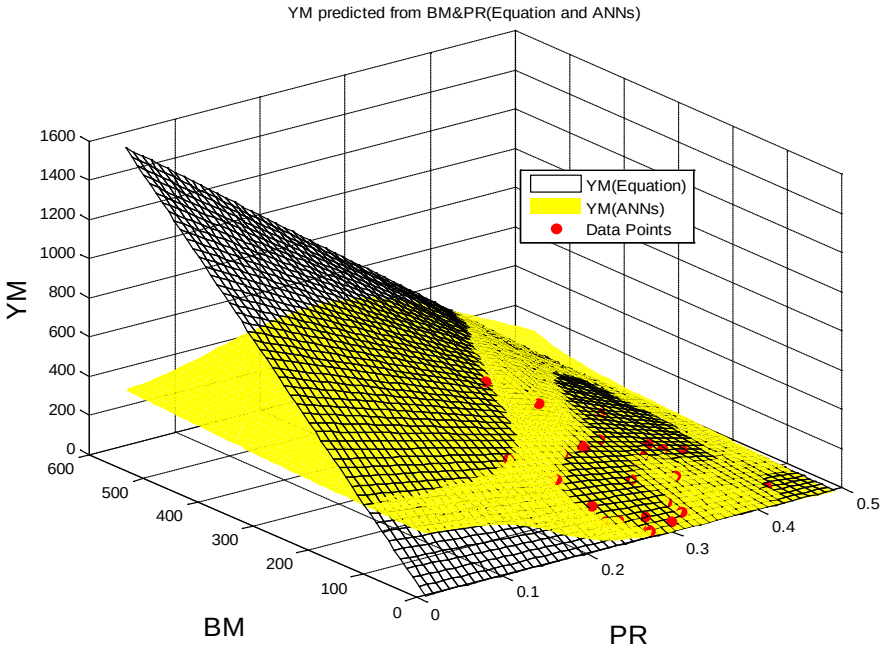
Symbol	Most common value in the source pool				Best experimental value				Ref.
	v	*K*	*E*	*G*	*v*	*K*	*E*	*G*	
		GPa				GPa			
Gd	0.259	38	55	*21.9*	*0.259*	*37.9*	*54.8*	*21.8*	[328]
Nd	0.281	31.9	41.2	*16.15*	*0.281*	*31.8*	*41.8*	*16.3*	[337]
Er	0.237	*44.4*	69.95	28.3	*0.237*	*44.3*	*69.9*	*28.3*	[328]
Ir	0.26	*365*	528	209.8	*0.246*	*354.67*	*540.19*	*216.74*	[140]
Li	0.36	11.2	*4.91*	*4.23*	*0.351*	*11.97*	*10.7*	*3.96*	[348]
Pr	0.281	28.8	*37.3*	*14.8*	*0.281*	*28.8*	*37.9*	*14.8*	[350]
Tc	0.301	289	*351*	*134*	0.309	281	322	123	[352]
Rh	0.26	*275.5*	*330*	149.45	*0.264*	*266.53*	*377.84*	*149.49*	[355]
Cs	*0.295*	*1.6*	1.7	0.65	*0.356*	*2.03*	*1.76*	*0.65*	[161]
Os	*0.25*	*376.45*	558.75	222.25	*0.231*	*405.3*	*654*	*265.6*	[362]
Re	*0.293*	*367.5*	463	180	0.291	366.9	460.8	178.5	[370]
Ru	*0.286*	*286*	432	173	0.245	310.7	475.2	190.8	[370]
Sc	*0.279*	*56.6*	74.4	29.1	*0.28*	*57.1*	*75.5*	*29.5*	[374]
Tm	*0.213*	*44.5*	74	30.5	*0.213*	*44.5*	*74*	*30.5*	[377]
Y	*0.252*	*41.2*	64.4	25.6	*0.246*	*41.5*	*63.3*	*25.4*	[379]
Th	*0.27*	54	*78.3*	30.8	*0.3*	*59.59*	*71.94*	*27.7*	[385]
Ho	*0.231*	*40.2*	*64.8*	26.3	*0.231*	*40.2*	*64.8*	*26.3*	[320]

Symbol	Most common value in the source pool				Best experimental value				Ref.
	v	*K*	*E*	*G*	*v*	*K*	*E*	*G*	
		GPa				GPa			
Zr	*0.34*	*89.8*	*94.95*	35	*0.333*	*95.7*	*96.1*	*36.1*	[391]
Hf	*0.3*	109	*140*	*54.9*	*0.281*	*108.6*	*143*	*55.8*	[329]
Na	*0.34*	6.52	*6.8*	*2.915*	*0.365*	*6.88*	*5.58*	*2.05*	[394]
Pu	*0.195*	54	*96*	*43*	*0.186*	*54.72*	*103.09*	*43.46*	[397]
Ce	0.245	*21.7*	*33.6*	*13.5*	*0.244*	*18.18*	*27.94*	*11.23*	[404]
In	0.45	*41.1*	*10.8*	*3.68*	*0.447*	*42.4*	*13.6*	*4.7*	[406]
K	0.35	*3.14*	*3.53*	*1.3*	*0.38*	*3.33*	*2.49*	*0.91*	[412]
Tl	0.45	*35.8*	*7.975*	*2.775*	*0.429*	*35.6*	*15.3*	*5.3*	[414]
Be	*0.032*	*117.785*	*291.8*	*139*	*0.05*	*116.8*	*315.2*	*150.1*	[419]
Cd	*0.3*	*46.7*	*58.6*	*24*	*0.303*	*53*	*62.6*	*24*	[424]
Eu	*0.152*	*8.3*	*18.2*	*7.9*	*0.155*	*8.85*	*18.2*	*7.9*	[301, 389]
Ga	*0.47*	*42.6*	*9.81*	*6.67*	*0.23*	*48.1*	*90.99*	*38.4*	[393, 395]
La	*0.28*	*27.9*	*37.5*	*14.9*	*0.28*	*27.9*	*36.6*	*14.3*	[119, 303]
Lu	*0.2605*	*47.6*	*68.6*	*27.2*	*0.261*	*47.6*	*68.4*	*27.2*	[433]
Rb	*0.328*	*2.5*	*2.35*	*0.91*	*0.372*	*2.32*	*1.78*	*0.65*	[435]
Sm	*0.274*	*37.8*	*49.7*	*19.5*	*0.284*	*37.88*	*48*	*18.7*	[301, 306]
U	*0.22*	*100*	*208*	*97.7*	*0.359*	*255.8*	*217.09*	*79.9*	[402, 406-408]
Yb	*0.207*	*30.5*	*23.9*	*9.9*	*0.207*	*13.61*	*23.8*	*9.9*	[301, 344]

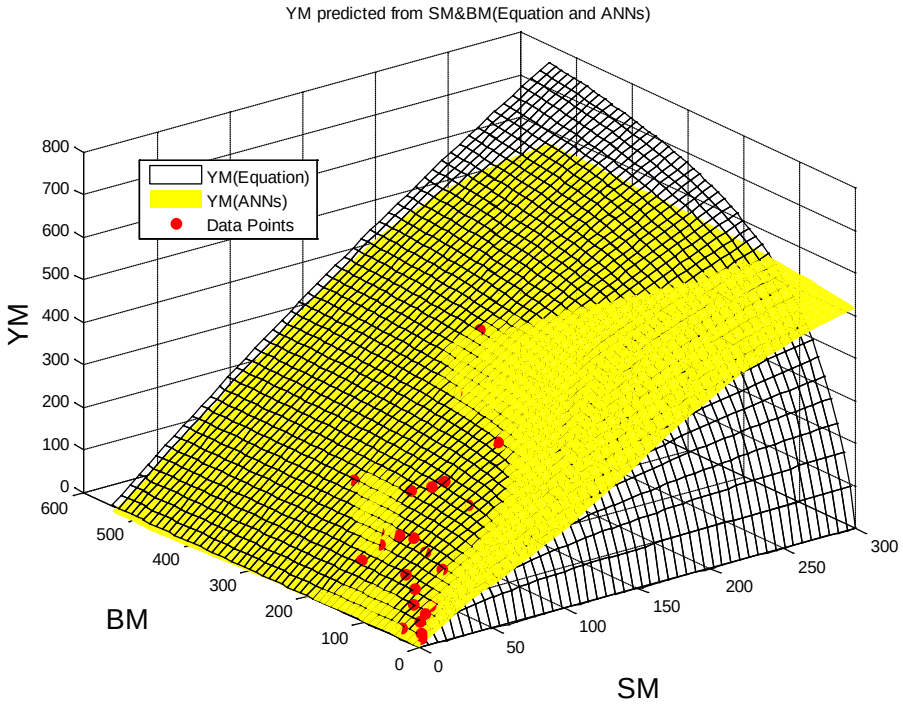
Appendix XVIII

The ANN modelling of elastic property correlations in 3D dimensions:

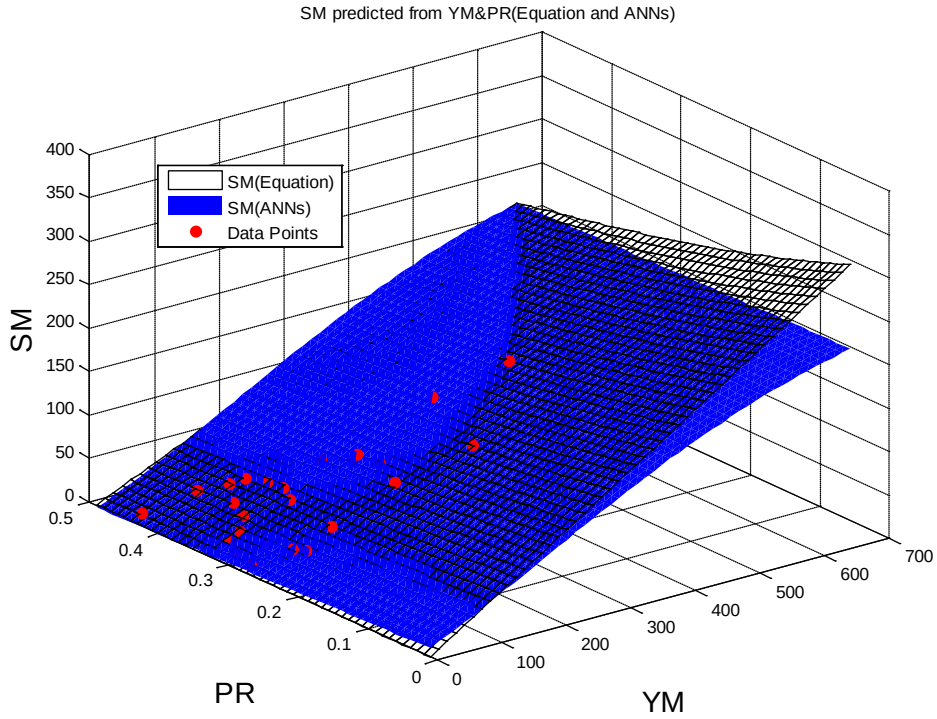




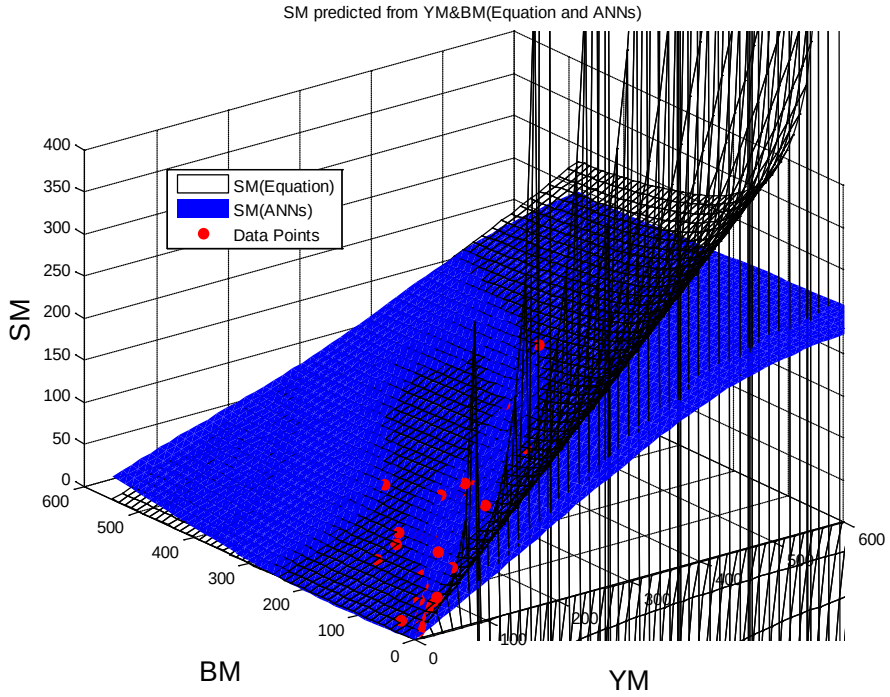
(b)


YM predicted from SM&BM(Equation and ANNs)
YM(Equation)
YM(ANNs)
Data Points
YM
BM
SM
800
700
600
500
400
300
200
100
0
600
500
400
300
200
100
0
0
50
100
150
200
250
300

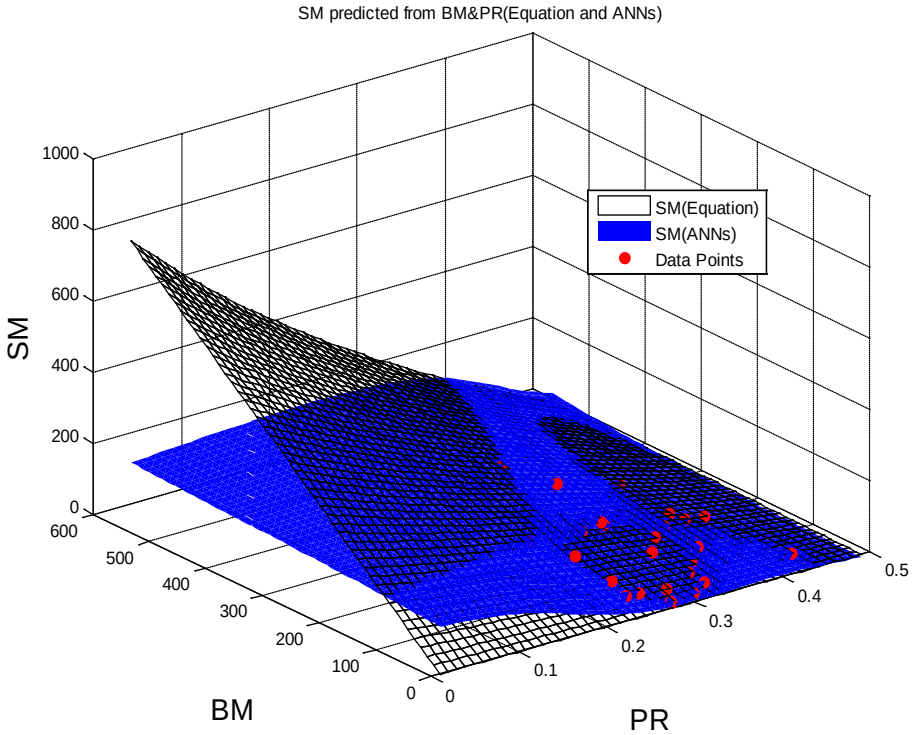
(c)


SM predicted from YM&PR(Equation and ANNs)
SM(Equation)
SM(ANNs)
Data Points
SM
PR
YM
400
350
300
250
200
150
100
50
0
0.5
0.4
0.3
0.2
0.1
0
0
100
200
300
400
500
600
700

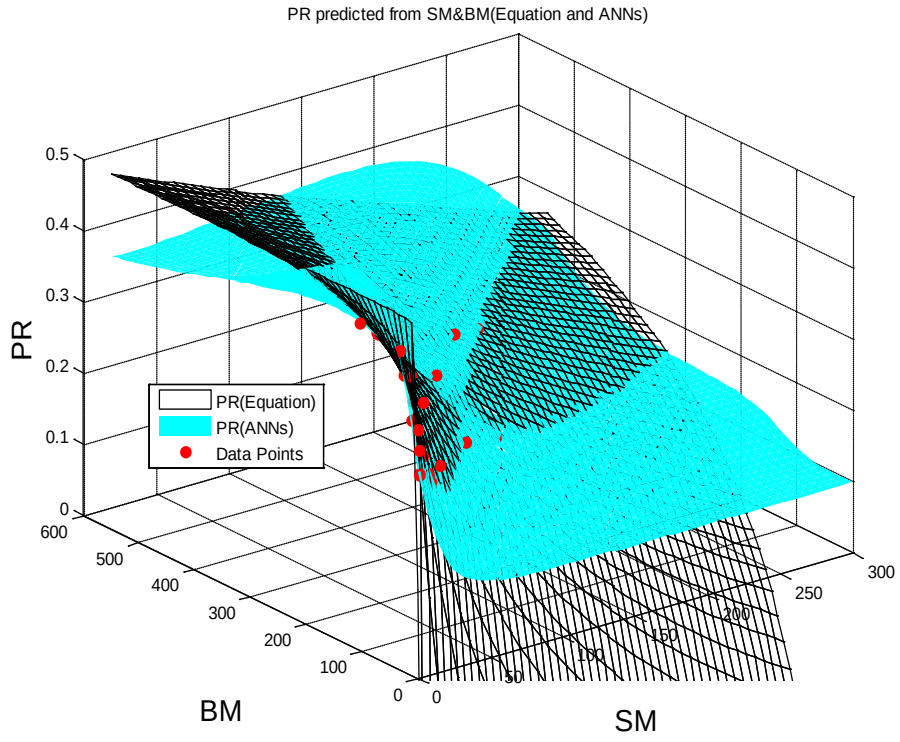
(d)


SM predicted from YM&BM(Equation and ANNs)
SM(Equation)
SM(ANNs)
Data Points
SM
BM
YM

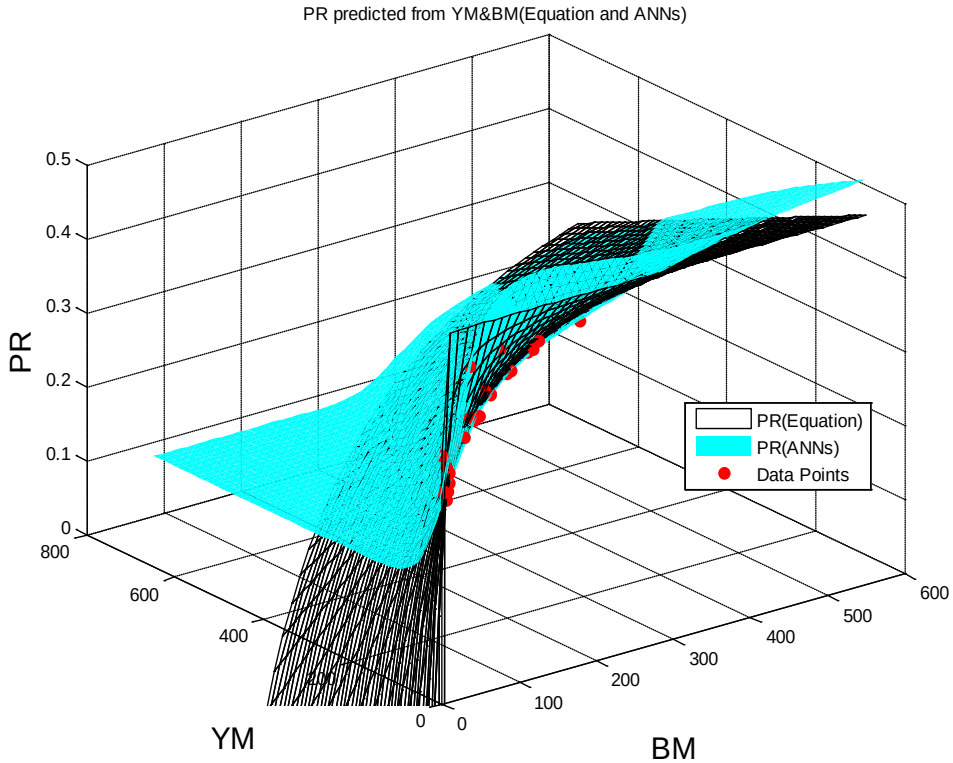
(e)


SM predicted from BM&PR(Equation and ANNs)
SM(Equation)
SM(ANNs)
Data Points
SM
BM
PR

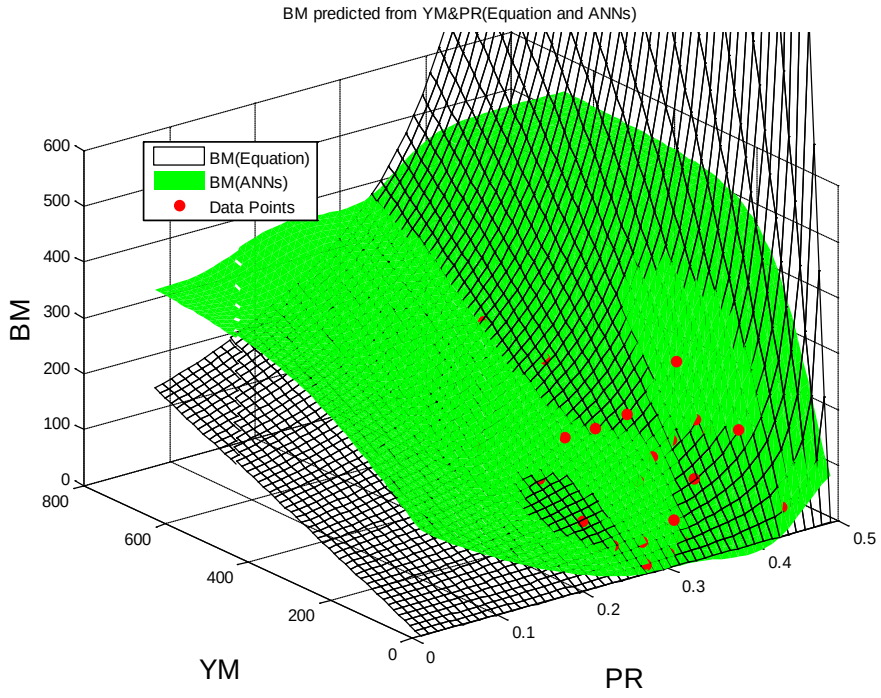
(f)


PR predicted from SM&BM(Equation and ANNs)
PR
BM
SM
PR(Equation)
PR(ANNs)
Data Points
0.5
0.4
0.3
0.2
0.1
0
600
500
400
300
200
100
0
0
50
100
150
200
250
300

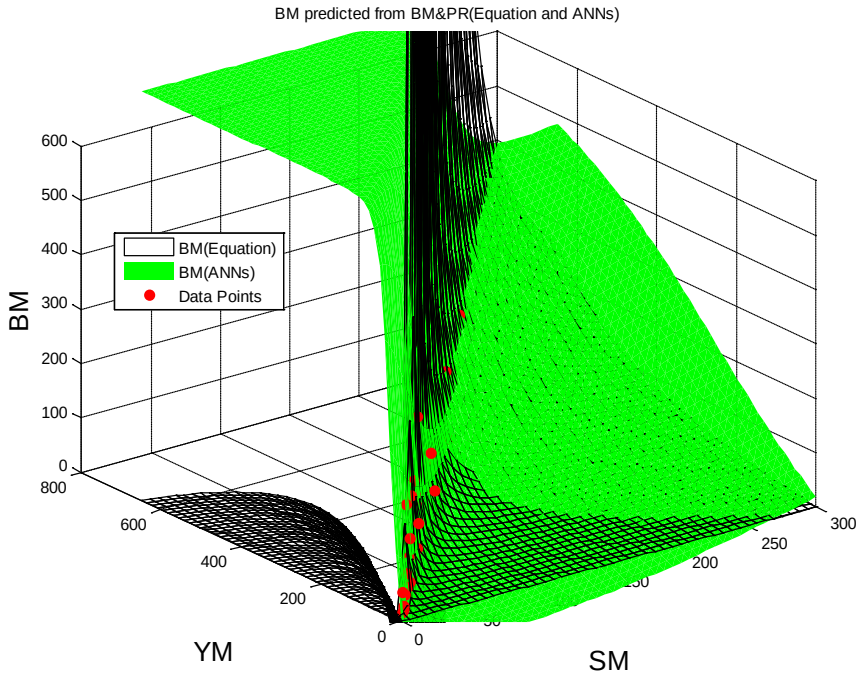
(g)


PR predicted from YM&BM(Equation and ANNs)
PR
YM
BM
PR(Equation)
PR(ANNs)
Data Points
0.5
0.4
0.3
0.2
0.1
0
800
600
400
0
0
100
200
300
400
500
600

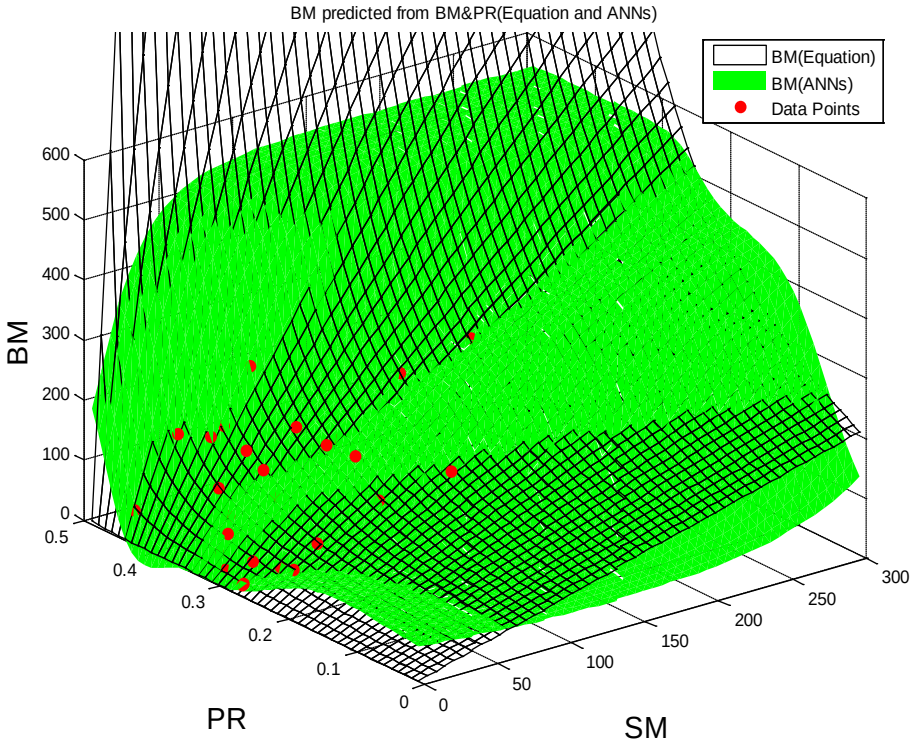
(h)


BM predicted from YM&PR(Equation and ANNs)
BM(Equation)
BM(ANNs)
Data Points
BM
YM
PR
600
500
400
300
200
100
0
800
600
400
200
0
0
0.1
0.2
0.3
0.4
0.5

(i)


BM predicted from BM&PR(Equation and ANNs)
BM(Equation)
BM(ANNs)
Data Points
BM
YM
SM
600
500
400
300
200
100
0
800
600
400
200
0
0
200
250
300

(j)



(k)

Figure A-1 Ternary order correlations between elastic properties are reflected in 3D dimensions. Young's modulus is denoted as YM, shear modulus is denoted as SM, bulk modulus is denoted as BM and Poisson's ratio is denoted as PR. The correlating equations, which are listed in Table 1-2, are indicated as mesh grids. While the correlations derived from ANNs are indicated as colourful surfaces. Yellow is for Young's modulus, blue is for shear modulus, green is for bulk modulus, and cyan is for Poisson's ratio. The values of 27 elements (which are used to construct ANNs) are plotted as red data points.