COMPLEX MECHANICAL PROPERTIES OF STEEL

Radu Calin Dimitriu
Department of Materials Science and Metallurgy
University of Cambridge
Churchill College

Facilis descensus Averno

Publius Vergilius Maro

PREFACE

This dissertation is submitted for the degree of Doctor of Philosophy at the University of Cambridge. The research reported herein was conducted under the supervision of Professor H. K. D. H. Bhadeshia in the Department of Materials Science and Metallurgy, University of Cambridge, between October 2005 and November 2008.

This work is to the best of my knowledge original, except were acknowledgement and references are made to previous work. Neither this, nor any substantially similar dissertation has been or is being submitted for any degree, diploma or other qualification at any other university or institution. This dissertation does not exceed the word limit of 60,000 words.

Part of the work described herein has been published or has been accepted to be published in the following publications:

R.C. Dimitriu and H.K.D.H. Bhadeshia. Hot Strength of Creep Resistant Ferritic Steels and Relationship to Creep Rupture Data. *Materials Science and Technology*, 23:1127-1131, 2007.

R.C. Dimitriu, H.K.D.H. Bhadeshia, C. Fillon, and C. Poloni. Strength of Ferritic Steels: Neural Network and Genetic Programming. *Materials and Manufacturing Processes*, 24:10-15, 2009.

H.K.D.H. Bhadeshia, R.C. Dimitriu, S. Forsik, J.H. Pak and J.H. Ryu. On Performance of Neural Networks in Materials Science. *Materials Science and Technology*, in press.

Radu Calin Dimitriu

January 2009

Acknowledgements

I would like to express my sincere thanks and gratitude to my supervisor Professor H. K. D. H. Bhadeshia for his support, guidance and enthusiasm.

I also want to acknowledge all the Phase Transformation and Complex Properties Research Group from the Department of Materials Science and Metallurgy, University of Cambridge, especially Mathew, Mohamed, Stephane, Saurabh, Amir, Gonzalo, Richard, for their assistance and the relaxing tea and coffee breaks.

I am greatly indebted to my wife Anca, my parents Octavian and Domnica, my family Oana, Grigore, Maria, Aurelian, Carmen, Andra, Maria, Ioan, Adrian, Monica, Marinel and Raluca, my friends Marius, Ioan, Ovidiu, Marcel and Mihai for their assistance, devotion and understanding throughout my life.

I express my sincere gratitude to the European Union which through the Marie-Curie Actions sponsored my Early Stage Career Fellowship which made it possible for me to study in Cambridge.

ABSTRACT

Whereas considerable progress has been reported on the quantitative estimation of the microstructure of steels as a function of most of the important determining variables, it remains the case that it is impossible to calculate all but the simplest of mechanical properties given a comprehensive description of the structure at all conceivable scales. Properties which are important but fall into this category are impact toughness, fatigue, creep and combinations of these phenomena.

The work presented in this thesis is an attempt to progress in this area of complex mechanical properties in the context of steels, although the outcomes may be more widely applied. The approach used relies on the creation of physically meaningful models based on the neural network and genetic programming techniques.

It appears that the hot–strength, of ferritic steels used in the power plant industry, diminishes in concert with the dependence of solid solution strengthening on temperature, until a critical temperature is reached where it is believed that climb processes begin to contribute. It is demonstrated that in this latter regime, the slope of the hot–strength versus temperature plot is identical to that of creep rupture–strength versus temperature. This significant outcome can help dramatically reduce the requirement for expensive creep testing.

Similarly, a model created to estimate the fatigue crack growth rates for a wide range of ferritic and austenitic steels on the basis of static mechanical data has the remarkable outcome that it applies without modification to nickel based superalloys and titanium alloys. It has therefore been possible to estimate blindly the fatigue performance of alloys whose chemical composition is not known.

Residual stress is a very complex phenomenon especially in bearings

due to the Hertzian contact which takes place. A model has been developed that is able to quantify the residual stress distribution, under the raceway of martensitic ball bearings, using the running conditions.

It is evident that a well–formulated neural network model can not only be extrapolated even beyond material type, but can reveal physical relationships which are found to be informative and useful in practice.

Contents

Nomenclature

α control parameter

β control parameter

δ delta ferrite

ϵ_0 creep rate

ϵ_f fatigue ductility coefficient

η fatigue ductility exponent

π constant (3.14159265...)

ρ dislocation density

σ uncertainty

σ_Y hot–strength

$\sigma_{0.2}$ 0.2% proof stress

σ_f flow stress

σ_i friction stress

σ_y stress necessary to propagate yield

θ bias

A number of models in the committee

a crack length

a_1 reliability constant

a_2 material fatigue properties constant

a_3 lubricant constant

C material constant

c_i chemical composition

D dynamic load

d average grain size

E elastic modulus

E_D test error

E_E calculated error

E_w regulator

f function

h neuron

i node

K stress intensity factor

k constant

K_c critical stress intensity factor

K_{eff} effective stress intensity factor

K_{Ic} fracture toughness

K_{max} maximum stress intensity factor

K_{min} minimum stress intensity factor

K_{open} stress required to open a crack

K_{th} stress intensity factor threshold

k_y constant

L life of the bearings

l exponent

LPE log predictive error

M objective function

MSE mean square error

N umber of cycles

n material constant

P equivalent load

p bearing geometry related exponent

Q activation energy

R universal gas constant

$RMSE$ root mean square error

S applied stress

S_{LM} martensite size partitioned by lower bainite

S_{UM} martensite size narrowed by upper bainite

SS solid solution strengthening

T temperature

t target value

T_A ambient temperature

T_a austenitising temperature

t_a austenitising time

T_C transition temperature

T_T test temperature

T_t tempering temperature

t_t tempering time

UTS ultimate tensile strength

V_B volume fraction of bainite

w weights

x input value

Y yield strength

y output value

Chapter 1

INTRODUCTION

The purpose of the work presented in this thesis is to assist the continuing progress of steel and by correlating the properties with the materials response to stress and strain facilitate the development of steels. The vast array of interconnected properties, explains the success of steel, and in this context the steels chemical composition and heat treatment influences the mechanical properties which in turn influences the response of the material to external forces.

The principal aim of modelling is to create increasingly complex models which will be able to calculate all the properties and can capture all the complex interactions in steels, thus reducing the need for trial and error development of steel and subsequent testing. Creating such a model requires vast amounts of data and it is impossible to accomplish this in one single try, so a more systematic approach is required where individual properties and interactions are successfully modelled, following this the data and knowledge of these models are combined to create a truly general model.

Work toward this aim has already been completed with models covering the creep strength of austenitic stainless steel [1], prediction of

Vickers hardness in austempered ductile irons [2], modelling precipitation sequences in power plant steels [3, 4], modelling irradiation hardening in low-activation steels [5], predicting the martensite–start temperature of steels [6], modelling the Charpy toughness [7], recristalization in mechanically alloyed materials [8] and many more. A schematic representation, which is by no means exhaustive, of the models created until present including the models described in this thesis is given in Fig. 1.1.

irradiation hardening fatigue crack growth rates precipitation during tempering martensite start temperature

creep strength strength ductility recrystallisation

Vickers hardness hot strength bainite start temperature

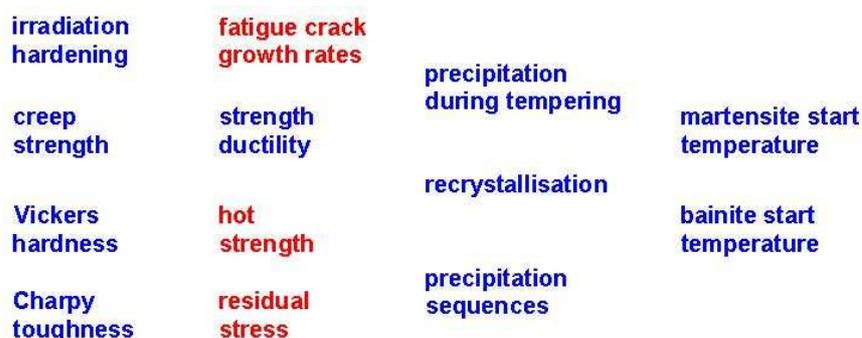Charpy toughness residual stress precipitation sequences

Figure 1.1: Schematic representation of the models created which will contribute towards a general applicable model. The red coloured models are described in the present thesis whilst the blue ones have already been developed.

Bayesian neural networks are an established materials modelling technique known for their ability to handle complex data hence the reason why they were used in the present work. The hot–strength of ferritic steels was modelled as a function of chemical composition and heat treatment whereas the fatigue crack growth rates of steels were modelled as a function of mechanical properties.

An overview of the problems associated with fatigue, strength, creep and residual stress of steel and their influencing factors is provided in Chapter 2.

Chapter 3 reviews the problems of modelling complex properties

and describes the Bayesian neural network and genetic programming methods used in the present work.

Chapter 4 describes the neural network model which enables the estimation of hot–strength presenting a connection between hot–strength and creep. This chapter also presents a comparison between neural networks and genetic programming.

Chapter 5 describes the construction and exploitation of a neural network model which calculates the fatigue crack growth rates in steels and its remarkable ability to calculate the rates in nickel, titanium and aluminium alloys.

The development of residual stress due to rolling contact fatigue conditions in ball–bearings is detailed in Chapter 6.

Chapter 7 reviews the conclusions reached from the previous chapters and explores further research pathways suggested by the present work.

Chapter 2

LITERATURE REVIEW

2.1 FATIGUE CRACK GROWTH

Metal subjected to continual and fluctuating stress will fail at a stress much lower than that required to cause fracture on a single application of the load. These classes of failures occurring under dynamic loading are labelled *fatigue failures* because they should occur only after a substantial service life. Fatigue failures can be dangerous because they sometimes occur with little warning, resulting in a brittle looking fracture, with no obvious deformation at the parted surface [9].

Fatigue failures occur only if three main factors are met: the maximum tensile stress has a sufficiently large value, there is a fluctuation in the stress applied and the number of cycles is high enough for failure to occur. In addition to this, stress concentration, temperature, corrosion, residual stress and basic mechanical properties influence fatigue failure [10].

2.1.1 Fatigue Failure Concepts

There are four concepts applied in fatigue, that are concerned with design against failure [11].

Infinite–Life Design this design criterion is based on keeping the stresses at a fraction of the fatigue limit, which in some materials such as steel defines the minimum stress at which a fatigue can occur. This is used for parts subjected to very large number of cycles of uniform stress.

Safe–Life Design is based on the assumption that the part is initially free from flaws and has a finite life in which to develop a critical crack. The fatigue life is assumed even at constant stress amplitude to exhibit statistical scatter. This concept is used in calculating the life of rotating components in jet engines, pressure vessels and bearings.

Fail–Safe Design assumes that fatigue cracks will not lead to failure before they can be detected and repaired. The fail–safe designs employ multiple load paths and crack stoppers built into the structure, along with strong regulations and criteria for inspection and detection of cracks. This method is used extensively in the aircraft industry.

Damage–Tolerant Design is based on the assumption that finite cracks will exist in an engineering structure. Fracture mechanics are employed to determine whether the cracks will grow large enough to cause failure, between regular inspections. This method is applied to materials with high fracture toughness, slow crack growth and when reliable non–destructive evaluation methods are available.

2.1.2   *S-N* Curve

Interest in fatigue started before 1870 when the first investigation into characterising the fatigue behaviour of materials was published [12]. Wohler found that fatigue occurs by crack growth from surface defects until the product can no longer support the applied load. This led to the development of a relationship between the magnitude of applied cyclic stress and the number of the cycles that a component could withstand before complete failure. The relation takes the form of *S-N* curves or also called Wohler curves, Fig. 2.1, where S represents stress and N the number of cycles to failure at that stress.
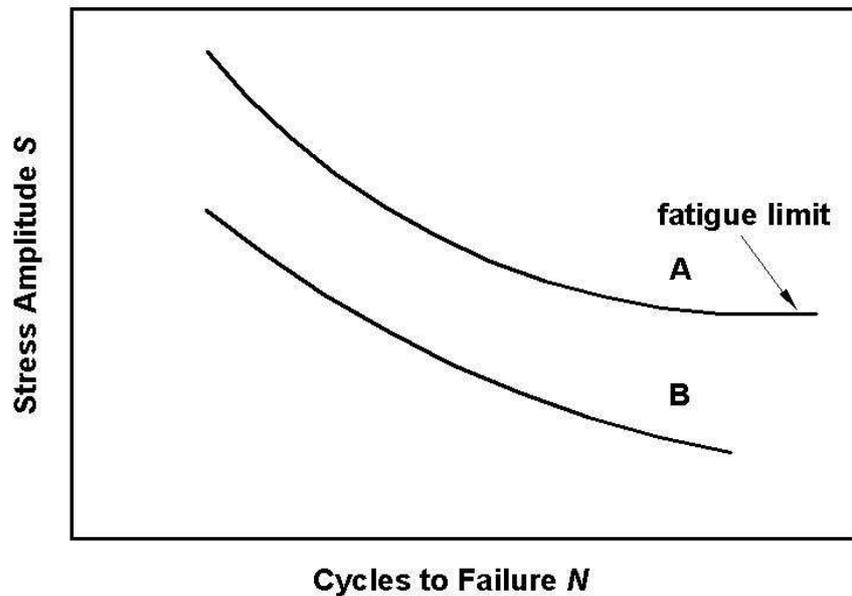

Figure 2.1: Schematic $S - N$ curve. Curve A is typical for steel, curve B is for aluminium. For steel, there is a threshold stress below which, the life of the component can be infinite. Aluminium, must eventually fail regardless of the amplitude of the applied stresses.

The *S-N* curve is the easiest and best known method of presenting fatigue data. The stress value plotted can be the maximum, minimum or the alternating stress and usually is measured in reversed bending so

that the mean stress is zero. The *S-N* curves represent fatigue failures for a large number of cycles ($N > 10^5$ cycles), the stress conditions are elastic on a gross scale, but the metal deforms plastically on a localised scale.

For steel the *S-N* curve becomes horizontal at a limiting value of stress, known as the *fatigue limit*. The material can, when the stress is below this limit, withstand an 'infinite' number of cycles without failure. Non–ferrous[1] metals and copper alloys have an *S-N* curve which slopes gradually downward with increasing number of cycles. These materials do not have a fatigue limit so design is based on *fatigue strength* selected according to the prescribed life of the component.

The procedure for creating the *S-N* curve is to test 8 to 12 specimens at different stress levels. The first specimen is tested at high stress (two thirds of the static tensile strength) where failure is expected after a short number of cycles. The stress is then decreased for subsequent specimens until one or two do not fail in the specified number of cycles (at least 10^7). The highest stress at which failure does not occur is considered the fatigue limit.

The *S-N* curves are generated by test specimens that have a very good surface finish, and virtually no surface defects. Nevertheless, in practice the majority of components do not have such a good surface finish, so it is better to evaluate the life of the components based on the ability of the component to function with a crack that has a certain length. The crack can grow with each cycle at a rate which defines the material resistance.

---

[1]An exception are the Al-Mg alloys which have a fatigue limit, since Mg causes similar effects to C in steel *i.e.* segregates to dislocations.

2.1.3   Fatigue Crack Growth

Fatigue is influenced by material properties such as yield strength, elongation and fracture toughness; imposed parameters such as the cyclic stresses applied, their magnitude, pattern and frequency must clearly also influence life. Other important factors are the size of internal and surface defects, and the orientation in relation to the applied stresses. Service temperature and corrosion also play an important role in the development of fatigue [13].

Measuring the Crack Length

Fatigue crack growth rate measurements are conducted on machines capable of applying constant–load amplitude cycles at different frequencies. The standard method for constant–amplitude fatigue crack growth rate is described in ASTM E 647 [14]. Specimens are pre–cracked by cycling at a load amplitude equal to or less than the test–load amplitude. The crack length is measured as a function of the number of cycles as shown in Fig. 2.2. At a known crack length a it is possible to calculate ΔK (stress intensity range) as a function of load and specimen geometry.

The basic data in fatigue crack propagation tests are the number of cycles and crack length. The number of cycles is readily monitored by electronic or mechanical counters. There are several methods employed to determine the crack length[13, 15–17]:

- direct visual measurement, usually for thin sheet specimens where crack dimensions can be visually inspected;

- ultrasonic sensing, where a probe is attached to the top edge of the specimen and is moved along the surface such that a constant
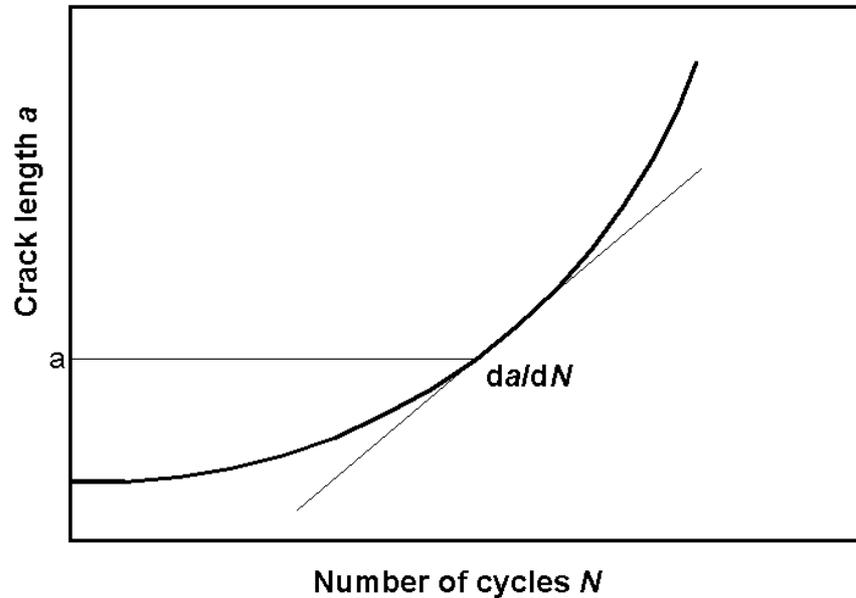
Figure 2.2: The basic data: crack length versus number of cycles.

relationship between the crack tip and the position of the probe is maintained, the movement of the probe provides the amount of crack advance;

- electrical potential drop measurements a constant d.c. current is passed through the specimen and the change in resistance of the specimen as the crack grows is detected by measuring the change in potential across the opened notch;

- the compliance (displacement per unit load) method a.c. current is passed through the specimen and a clip–gauge mounted across the crack mouth measures the compliance as the crack grows.

Crack Propagation

Fatigue failures have been studied extensively over many decades [12, 18–20]. Most begin with the slow growth and propagation of pre-

existing cracks or defects, under fluctuating stress. Once the crack reaches a critical size, the remaining cross-section cannot withstand the applied stress, leading to sudden and complete failure. For this reason crack propagation relations are used to enable the implementation of fail–safe design [21].

Fatigue crack propagation became a tool in an engineering context with the comprehension that crack length versus the number of cycles at different stresses can be represented by a general plot of $\log \frac{da}{dN}$ versus $\log \Delta K$ as illustrated in Fig. 2.3.
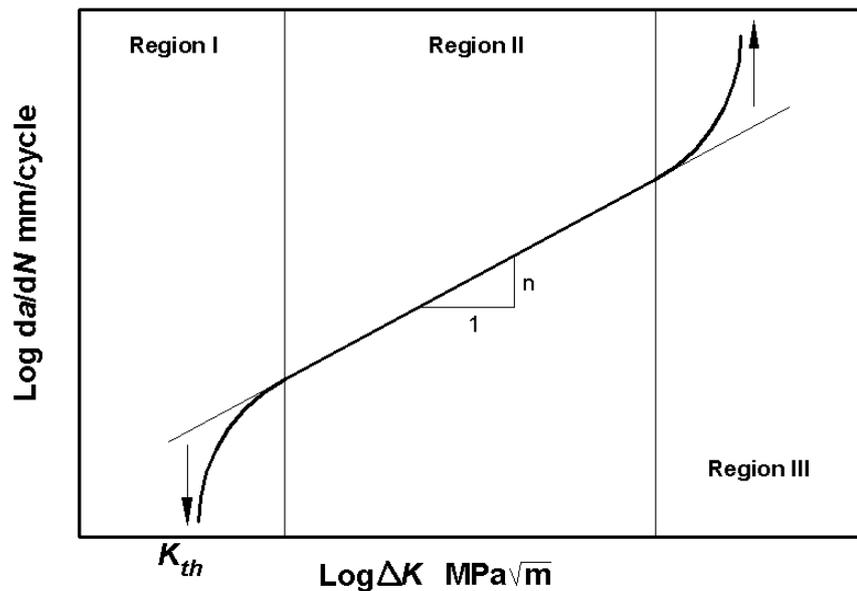

Figure 2.3: Fatigue crack growth rate versus ΔK.

The stress intensity factor K, is used to predict the stress state near the tip of a crack, caused by a remote load or residual stresses. When this stress state becomes critical a crack will grow and the material will fail. It is a theoretical construct applicable to a homogeneous elastic material, used to provide a failure criterion.

There is a minimum threshold in the stress intensity required for the

crack to grow, after this point the curve can be separated into three regimes. In the first region, the crack growth rate varies sharply with ΔK once the ΔK_{th} is exceeded. The second 'Paris law' region is where Equation 2.1 is applicable with a linear relationship between $\log \frac{da}{dN}$ and logarithm of K in this region most of the fatigue life is consumed. Finally the third region is where K_{max} approaches the critical K for instability; this zone is associated with a rapid growth of the crack finally leading to fracture by mechanisms akin to static testing, for example in tension alone [22].

The equation by which region two is governed:

$$\frac{da}{dN} = C(\Delta K)^n \tag{2.1}$$

where a is the crack length, N is the number of fatigue cycles, ΔK is the cyclic stress intensity range whereas C and n are materials constants [22].

The graph presented in Fig. 2.3 is obtained from the data used to plot the graph in Fig. 2.2 as follows: at a crack length a it is possible to calculate ΔK from the maximum and minimum loads applied. At the same crack length the value of da/dN is determined from the gradient of the curve in Fig. 2.2. It is possible to obtain a figure similar to Fig. 2.3 by plotting $\log \frac{da}{dN}$ *vs* $\log \Delta K$.

Log $\frac{da}{dN}$ versus $\log \Delta K$ graphs are considered a good way of representing crack propagation given that defects will be present within the material. *S-N* curves in contrast are considered a representation of crack initiation due to the fact that at high stress the specimen spends little time cracked and failure arrives quickly.

2.2 ROLLING CONTACT FATIGUE

Contact fatigue is a common type of failure encountered in bearings, rails, gears and valves. It differs from structural fatigue (bending or torsion) in that the cyclic stress originates in Hertzian contact, when a curved surface rolls over another curved or flat surface under normal load. The life of the structures subjected to contact fatigue is limited by surface disintegration. The parameters that influence fatigue are contact pressure, material properties, lubrication properties, surface roughness, amount of relative sliding or slip in the contact, residual stresses and inclusion size and nature [23, 24].
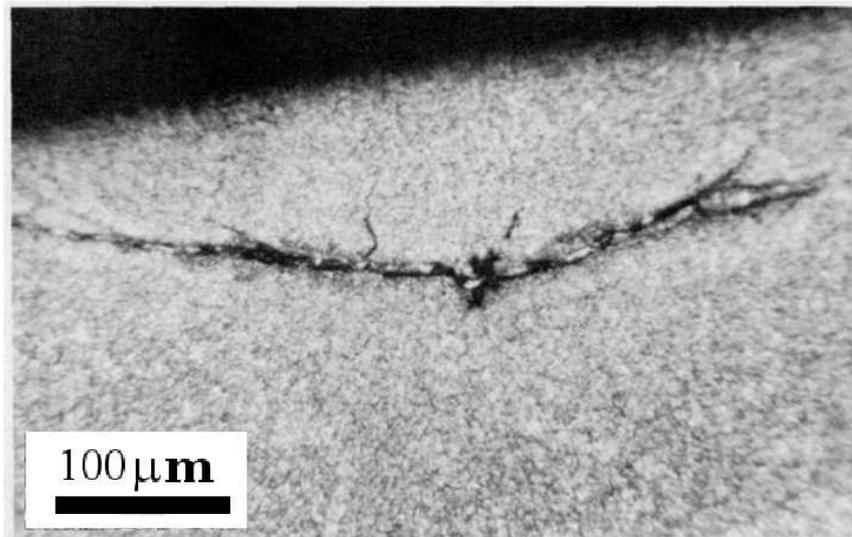

Figure 2.4: Optical micrograph of a crack originating at an oxide–type inclusion below the surface of a spur-gear tooth [23].

Failures, as the one presented in Fig. 2.4, due to rolling contact fatigue will initiate from an inclusion or at the interface between layers with different properties in the matrix. The initiation of the failure will be triggered by the cyclic stresses developed below the surface near the contact zone [25]. The useful life of a component is a direct function of

tensile stress, stressed volume, alternating shear stress, number of stress repetitions, depth of alternating shear stress and residual stress [26].

2.2.1 Residual Stress

Residual stress is that which remains in a body after processing or use [27, 28]: it can be detrimental as it may reduce the tolerance of the material to an externally applied force, such as the load that a structure has to endure. In general, tensile residual stress reduces fatigue life, and compressive residual stress improves it. However, the details of how the residual stress patterns affect the life of components are not easy to assess, because the residual stresses are self equilibrating over a variety of distances (depending on the type of residual stress) and also due to the multitude of factors influencing them [29].

In an ideal situation a residual stress will superimpose on an elastic stress and this will result in an enhanced fatigue life for a component. In the case of bodies of revolution the residual stresses which needs to be taken into account are the tangential and axial components, because the radial stress is generally minor [30].

Causes of Residual Stress

Residual stresses can be introduced by thermo–mechanical processes such as machining, coating, sand blasting and shot peening, which induce short–range residual stress or by heat treatment, rolling, forging, and welding which induce long–range residual stress.

The stress can develop during the heterogeneous processing of steel due to the a volume expansion associated with the transformation from austenite to various body-centred cubic phases [31]. Heat treatment of

the material can add to the residual stress. During carburizing, residual compressive stress is developed in the hardened surface layers, which expand owing to their higher carbon content; the core resists this expansion thus putting the surface layers in compression [32]. Grinding, hard turning and other machining operations induce residual stress in the component, due to the mechanical deformation of the surface layers [33, 34]. Tensile residual stress is found in the surface layers because of the tempering owing to the heat generated by the grinding, which results in shrinkage, while the base material remains unaffected [35–37].

There are many causes why a component subjected to contact fatigue can fail. If the lubrication is not appropriate, the component will be subjected to wear. When the lubrication is not filtered, indents will appear in the contact surfaces. If the load is too high, seizure will occur, but if all of these are avoided then the only alternative and viable cause of failure is due to the continuous accumulation of residual stress until spalling occurs. Such accumulation can be further enhanced by plastic deformation and strain or stress-induced phase transformations due to service load [38–40].

The residual stress pattern usually found beneath the raceway of new and used bearings is illustrated in Fig. 2.5. The characteristic feature of the residual stress for a new bearing is the high compressive stress at the surface. The stress then decreases with depth, gradually levelling around zero [41, 42].

For the used bearings, as shown in Fig. 2.5, an increase in the number of cycles the compressive residual stress peaks at a depth between 0.1 and 0.5 mm, even though all the other parameters involved (load, speed, temperature, lubricant) were kept constant [41, 43].
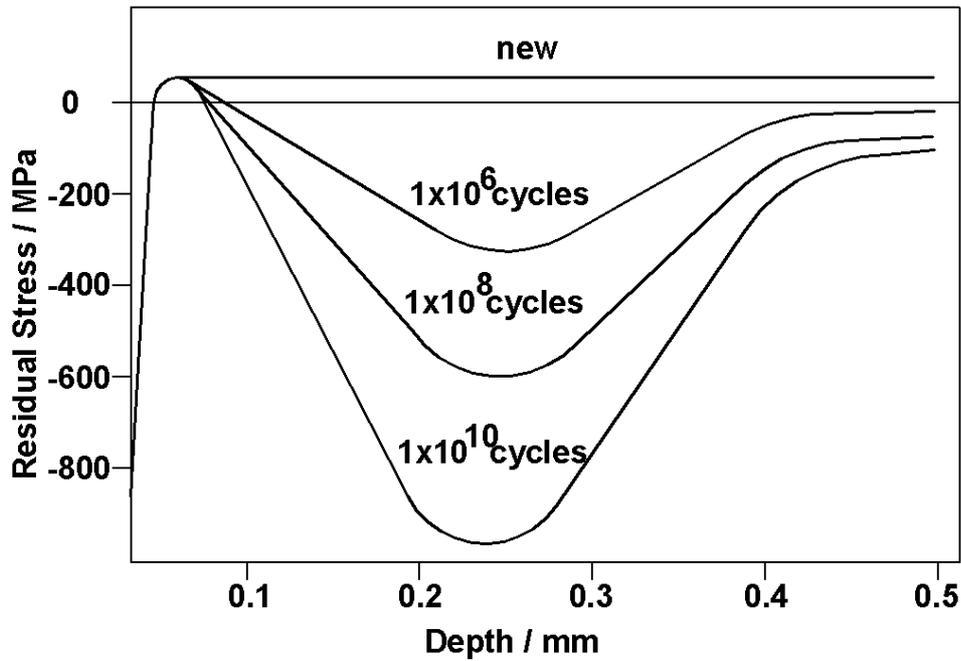
Figure 2.5: A typical residual stresses profile for new and used bearings [41].

Classification of Residual Stress

Residual stress is classified according to the scale over which it self-equilibrates. Long–range stresses (type I) equilibrate over macroscopic dimensions. Type II residual stress equilibrates over a number of grain dimensions. Type III stresses on the other hand exist over atomic dimensions and balance within the grain [27, 28].

Macrostress (type I stress or body stress) can arise from a non-uniform expansion, contraction or shear. It can result from mechanical, chemical or thermal operations performed on the body. Chemical changes propagating from the surface to the interior give rise to macrostress. Thermal volume changes induce residual stress if the body is stress–free at the time when the temperature distribution in it is not

uniform and then becomes so. Macro residual–stresses can accumulate during cyclic loading, because there will be some plastic deformation, the material will strain harden, its capacity for plastic stress redistribution will decrease and the stress amplitude will rise in spite of the constant amplitude of the load acting upon the structural part [44].

Microstress (type II and III) results from microstructural inhomogeneities such as inclusions with thermal expansion coefficients different from that of the surrounding matrix. There are many types of microstresses, due to the grain structure (thermal stresses due to anisotropic thermal expansion of the grains), from the inherent inhomogeneity of deformation process (stress accompanying martensitic transformation) and microstresses around inclusions [45].

Measurements of Residual Stresses

There are many techniques available for measuring residual stress, varying in accuracy, spatial resolution, penetration below the surface and completeness of information [28, 46, 47].

Mechanical stress measurements are based on the destruction of the equilibrium state in the component. The stress is determined via the progressive relaxation of the component. Mechanical measurements of this type are sensitive only to macroscopic residual stress. Hole drilling, ring core technique, bending deflection method, sectioning method and compliance methods fall in this category.

Diffraction measurements are based on the study of variations in the inter–planar spacing of the polycrystalline material using (electron diffraction, neutron diffraction and X-ray diffraction). They can be used to study all three kinds of residual stress.

Magnetic stress measuring methods rely on the interaction between

magnetisation and elastic strain in ferromagnetic materials. Two magnetic stress-measuring methods are in use, the Barkhausen noise method and the magnetostriction method.

Ultrasonic techniques are based on variations in the flight time difference of ultrasonic waves in the material. The ultrasonic and magnetic methods are not destructive and are sensitive to all three kinds of residual stresses but cannot distinguish between them.

2.3 STRENGTH OF STEELS

Since iron was discovered, blacksmiths and later metallurgists tried in various ways to enhance its properties [48–50]. One of the most important properties of steel is its strength. There are different ways by which the strength can be increased, and some of them are described in the following paragraphs.

2.3.1 Strengthening Mechanisms

Solid Solution Strengthening

Solid solution strengthening is a mechanism which relies on adding a foreign solute into the base structure. The approximate rules according to the way an element dissolves in a base were formulated by Hume-Rothery [51, 52]:

- if a solute differs in its atomic size by more than about 15% from the host, then it is likely to have a low solubility in that metal. The 'size factor' is said to be unfavourable;

- if a solute has a large difference in electronegativity (the ability to attract electrons) or electropositivity (the ability to donate electrons) when compared with the host, then it is more likely to form a compound. Its solubility in the host would therefore be limited;

- a metal with a lower valency (a measure of the number of chemical bonds formed by the atoms of a given element) is more likely to dissolve in one which has a higher valency.

Depending on the size of the solute atom, a substitutional or an interstitial solid solution can be formed. In both cases, the overall crystal structure of the base metal is unchanged, apart from lattice parameter variations as a function of composition. Substitutional solid solution forms when the solute atom is comparable in size to that of the solvent. Substitutional solutes cause symmetrical distortions in their vicinity, resulting in hydrostatic stresses which lead to moderate interactions with dislocation stress fields. The strengthening achieved by substitutional solute is greater the larger the difference in atomic size of the solute from that of the iron. Because the substitutional atoms replace the iron atoms, the distortion in the lattice is limited so the strengthening effect is smaller than for the interstitial solid solution [54]. Fig. 2.6 represents sizes of various alloying elements, both interstitial and substitutional, compared to iron [53].

Interstitial solid solution forms when the size of the solute is markedly smaller than that of the solvent. Depending on the symmetry of the interstitial site, this can lead to anisotropic strains with huge interactions with dislocations and have a substantial hardening from small amounts of solute. In the case of iron, carbon and nitrogen are the most important interstitial strengtheners, they have the tendency to reside in the octahedral sites Fig. 2.7 where they distort the lattice, resulting a powerful interaction with the dominant shear component of a dislo-
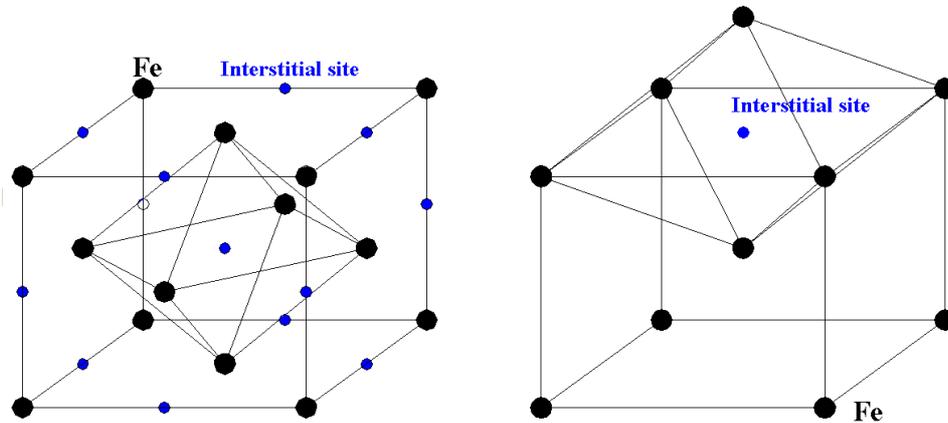
Figure 2.6: Atomic size of alloying elements in iron and their solubility: the squares represent the insoluble elements, the filled triangles and the filled circles are the elements that expand the γ field, whilst the circles and the filled squares represent the elements which restrict the γ field. The dashed lines represent a deviation of 15% from the atomic size of iron [53].

Figure 2.7: On the right the rectangular octahedron interstice in ferrite, whilst on the left the octahedron interstice in ferrite. Two of the axes are longer than the third (vertical) which leads to a tetragonal distortion when the site is occupied by an interstitial atom [55].

cation strain field. Thus, carbon in ferrite causes a tetragonal strain and intense hardening, whereas the same solute in the regular octahedral interstices of austenite causes only isotropic strains and minimal hardening. The maximum solubility of carbon in austenite under equilibrium condition is about 2 wt% at a temperature of 1150°C, whereas its solubility in the smaller octahedral interstice in ferrite is only 0.025 wt% at a temperature of 720°C [56, 57].

Dislocations

Dislocations are of two primary types, edge and screw dislocations; mixed dislocations, shown in Fig. 2.8 represent a combination of these. Dislocations are lines between slipped and unslipped regions on a slip plane, their movement causes the relative shear displacement of the atoms on opposite sides of the glide plane [58].

An edge dislocation is a defect where an extra half-plane of atoms

is introduced partly through the crystal, distorting nearby planes of
atoms. The boundary is perpendicular to the slip direction and its
passage causes a relative displacement in a direction perpendicular to
itself.

A screw dislocation is a defect in the crystal structure where half
way through the crystal there is a cut and the two parts have slipped.
The boundary is parallel to the slip direction and its passage causes
relative displacement in the direction parallel to itself. It also gets its
name because of the path taken around it following the atomic planes
results in a helix or a screw.
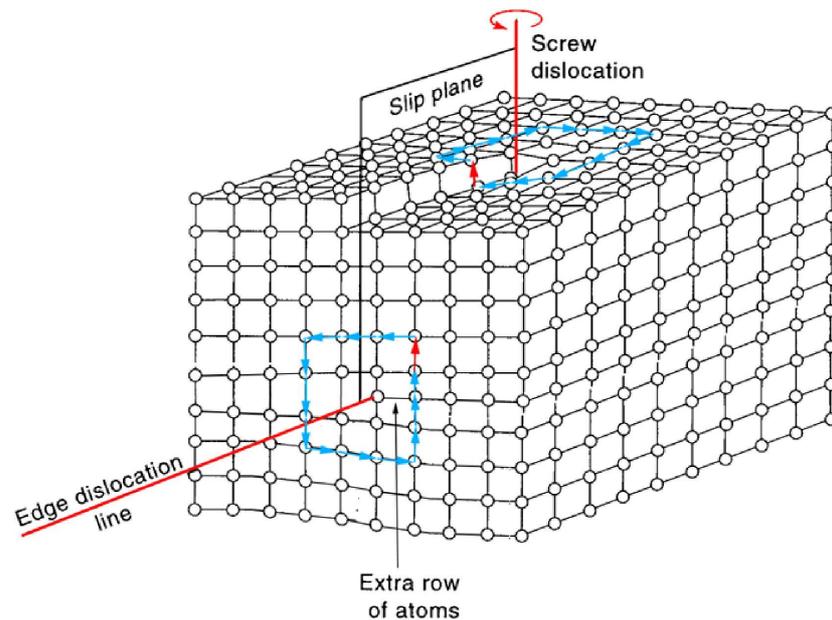
Figure 2.8: A schematic representation of a mixed dislocation (formed
from a screw dislocation on top and an edge dislocation on the side)

The Burgers vector of a dislocation is a crystal vector, specified
by Miller indices, that quantifies the difference between the distorted

lattice around the dislocation and the perfect lattice. To determine the vector, in a faulted crystal a deformed rectangle, whose lengths and widths are integer multiples of the interatomic vector, is drawn on a plane which encompasses the dislocation line, the number of lattice vectors travelled along each side of the loop is noted. Than the same circuit will be performed in the crystal without the dislocation (perfect crystal) than a closure failure will occur, this closure failure in the perfect crystal represents the Burgers vector presented in Fig. 2.9 by the green arrow which define the direction and magnitude of the Burgers vector [59].
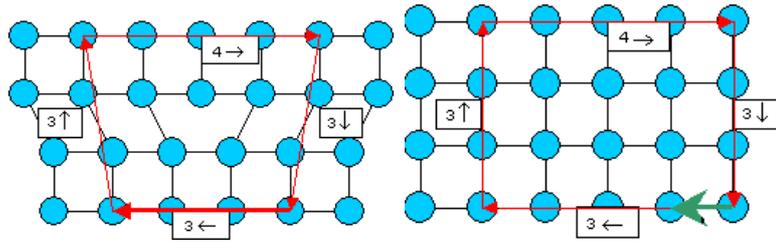

Figure 2.9: Schematic representation of a Burgers vector [60]

One of the ways by which dislocations affect the strength of steels is through the flow stress σ_f which is related to the dislocation density ρ by the relation:

$$\sigma_f = k\sqrt{\rho} \tag{2.2}$$

where k is a constant incorporating the shear modulus and the Burgers vector.

There are arrays of structures which act as pinning points for dislocations and hence lead to hardening [61]:

- point defects such as vacancies or solute atom create stress fields that interfere with the movement of dislocations;

- precipitates create barriers which dislocations have to overcome, either by cutting through the encountered particles or by looping;

- grain boundaries are also barriers to dislocations motion which require lattice ordering to move through a material, so the grain boundary stops the dislocation from moving.

Grain Size Refinement

Grain size refinement is one of the most useful strengthening processes because it increases the toughness and the strength of the material at the same time, as proposed by Hall and by Petch [62, 63]. Toughness is the material's resistance to fracture which can be transgranular when the fracture passes through the grain, or intergranular when the fracture follows the grains of the material. Both kinds of fracture will require more energy to propagate the crack to a big enough size to cause rupture if the grains are smaller.

The relationship is based on the assumption that the grain boundaries act as obstacles for slip dislocations, causing them to pile up at the interfaces. The number of dislocations possible in a pile up is greater when the grain size is large. The pile-ups therefore produce a stress concentration in the adjacent grain, the magnitude of which increases with the number of dislocations in the pile–up and the applied stress. Thus, in coarse–grained materials the stress concentration will be greater than in those which are fine–grained, making it more difficult to transmit plasticity across fine grains [64]. The Hall-Petch relationship is [62, 63]:

$$\sigma_y = \sigma_i + k_y d^{-\frac{1}{2}} \qquad (2.3)$$

where σ_y in the stress required to propagate yielding across grains, σ_i

is the friction stress opposing the movement of dislocations within the grains, k_y is a constant and d is the average grain size.

The size of the grain for which this theory applies cannot be decreased indefinitely because it becomes more difficult for grains to accommodate dislocations and the grains may then undergo sliding at the boundaries. There is considerable interest in developing nano–structured steels with sizes as small as 20 nm [65, 66]. It is unlikely that the Hall-Petch equation applies at such small sizes since the concept of dislocation pile–ups over very small distances than becomes difficult to justify.

Downey *et al.* working on 0.012 C, 0.153 N, 17.52 Cr, 12.86 Ni, 0.071 Nb in wt% steel observed that twins encompass a large fraction of the grain boundaries so they developed a model based on the Hall–Petch relationship that takes into account the twins as well as grain size and chemical composition. The relationship is based on the assumption that twin interfaces are as effective barriers as grain boundaries. In view of the fact the twin interfaces do allow the partial transmission of piling dislocations, this prediction is expected to produce an upper bound yield strength value [67].

The Hall–Petch relation considers only the mean grain size whereas in reality the grains form a population of stochastic nature with different sizes and shapes. Berbenni *et al.* developed a computer model which takes such variations into account. Based on grain to grain accommodation and it was found to be more relevant than analytical models treating the yield stress as simple mixture rules of components [68].

An investigation by Song *et al.* revealed that ultra fined grained steels with dual phase microstructure do not follow the Hall–Petch relationship, because a relatively small increment in stress is achieved in the dual phase steel when the ferrite grain size is refined from 19.4 to 0.8 μm [69]. Louchet *et al.* also working on nanometrer grain–sized

steels observed that the Hall–Petch relation does not cover very small grained steels due to the fact that instead of involving intermittent and correlated motion of interacting dislocations, strain proceeds through uncorrelated events of individual dislocation nucleation and propagation [70].

Bata and Pereloma developed a model to explain the Hall Petch relation which addresses the fundamental premise that application of an average stress is sufficient to induce the emission of a straight dislocation from a boundary, even though that stress is far below that required to overcome the highest energy barrier encountered by the emitted dislocation [71]. The model's validity was questioned because the averaging process emphasises weak long-range interaction of the emitted dislocation with the original boundary, is susceptible to perturbation by the dislocation structure introduced during straining, and overlooks the effect of boundary structure defects as heterogeneous dislocation sources [72].

Precipitation Hardening

Precipitation hardening is a process which relies on small particles dispersed in the materials matrix, which impede the movement of dislocations, to strengthen the material. The finer and numerous the precipitates are the more effective they are in strengthening a material [73].

The process to produce a fine dispersion of precipitates requires two steps. Firstly a supersaturated solution is produced by annealing at high temperature, followed by quenching to a lower temperature and an isothermal transformation (the process of annealing and quenching a supersaturated solution is called solution treatment) this is the first stage in the precipitation hardening process [74]. Then precipitation

of particles from the supersaturated solution occurs by nucleation and growth (this is the aging treatment *i.e.* holding at an intermediate temperature) [58].

The determining parameter is holding time and temperature. Too long time results in loss of hardness (called overaging), whilst aging for a too short period results that the optimum properties are not achieved. Too high temperature result in an increase in hardness in a very short time followed by a decrease, a too low temperature result that the process of precipitation may take too long or even be delayed until the optimum temperature is achieved [64].

2.4 CREEP

Creep is the term used to describe the tendency of a material to deform permanently with the aid of diffusion, in order to relieve stresses which may be much smaller than the yield strength measured in a tensile test. The rate of this deformation is a function of the material properties, exposure time, temperature and the applied load. Creep naturally becomes important when evaluating components that operate under high stresses and temperatures.

2.4.1 The Creep Curve

To determine the creep curve of a metal, a constant load is applied to a tensile specimen maintained at a fixed temperature, and the resulting strain is determined as a function of time.

A typical creep curve is illustrated in Fig. 2.10. Following an initial rapid elongation of the specimen the creep rate (ϵ_0) decreases with time, then reaches a steady state in which the creep rate changes little with
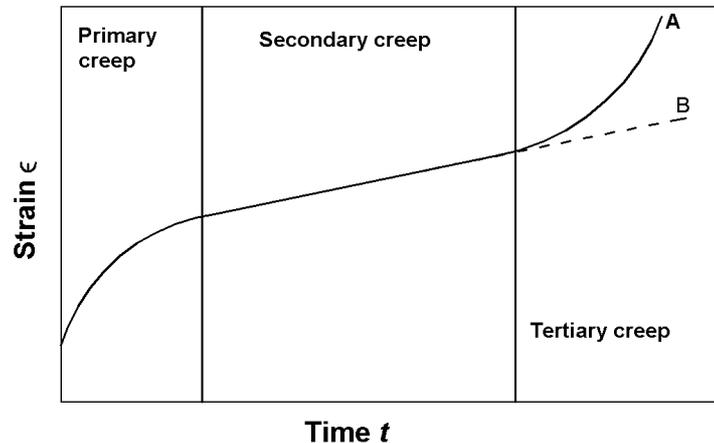
Figure 2.10: A typical creep curve showing the three steps of creep. Curve A constant load test; curve B constant stress test [75].

time, followed by accelerated creep which leads to fracture.

It is usual practice to maintain the load constant during creep test, leading to a curve similar to A in Fig 2.10. As the specimen elongates and decreases in cross–sectional area, the axial stress increases until fracture eventually occurs. Methods of compensating for the changes in the dimensions of the specimen to ensure constant-stress conditions have been developed [76, 77]. in these conditions the onset of stage three is delayed, curve B in Fig 2.10.

A constant–stress creep curve can be considered to represent the superimposition of two separate processes following elastic deformation [78]. The first is *transient creep* (Fig. 2.11c) with the creep rate decreasing with time. Added to this is a constant–rate *viscous creep* component shown in Fig. 2.11d. The strain ϵ_0 occurs practically instantaneously upon the application of the load and it is not entirely elastic.

There are three stages in the development of the creep curve. In the first stage, *primary creep*, the resistance of the material increases by

(a) Constant–stress creep curve.

(b) Instantaneous strain
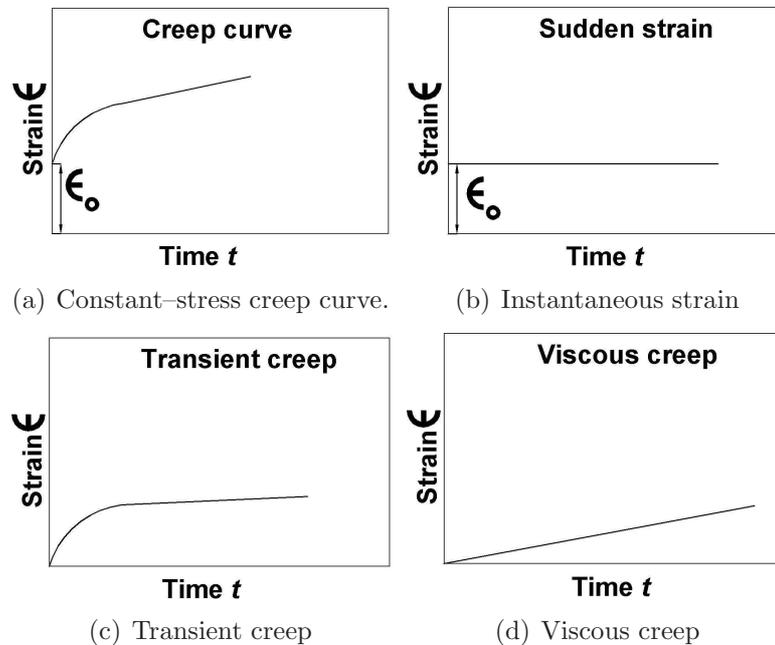
(c) Transient creep

(d) Viscous creep

Figure 2.11: The processes which determine the shape of the creep curve [78].

virtue of its own deformation. For low temperatures and stresses (*i.e.* lead at room temperature) primary creep is the dominant process. The second stage, *secondary creep*, occurs at a constant rate due to balance between strain hardening and recovery. The average value of this steady state creep rate is called the minimum creep rate. *Tertiary creep* occurs mostly in constant–load creep tests at stresses which are high in the context of the test temperature. It occurs when there is a reduction in the cross–sectional area either because of necking or internal void formation [75].

Chapter 3

TECHNIQUES

There are material properties that can be modelled simply and elegantly in an analytical way, such as crystallization kinetics or the tensile properties of composite materials. However, in some cases the property arises from many interactions, not all of which are well understood. Even if the mechanisms are known it may not be clear how they work together, and simplification of the problem may be unacceptable from an engineering point of view.

Some properties can be estimated on the basis of meaningfully compiled empirical models that have been fitted to experimental data. The most general regression method involves neural networks, which have proved to be incredibly powerful in the creation of new materials [79–84].

3.1   NEURAL NETWORKS

Neural networks consist of interconnected processing elements called nodes that work together to produce an output function. Processing

the information by the network is characteristically done in parallel rather than in series, since it relies on its member neurons collectively to perform its functions. Neural networks can be used to model complex relationships between inputs and outputs or to find patterns in the data.

3.1.1 Bayesian Neural Networks

An artificial neural network is a method for fitting a function to a number of points in the data space [85]. More technically, it is a parametrised non–linear model which can be used to perform regression in which a flexible non–linear function is fitted to experimental data.

Neural networks have been studied intensively [86–90] and for a long time [91]. The particular features of the method employed here, which is due to MacKay [92, 93], are described in the following paragraphs.

Assume that the material property of interest in this case is the ultimate tensile strength (UTS) of a particular steel, and that it can be expressed as a non linear function f, of a number of experimental available variables in a database:

$$UTS = f(c_i, t_t, T_t, t_a, T_a, T_T, ...) \qquad (3.1)$$

where c_i represents the chemical composition of the steel, t_t is the tempering time, T_t is the tempering temperature, t_a is the austenitising time, T_a is the austenitising temperature, T_T represents the test temperature and "..."represents all the other parameters that might influence the UTS.

The reason for using a Bayesian neural network [92, 93] is that it makes as few assumptions as possible about the form of the fitting function, whilst trying to infer and thus mimic its shape. The only

assumption it makes is that the function modelled is continuous and differentiable. The method also has an outstanding treatment of uncertainties. It has been shown that a three-layer neural network which has a sufficiently high degree of complexity can represent any non–periodic complex function [92]. The neural network is very flexible and can capture interactions between the parameters.
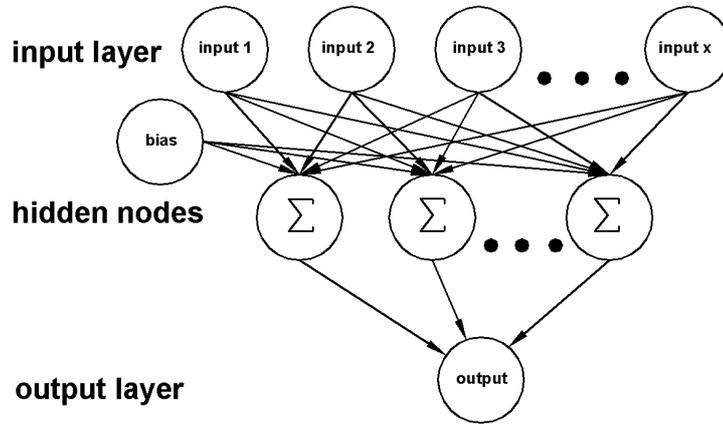

Figure 3.1: Structure of a three-layer neural network. The model complexity is controlled by the number of hidden nodes in the second layer.

The three-layer feed-forward neural network used in the present work and generally used for material property applications is shown schematically in Fig. 3.1. The first layer consists of the inputs to the network, the second of a number of non–linear operators (h – neurons) whose arguments are provided by the first layer in the network. Each consists of an activation function, that describes the output by a non-linear, continuous and differentiable function. In the present work *tanh* was used due to its flexibility and its invaluable property of being able to combine two or more functions to obtain highly flexible functions. The target to the output y, can be any function, but is commonly a linear one. The activation function for a node i is given by [92, 93]:

$$h_i = \tanh \left(\sum_j w_{ij}^{(1)} x_i + \theta_i^{(1)} \right) \tag{3.2}$$

and the output weighting function is:

$$y = \sum_i w_{ij}^{(2)} h_i + \theta^{(2)} \tag{3.3}$$

The x_i are inputs, and w_{ij} the weights which define the network. The superscripts $^{(1)}$ and $^{(2)}$ denote weights and biases in the hidden layer and in the output layer. The aim of training a network is to find the optimum values for w. The parameters θ are known as *biases*, and are treated internally as weights associated with a constant input set to unity.

The inputs are normalised within a range of $\pm$ 0.5:

$$x_j = \frac{x - x_{min}}{x_{max} - x_{min}} - 0.5 \tag{3.4}$$

where x is the un-normalised input, x_{min} and x_{max} are the minimum and maximum values in the database for that specific input, and x_j is the normalised value. The purpose of normalising is to be able to compare the sensitivity of the output to the inputs without biasing the comparison due to the differing magnitudes of the variety of inputs.

The complexity of the neural network model increases as the number of hidden units increases. The trained network is transparent because the inputs are known, the output is known and the weights can be examined, although they may be difficult to interpret directly because of the non–linearity of the model. The easiest way to find the interactions between inputs and output in a model is to use it to make predictions and visualise the behaviour which emerges from various combinations of inputs.

Overfitting

Due to the fact that neural networks are powerful and flexible, there is a danger of *overfitting* the model to an unjustified level of accuracy. Training a network involves finding a set of weights and biases which minimise an *objective function*, which balances complexity and accuracy, typically:

$$M(w) = \alpha E_w + \beta E_D \tag{3.5}$$

in which E_w is a *regulator*, designed to force the network to use small weights and limit the number of hidden units:

$$E_w = \frac{1}{2} \sum_{ij} w_{ij}^2 \tag{3.6}$$

and E_D is the overall error between target output values and network output values, given by:

$$E_D = \frac{1}{2} \sum_{k} (t^k - y^k)^2 \tag{3.7}$$

where t^k is the set of targets for the set of inputs x^k, while y^k is the set of corresponding network outputs. α and β are control parameters which influence the balance between a simple but inaccurate model, and an over–complex, also inaccurate model as shown in Fig. 3.2. MacKay's algorithm allows the inference of these parameters from the data, permitting control of the model complexity [92].

To accomplish training without overfitting, the data are randomly split into two parts a *training set* and a *test set*. The model is trained on the former set, and then its ability to generalise is compared against the test set of data. Fig. 3.3 shows how increasing the complexity continuously lowers the training error (the mismatch between model

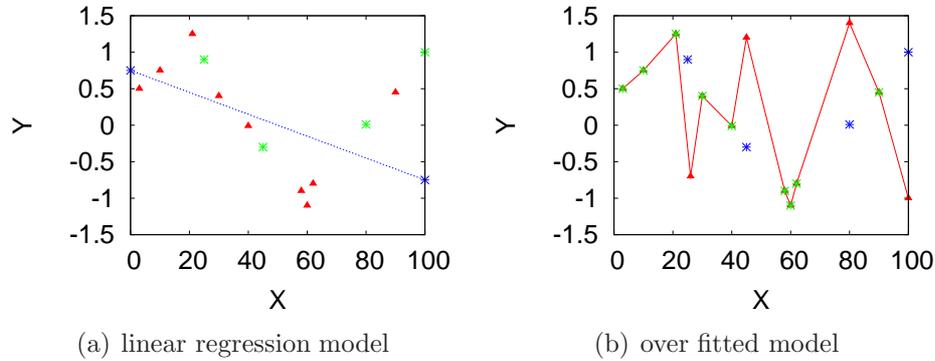(a) linear regression model   (b) over fitted model

Figure 3.2: Underfitting and overfitting. A set of data was split into a training (the filled triangles) and a testing (the stars) set based on the training set two models were trained. The first one (a) represents linear regression and it can be seen that it only gives a poor representation of both the training and the testing data. The second model (b) represents a complex function which gives an excellent representation of the training data but a poor one for the testing data.

predictions and the training dataset), while the test error (the mismatch between model predictions and the test dataset) goes through a minimum and increases again. As the model's complexity increases, overfitting causes the test error to increase as the number of hidden units increases. The aim of the training is to minimise the test error against the dataset and against new data not seen by the model.

For the present work, the fitting method is based on a Bayesian framework and treats training as an inference problem, allowing estimates to be made of the uncertainty of the model fit. Rather than trying to identify one best set of weights, the algorithm infers a probability distribution for the weights from the data presented. The performance of different models is best evaluated using the *log predictive error* (LPE) instead of the test error. This error penalises wild predictions to a lesser extent when they are accompanied by appropriately large error–bars and it is defined by:
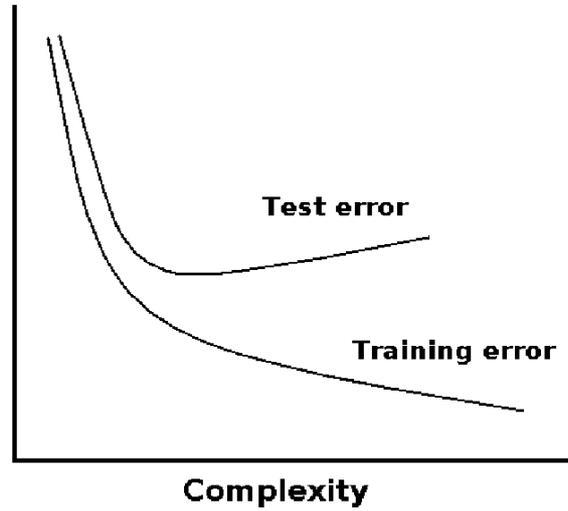
Figure 3.3: Comparison of error on training and testing sets as a function of network complexity, illustrating the problem of over complex models.

$$LPE = \frac{1}{2} \sum_k \left[\frac{\left(t^k - y^k\right)^2}{\left(\sigma_y^k\right)^2} + \log\left(2\pi\left(\sigma_y^k\right)^2\right) \right] \qquad (3.8)$$

where t and y were defined previously, and σ_y^t is related to the uncertainty of fitting for the set of inputs x^k.

In the training process a large number of models (for 20 hidden units and 5 seeds 100 models result after training) are created. These models perform differently when making predictions and are ranked according to the log predictive error. However more accurate predictions are obtained if instead of just using the best model a number of models is used to make predictions on unseen data. This is called a *committee* of models. The optimum number of models which will make up the committee is determined according to the combined test error of all the members of the committee. The prediction $\overline{y}$ of a committee of models is the mean prediction of its members, and the uncertainty is:

$$\sigma^2 = \frac{1}{A}\sum_l \left(\sigma_y^l\right)^2 + \frac{1}{L}\sum_l (y^l - \overline{y})^2 \qquad (3.9)$$

where A is the number of models in the committee and the exponent l refers to the model used to give the corresponding prediction y^l. During training it is usual to compare the performance of increasingly large committees on the testing set. Usually the error is minimised by using more than one model in the committee, and then the selected models are retrained on the entire database.

Significance

Another advantage of the method is an indicator of the network-perceived significance of each input, which is a function of the regularisation constants for the weights associated with an input σ_w. The regularisation is similar to a partial correlation coefficient in that it represents the amount of variation in the output that can be attributed by a particular input.

To determine the sensitivity of the model to individual input parameters, predictions must be made by varying one parameter only whilst keeping all the others constant. In some cases where an input is a function of one or more of the other inputs (for example, both temperature T and an Arrhenius function $\exp(\frac{1}{T})$) varying one of these parameters in isolation is not physically meaningful.

The network structure allows the assessment of input parameters based on physical models to be included in the training data, and those parameters which are not useful in explaining the output will have much lower significances than those that *are* useful as shown in Fig. 3.4 (the depth has a much bigger influence on the output than the hoop stress).
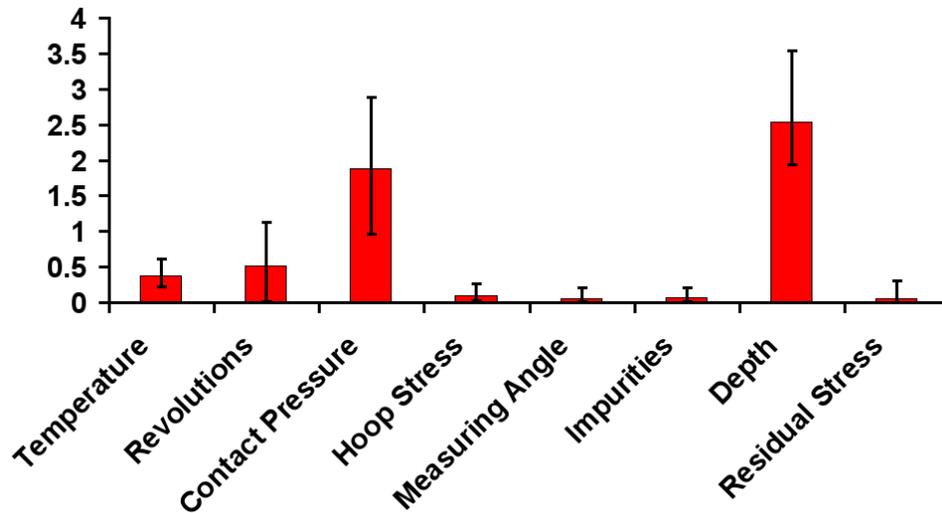
Figure 3.4: Example of a selected set of σ_w values from a committee. In this case the committee has 13 members, the error–bars represent the upper and lower limit from the members in the committee [94].

Uncertainty

When dealing with neural networks it is important to distinguish between two kinds of error. *Noise* is when the outcome of an experiment which is repeated a number of times results is slightly different results. This is because there may be variables which are not controlled. Suppose that a model has been created in such a way that overfitting has been avoided, the noise in the output can be assessed by comparing the predicted values of the output versus those measured. The noise is a constant measure and does not help much in assessing the generalization behaviour of the model.

The second kind of error is the uncertainty which originates in the ambiguity in the mathematical functions capable of representing the same experimental data. It is likely in complex problems that there may exist many mathematical functions which adequately represent

the same set of empirical data but which behave differently in extrapolation. This difference in behaviour is characterised by a *modelling uncertainty*. Fig. 3.5 illustrates how a number of different functions might adequately represent the same data, but which behave differently in regions where data are sparse or noisy.
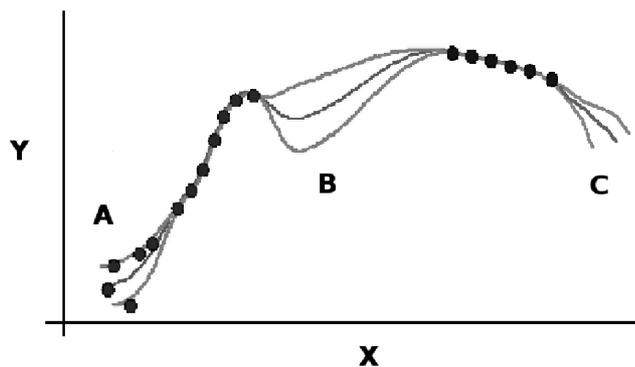

Figure 3.5: Schematic illustration of the uncertainty in defining a fitting function in regions where data are sparse (B) or noisy (A). In the (B) region the outer lines represent the error–bars due to uncertainties determining the weights. Outside the range of data the extrapolation is increasingly uncertain (C). The regions (B and C) will provide the most informative experiments [85].

The magnitude of the modelling uncertainty varies as a function of the position in the input space. It is an excellent way of identifying domains where further experiments are needed to reduce uncertainty or where the predictions should be used alongside an assessment based on other techniques, known trends or experience. The minimum noise level that the fitting procedure can achieve in the output can be fixed if it is known prior to the training process. Doing so helps to avoid overfitting.

Improving The Model

Artificial neural networks obviously perform best with good quality data, which are physically meaningful. It is also necessary to optimise the number of input variables, to ensure that they form a physically meaningful set. An over-ambitious set will however limit the quality data available for analysis since many reported experiments do not cover all the necessary input variables. A pragmatic approach ensures sufficient data and variables to capture the complexity of the problem.

Fig. 3.6a shows the model-perceived noise (σ) for a database with 16 input parameters and Fig. 3.6b shows the model-perceived noise (σ) for a database with 11 input parameters. The noise is smaller, for the 16 input parameter model because for the 11 input parameter model 5 important inputs were eliminated before training. The figure Fig. 3.6a represents the noise level from training a hot–strength model as a function of chemical composition, heat treatment, and test temperature whereas for Fig. 3.6b only chemical composition was chosen as input parameters.
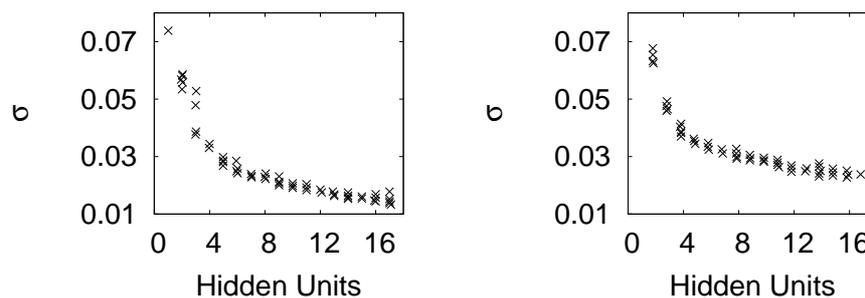
(a) noise level for 16 variable model   (b) noise level for 11 variable model

Figure 3.6: Perceived level of noise for the same problem but with different numbers of input parameters.

The maximum and minimum values for each input should be checked before commencing the modelling, to ensure that these limits are mean-

ingful (for example a maximum value for carbon at 88 wt% would obviously be a compiling error). Table 3.1 shows steel related data used for training a neural network.

Input	Units	Minimum	Maximum	Average	Deviation
C	wt%	0.04	0.48	0.1514	0.0627
Si	wt%	0.03	0.86	0.3624	0.143
Mn	wt%	0.35	1.73	0.6091	0.2402
S	wt%	0.005	0.029	0.0149	0.0056
P	wt%	0.001	0.1	0.0101	0.0085
Ni	wt%	0.0001	2.64	0.2216	0.4075
Cr	wt%	0.0001	12.9	3.603	4.114
Mo	wt%	0.005	2.03	0.6944	0.3855

Table 3.1: The minimum–maximum table.

A uniform spatial distribution of data minimises the uncertainties for interpolation and extrapolation. Fig. 3.7 illustrates how the uniformity with which the data are distributed may not be the same for all the inputs. In the case of nickel, it is likely that predictions may be associated with large uncertainties in the concentration range 1-2 wt%.

(a) Data distribution for copper
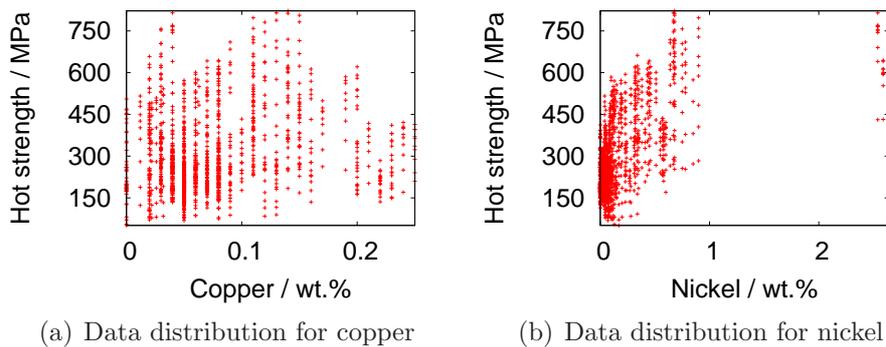
(b) Data distribution for nickel

Figure 3.7: Data distribution for different input parameters.

Experimental data should, prior to training be assessed for accuracy and reliability. Outliers such as the ones shown in Fig. 3.8, are those

points whose 95% confidence limits are so far away from expectation that they are unlikely to be right. Such data should be tracked back to their origin to see if there is a mistake in the process of collecting and compiling the database. If no errors are found then some reasoning has to be applied to ensure that the model construction is correct. A model that has a lot of outliers is likely to behave poorly on unseen data. The particular reason for the outlier illustrated in Fig. 3.8 is because there was a compiling error in the database.
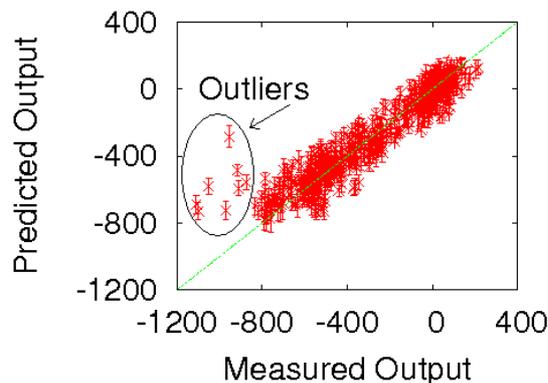
Figure 3.8: Example of outliers.

The input parameters can be inserted in the database in raw form, for example the chemical composition of a steel, or in a functional form, for example the free-energy of transformation calculated from the composition. Indeed both the raw data and the functional forms can be included as inputs to avoid biasing the model.

An input logarithm parameter may be included in the database if that parameter is linked to a time dependent process, for example the time period of heat-treatment is an input, and then along with the time a logarithm of time may be included $\log(time)$ as an input [95].

The logarithm of an input should also be used if it is desired to obtain better spread of the data. For example Fig. 3.9a shows the

distribution of residual stress developed in the raceway of a bearing as a function of revolutions; it can be seen that the data are clustered in two zones leaving a gap in the middle of the figure. If the logarithm of the revolutions is considered for the same data it can be seen Fig. 3.9b that the distribution is much more proportional resulting in smaller modelling uncertainties and a better deviation of the fitting function [2].
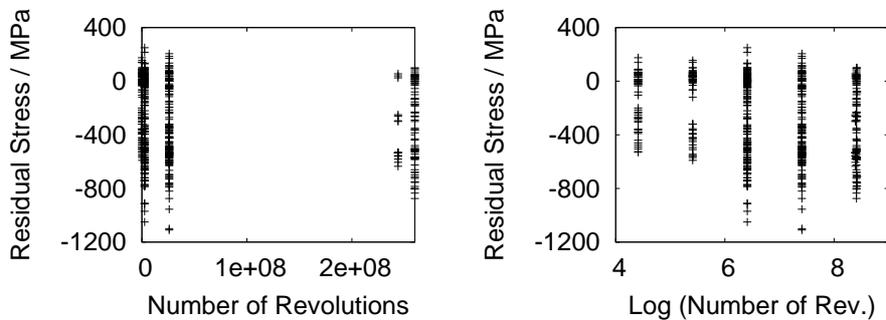
(a) Data in raw form, with an imbalance in the distribution
(b) Logarithm of the raw data resulting in a proportional distribution

Figure 3.9: The difference in data distribution if the logarithm of an input is considered.

For thermally activated processes, temperature is expected from theory to act via the function:

$$\exp\left(-\frac{Q}{RT}\right) \tag{3.10}$$

where R is the universal gas constant, T is the temperature in Kelvin and Q is an activation energy. It was used by Shen [96] when modelling the rate of creep deformation because it builds in a dependence of the creep deformation rate on the activation energy. It could be argued therefore that it incorporates a physical relationship based on scientific understanding.

If a combination of two or more existing variables has a particular significance, it can be added to the database. For example, the following

could be added as an input from kinetic theory, a function of time and temperature (called 'kinetic time' [97]):

$$time \times \exp\left(-\frac{Q}{RT}\right) \qquad (3.11)$$

A neural network, like any empirical method, is a mathematical function and is not constrained to a particular range of output values. As such, it can predict unphysical values for the output, for example, the volume fraction of bainite can only take values between zero and one, in order to avoid the output taking values outside this range, a function is imposed on the output [2]:

$$f_{output} = \ln\left(-\ln\left(1 - \frac{z_{max} - z}{z_{max} - z_{min}}\right)\right) \qquad (3.12)$$

where z is the output and z_{max} and z_{min} are values set by the user. In the case of volume fraction z_{max} has the value of unity and z_{min} is zero.

3.2 GENETIC PROGRAMMING

In genetic programming, problems are viewed as the discovery of computer programs through an evolutionary search process [98, 99]. A computer program can be a formula. Brajlih *et al.* used genetic programming to acquire a formula which expresses a material shrinking coefficient [100]. It can be a plan; Shimizu *et al.* implemented a decision making plan based on genetic programming, which decides the optimal disassembly sequence in a recycling plant [101]. It may be a control strategy; Beldek and Leblebicioglu used genetic programming to create a control strategy which navigates particles in a area filled with obstacles [102]. Another alternative is a decision tree; Estrada *et al.* have successfully implemented a genetic programming method

which induces a decision tree which is able to detect interactions in genetic mutated viruses [103]. Whigham and Crapper used a learning model for rainfall-runoff based on genetic programming [104].

The learning problems can be divided into two kinds. The first one uses a dataset such as input/output mappings for supervised learning the information source provides examples of the inputs and the correct outputs during the training period, the supervised learning was used successfully in image classification [105]. The second deals with real or simulated robotic behaviour in a specific environment for example robots whose behaviour is determined by a robotic controller. The controller programs are automatically generated with the help of genetic programming [106].

A population of computer programs is randomly generated, each of which is then assessed by a fitness function which quantifies the ability to solve a given problem. To make possible the application of genetic programming it is necessary to define a fitness function, for that problem the definition of a quantitative measure of the performance of a possible solution. Fitter programs are selected for recombination to produce a new population of programs using genetic operators such as crossover and mutation [107]. This step is iterated for a number of generations until, the optimum solution has been reached, or an economical computational effort has been exhausted. The cycle is illustrated in Fig. 3.10.

As already explained, the method relies on a population of candidate solutions or programs generated at random. The population size is specified by the user and depends on the difficulty of the problem. A large population is needed for complex phenomena to maintain diversity and to avoid premature convergence [98, 108].
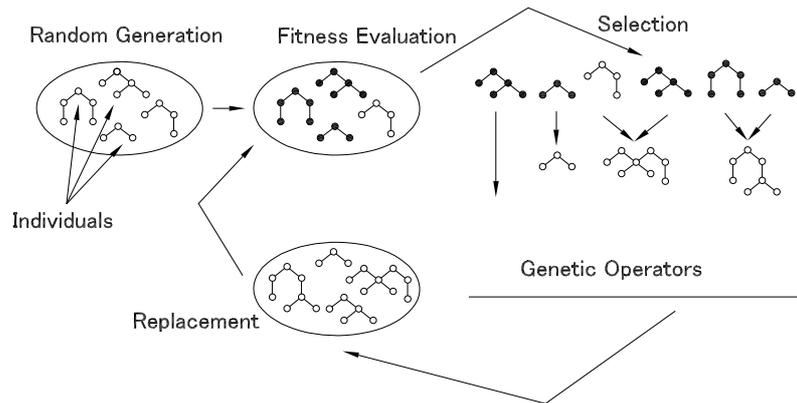
Figure 3.10: The evolutionary cycle in genetic programming.

Representation

The programs can be represented as abstract syntax trees, where a branch node is an element from a function set which may contain arithmetic and, or logic operators, or elementary functions with at least one argument. A leaf node of the tree is an element from a terminal set, which contains variables, constants and functions with no arguments. There are other representations such as linear genetic programming which is the evolution of computer programs written as linear sequences of instructions resulting in an acceleration of execution time and evolutionary progress. This was used in the selection and purchase of market shares [109] and in determining the optimal spatial distribution of agricultural ammonia emissions in order to minimize atmospheric nitrogen deposition in nature reserves [110].

Grammatical evolution is a different type of genetic programming used to provide alternatives to the syntax tree. This is done by ob-

taining the solution according to a user-specified grammar resulting in a restricted search space and faster times for solving the problem with the domain knowledge of the problem being incorporated into the solution. This has been used to optimize neural networks for use in gene detection studies [111].

The function and terminal set are provided by the user, who specifies in advance blocks that can be used for constructing a solution to the problem. The identification of the elements that will be included in the function and terminal sets should be chosen so that the functions are sufficiently complex to express the solution [112].

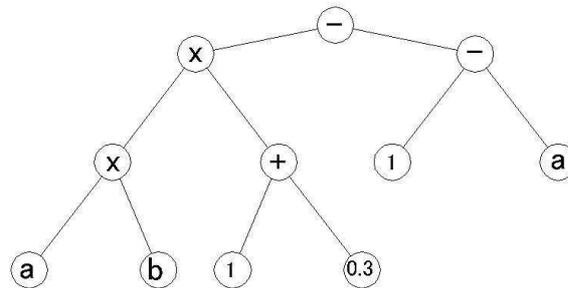Fig. 3.11 represents an example of abstract syntax tree for a mathematical expression.

Figure 3.11: A tree representation of $a \times b \times (1 + 0.3) - (1 - a)$.

Fitness Function

This provides a measure of fit and is a criterion that helps select the optimum function. Fitness is usually evaluated over a set of cases. The fitness functions most commonly used are the error variance (mean square error) and the error standard deviation (root mean square error) [113, 114]:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - t)^2 \qquad (3.13)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\frac{y_i - t}{t}\right)^2} \qquad (3.14)$$

where y_i is the predicted value and t is the target value.

Selection

The objective of the selection is to apply the survival of the fittest mechanism on candidate solutions:

- roulette-wheel selection where the probability that an individual will be selected depends on his normalised fitness value [115];

- ranking selection which is based on the ranking of the fitness values of the individuals in the population [116];

- tournament selection where n individuals with $n \geq 2$ are chosen in the population, and the individual with the highest fitness value is selected [117];

- elitism, which copies the best few individuals in the next generation, can increase performance and avoids losing the best known solutions [118].

Genetic Operators

Genetic operators perturb the structure of the trees by combining or modifying two or more parental individuals to create possibly better solutions.

The simplest operator in genetic programming is duplication since it only replicates an individual without any modification.

Crossover is an operation for the probabilistic exchange of genetic information between two randomly picked parents, facilitating a global search for the optimum in the process. The crossover operator is schematically represented in Fig. 3.12.
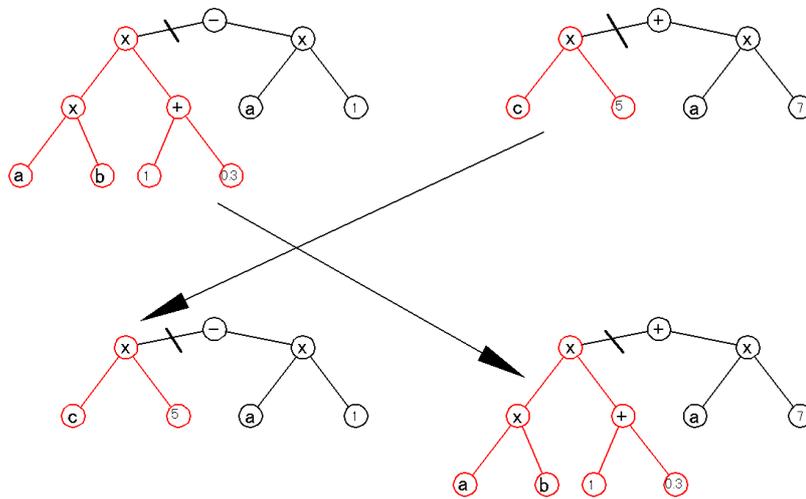

Figure 3.12: A sub–tree crossover.

Mutation is an operator inducing a small probabilistic change in the genetic make up, resulting in a local search. It begins with the selection of a point at random within the tree, and then the mutation operation replaces the function or the terminal set with the same type of element. This operation is useful for creating diversity without modifying the shape of the tree, presented in Fig. 3.13.

There are many variations of mutation and crossover available in the literature [119, 120]. It is believed that the crossover operator allows a quicker convergence than mutation but the subject is open for debate [121–124].
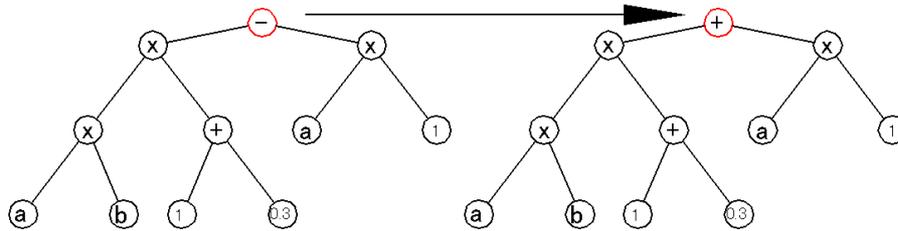
Figure 3.13: A one–point mutation.

3.3   THE PURPOSE

The purpose of any neural network and genetic programming model is to enable the quantitative expression and understanding of complicated problems.

It is usual to publish research in the form of papers which describe the work in sufficient detail to enable it to be reproduced. Research involving neural network/genetic programming models is similarly published but the work often cannot be independently reproduced from the publication alone, because, neither the data on which the model was created nor the coefficients needed to reproduce the mathematical structure can be incorporated in the paper.

Modern technology in the form of the world–wide web provides an elegant solution in that the numerical model itself or the data can be archived and made available for scrutiny, as a supplement to the normal research paper. The dissemination of the model has the added advantage that it will be exposed to a much wider audience with the potential of applications not originally envisaged by the creators of the model.

In Table 3.2 a marking scheme is presented, which has the purpose to encourage a useful guide to publication in this field [125]. The first

achievement is of course to do the work necessary to assemble the data and to create a model. Such a model must be considered potentially unreliable unless its behaviour and predictions are explored, both within the domain of the training data and beyond. The term 'prediction' in this context means conducting calculations in a regime where data do not exist.

Characteristic	Mark
Model or data disseminated	3
Modelling uncertainty	2
Prediction investigated experimentally	2
Predictions made and interpreted	2
Model created	1

Table 3.2: Marking scheme for models. *Prediction* refers to using the model in a regime of inputs not included in the data used to create the model. A *model* refers to a transparent mathematical function and associated coefficients. *Dissemination* implies making the model available to others.

The trends revealed by the model should be interpreted in terms of the known science of the problem. It is possible that a model captures false correlations and these can render it unsafe in application.

Some of the predictions of the network may be usefully investigated by conducting specific experiments.

It is important to realise that the extrapolation of a network cannot be satisfactorily explored without an indication of modelling uncertainty (discussed in Chapter 3 page 36). In a non–linear model extrapolation cannot be defined as being out of the range of the training data unless the latter are absolutely uniformly distributed in the input space. Modelling uncertainty, however it is calculated, must be considered an essential feature of any model.

Finally, the dissemination of the model or data is vital for scientific

progress and independent verification.

Chapter 4

HOT STRENGTH OF STEELS

The strength of a material refers to the material's ability to resist an applied force without undergoing permanent deformation. There were many attempts to model the strength of steel, in the following paragraphs mathematical models used to estimate the strength are described together with their drawbacks. In the second part of the chapter a more capable model is presented.

4.1 MATHEMATICAL MODELS FOR STRENGTH

There are many researchers who tried and succeeded in creating empirical models to express the strength of steels. Classic equations due to Irvine *et al.* quantify the strength in terms of alloying element contributions, grain size and δ–ferrite content [126].

$$
\begin{aligned}
0.2\% \quad \text{proof} \quad \text{stress} &= 4.1 + 23\text{C} + 1.3\text{Si} + 0.24\text{Cr} + \\
&+ 0.94\text{Mo} + 1.2\text{V} + 0.29\text{W} + 2.6\text{Nb} + 1.7\text{Ti} + \\
&+ 0.28\text{Al} + 32\text{N} + 0.16(\delta - \text{ferrite}) + 0.46d^{-\frac{1}{2}}
\end{aligned}
\tag{4.1}
$$

where d represents the average grain size which in the investigation[1] ranged from 2 to 15 mm$^{-\frac{1}{2}}$, δ–ferrite represents the volume % in the range 0 – 60%, the C, Si, *etc* represent the contribution in wt% of the respective components to strength; resulting in a 0.2% proof stress between 120 and 400 MPa.

Tomita and Okabayashi working on steels with mixed microstructures [2] (martensite and bainite) found that a mixture can perform better than the individual phases, with strength peaking as a function of volume fraction because the properties are affected by the size [127]:

$$
\sigma_{0.2}^{Mix} = \sigma_i + kS_{LM}^{-\frac{1}{2}} - (\sigma_i + kS_{LM}^{-\frac{1}{2}} - \sigma_{0.2}^{B})V_B
\tag{4.2}
$$

where $\sigma_{0.2}^{Mix}$ is the 0.2% proof stress (MPa) of the mixed microstructure steel, σ_i is the frictional stress, k is a constant, S_{LM} represents the martensite size partitioned by lower bainite, S_{UM} represents the martensite size narrowed by upper bainite, $\sigma_{0.2}^{B}$ is the 0.2% proof stress of bainite in the mixed structure and V_B represents the volume fraction of bainite.

A more general model which takes account of the number of intrinsic components was provided by Young and Bhadeshia [128]. The maxi-

---

[1]The base steel used for the investigation had the chemical composition C 0.091, Mn 1.06, Si 0.42, Ni 10.03, Cr 18.04, N 0.0105 wt% to which different alloying elements were added.

[2]The steel used for the investigation had the chemical composition C 0.4, Si 0.15, Mn 0.66, P 0.011, S 0.009, Ni 1.87, Cr 0.8, Mo 0.2 wt%, with heat treatments resulting in yield strengths of 800 to 1800 MPa.

mum in the strength of mixed microstructures is said to be due to the increase in the strength of martensite as carbon is partitioned into the residual austenite during the formation of bainite, and the plastic constraint in which the strength of bainite is enhanced by the surrounding more rigid martensite.

These models and others are valid in expressing the yield strength, but they usually are limited to a particular type of steel, they require extensive microstructural data and they do not extend beyond the range of the strength which was used in their development. All the empirical models assume a linear dependence between the various parameters and the strength, but there is no general justification for this approach. In order to remedy the situation it is necessary to address the complexity issue, and hence the resort to Bayesian neural networks. These are flexible mathematical functions which are not restricted to the type of data available. Because of this they can cover a huge range of parameters and have been proven to reveal new phenomena that have led to novel steels [79–84, 129].

4.2 NEURAL NETWORK MODEL

The Bayesian neural network method used to express the hot–strength of steels as a function of chemical composition and heat treatment, was described in Chapter 3. The method has been used previously with success for modelling the Charpy energy, weld cooling rate, strength of steel weld metals, the yield strength and ultimate tensile strength of nickel base superalloys, the behaviour of high temperature creep–resistant steels [87, 130–135]. A review of these applications has been given by Bhadeshia [85].

There have been studies in which the hot–strength of austenite has

been modelled, primarily as an aid to the simulation of the hot–rolling processes or that of the bending of steel during the continuous casting process [136–140]. There do not appear to be similar studies for creep–resistant ferritic steels, even though hot–strength is a parameter in the design of power plant components.

4.2.1 The Variables

There exist a lot of data in the literature on the strength of ferritic steels. The present analysis is based on published data on the hot–strength (0.2% proof strength) of ferritic creep–resistant steels including but not limited to the classical $2\frac{1}{4}$Cr1Mo, 5Cr, 9Cr1Mo and 12Cr1Mo type steels [141].

Fig. 4.1 illustrates the range of the input variables used to develop the model, plotted against the output – hot–strength. Fig. 4.1 does not have the role of showing functional dependencies; it is intended just to show the distribution of the data without correlation between the different variables. The data are not distributed uniformly, but the Bayesian framework of the neural network recognises this by associating large modelling uncertainties with the sparse or noisy domains.

The input parameters used to develop the hot–strength model are listed in Table 4.1 and include most of the parameters thought to influence the hot–strength: chemical composition, heat treatment, and test temperature. The microstructure is not included as a separate parameter because it is a consequence of heat treatment and chemical composition. The Bayesian framework of the neural network enables the assessment of the relevance of each individual input, so it is unnecessary to exclude any variable prior to the analysis because the variables which hold little influence on the output will be associated with small weights [142, 143].
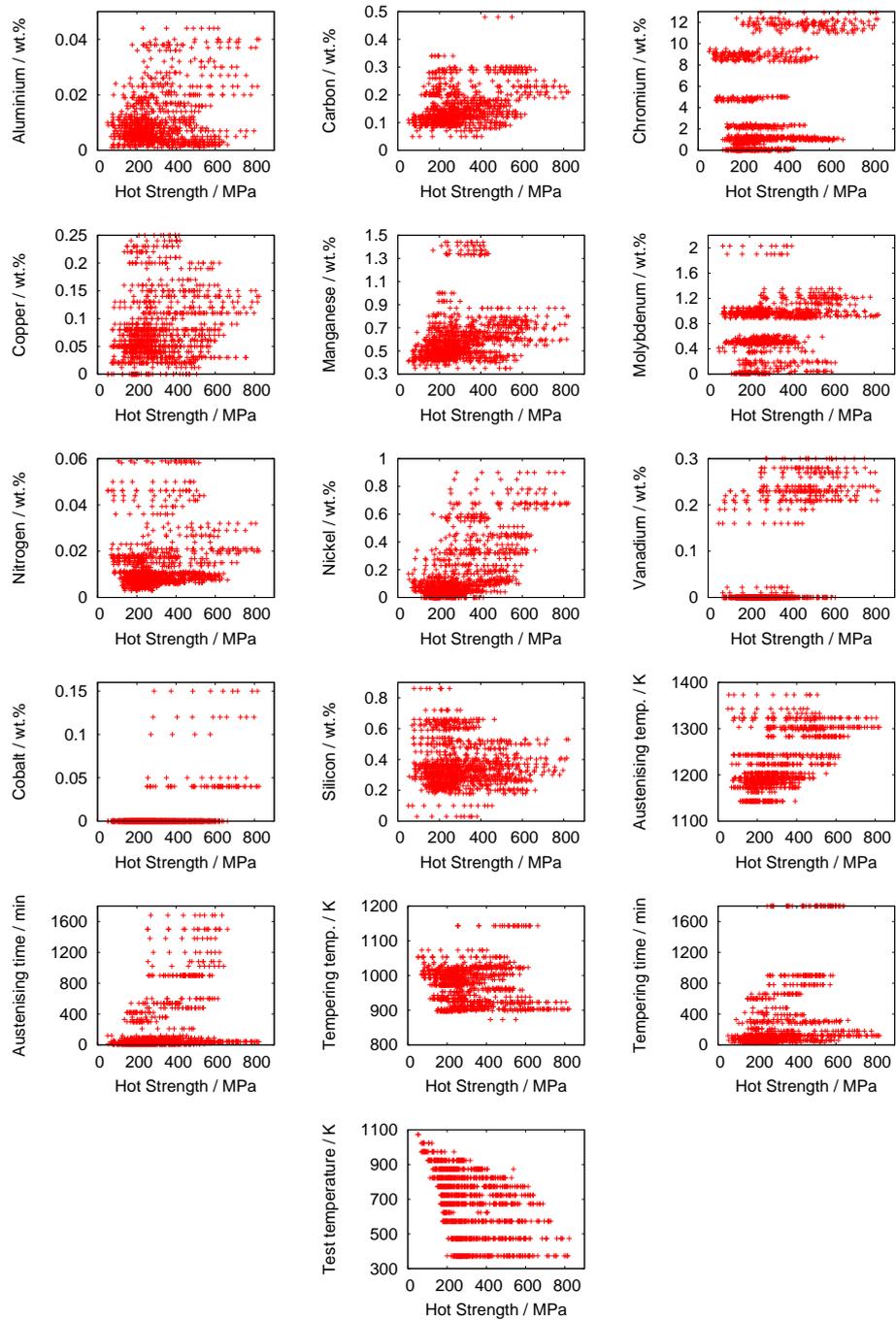
Figure 4.1: Visual illustration of the distribution of data used to create the model.

Variable	Minimum	Maximum
Aluminium / wt%	0.001	0.04
Carbon / wt%	0.09	0.48
Copper / wt%	0.0001	0.25
Chromium / wt%	0.0001	12.38
Manganese / wt%	0.38	1.44
Molybdenum / wt%	0.01	1.05
Nickel / wt%	0.0001	0.6
Nitrogen / wt%	0.001	0.04
Vanadium / wt%	0.01	0.3
Cobalt / wt%	0	0.15
Silicon / wt%	0.18	0.86
Austenitising time / min	10	5400
Tempering time / min	30	660
Austenitising temperature / K	1143.15	1243.15
Tempering temperature / K	898.15	1023.15
Test temperature / K	293.15	973.15
Hot strength / MPa	69	660

Table 4.1: The variables used to develop the model.

The hot–strength was modelled directly as a raw value, rather than as a functional form which would bound its values. This is because a bounding function leads to bias [144]. This is particularly the case when training the model with logarithmic values, which induces an upper and a lower limit for the predicted values and hence biases the model, this becomes extremely obvious when extrapolating over long ranges. To avoid this, the model was directly trained on the hot–strength data.

4.2.2 Training the Model

All the individual models can be ranked according to the test error or the log predictive error. The advantage of ranking them according to the log predictive error has been discussed in the previous chapter where was shown that combining a number of models to obtain a committee is a better procedure. Throughout this thesis the committee of X best models ranked according to log predictive error is used. The reason why the combined test error is presented for the committee of models is simply because it is easier to illustrate that the error is reduced when the committee is used.

The data were split into the training and testing sets, which were normalized as described in the Chapter 3. One hundred networks were trained, with hidden units ranging from one to twenty and five seeds in each case. As the number of hidden units increases so does the complexity and flexibility of the network, so the expected noise level from the training data decreased, Fig. 4.2a.

As can be seen in Fig. 4.2b, the ability of the models to generalise on the test data has an optimum at about ten hidden units, as does the log predictive error in Fig. 4.2c. The optimum number of models to form a committee was found to be thirteen as shown in Fig. 4.2d. The network perceived significances are shown in Fig. 4.3. Both the mean significance, and the upper and lower limits from the members of the committee are shown.

Fig. 4.3 shows that there is no one chemical component that has an overwhelming influence on the hot–strength: however carbon, nickel, chromium and molybdenum are most significant. This is expected as they are carbide formers, but the reason for nickel is not clear. A possible explanation for the low significance associated with the other elements could be their compositional narrow range. The range is illus-
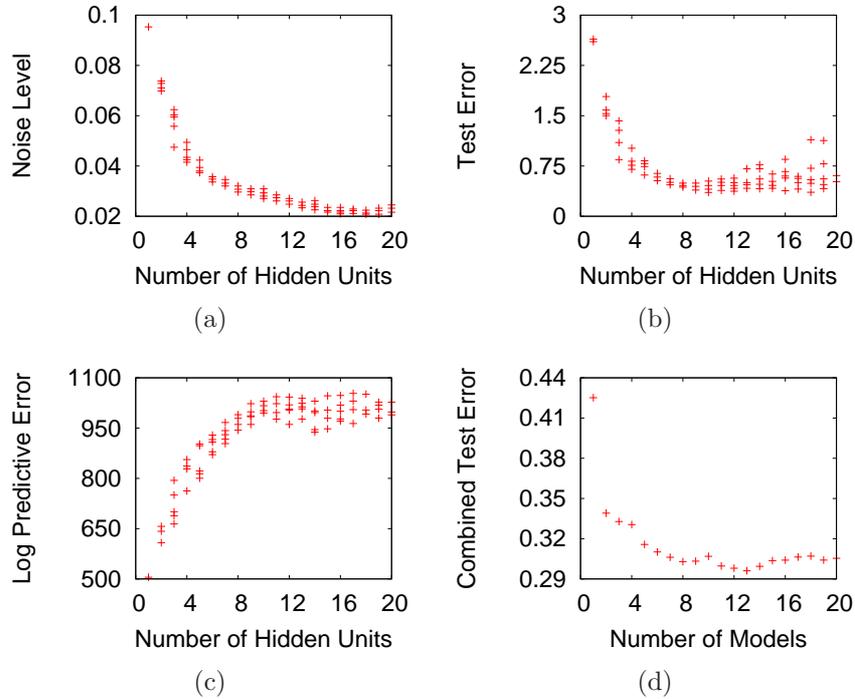
Figure 4.2: (a) Perceived level of noise for the training data (b) the test error (c) the log predictive error (d) the test error for the committees of different sizes.

trated in Fig. 4.1. The tempering heat treatment naturally has a large effect on the strength as does the test temperature.

The final trained committee tested on the whole dataset shows reasonable agreement between the predicted and measured values; even those few data which are badly estimated are accompanied by larger uncertainties as shown in Fig. 4.4. There is clear improvement over the predictions shown in Fig. 4.5a and b by the best single model.
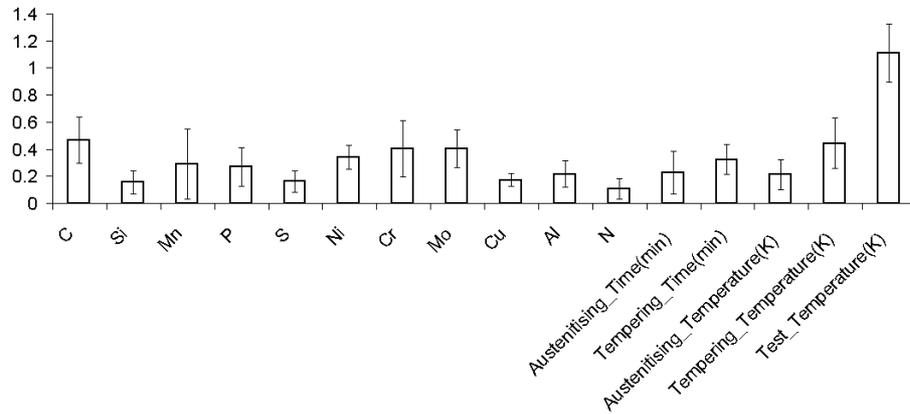
Figure 4.3: Model perceived significance for the committee of the models.
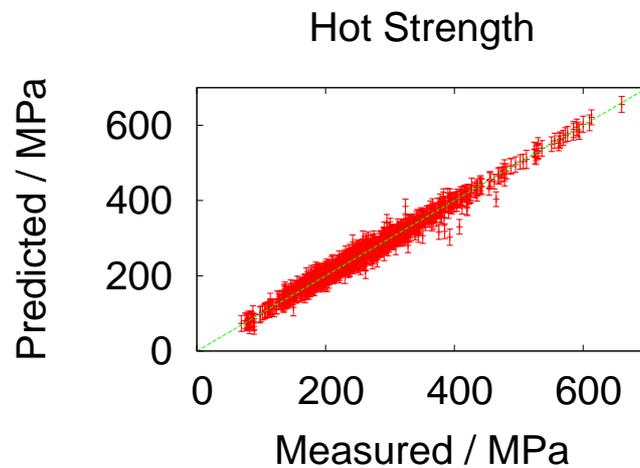
Figure 4.4: The performance of the committee of models on the whole database. For the committee model the standard deviation for predicted versus measured is 34 MPa, and there were only 28 points more than three standard deviations away from their measured values out of 1107 data points.

4.2.3 Interpretation of Hot–Strength

For all the predictions in the present thesis resulting from neural network models the modelling uncertainties are $\pm 1\,\sigma$.
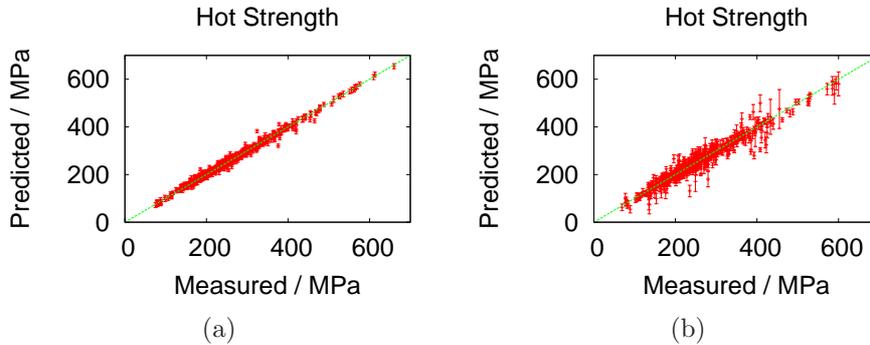
Figure 4.5: (a) The performance of the best single model on the training data, (b) the performance of the best single model on the testing data.

In order to assess the hot–strength model, three examples where considered, using steels in the $2\frac{1}{4}$Cr1Mo, 5Cr and 9Cr1Mo categories. In each case, phase diagrams were calculated using MTDATA [145] and the 'solution plus' thermodynamic database. The phases allowed in the calculation were cementite, M_3C_2, M_7C_3, $M_{23}C_6$, M_6C ('M'stands for metal atoms) and ferrite, including Fe, C, Si, Mn, Ni, Cr, Mo, Cu, Al, N as the components. The composition used to estimate the solid solution strengthening of ferrite was that calculated for the tempering temperature, since this is higher than the tensile test temperature and because the test itself is of a short duration[3].

With the calculated concentrations of substitutional solutes in the ferrite, their contributions were estimated as a function of test temperature using a model due to Sugden [146]. The small concentrations of interstitials in equilibrium with carbides were neglected since they are likely to be located at defects [147–150], and hence do not contribute to solid solution effects; at the tempering temperature the approximate amount of interstitials for the steels was C 0.009 and N 0.008 wt%.

---

[3]The hot–tensile tests are reported to have been carried out to Japanese standard JIS G 0567 – after reaching the specified temperature, the sample is held there for 15 min before testing.

The strength of pure iron with a coarse microstructure, as a function of temperature, is from Leslie [151]. The microstructural contribution from carbides and the tempered martensite or bainite plates, is then the difference between the neural network estimate and the SS+Fe curve, where SS stands for solid solution strengthening.

The hot–strength predictions for the three steels are presented in Figs. 4.6a,c,e. It is seen that hot–strength can be categorised into two regimes. The first is an almost linear decrease in strength with temperature approximately over the range 200–700 K. This occurs at a rate consistent with the loss of strengthening due to substitutional solutes and iron (*cf.* SS+Fe curve). The second part of the hot–strength is also approximately linear over the range 800–950 K, but the decrease in strength with temperature is much more dramatic. By extrapolating the low and high–temperature behaviour, T_C, which is the transition temperature between the two regimes is found to be 793 K, 845 K and 780 K for the $2\frac{1}{4}$Cr1Mo, 5Cr and 9Cr1Mo steels respectively.

The accelerated decrease in hot–strength when $T > T_C$ cannot be attributed to coarsening phenomena or microstructural changes, because in all the cases illustrated, the samples have been tempered at temperatures in excess of 990 K, which is much higher than the tensile test temperatures. Neither can it be associated with any similar behaviour in the solute strengthening or the strength of pure iron, both of which are almost monotonic straight lines.

The phase diagrams (Figs. 4.6b,d,f) show that there is no dramatic or consistent change in equilibrium phase fractions at T_C. The remaining possibility is that it becomes easier for dislocations to overcome obstacles by a thermally activated mechanism beyond T_C within the time scales of the experiments.
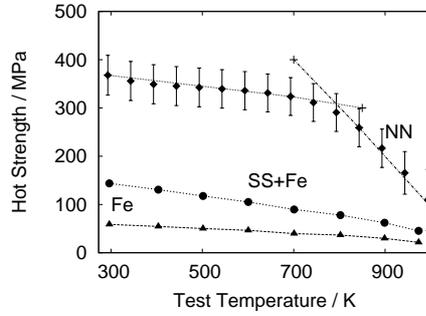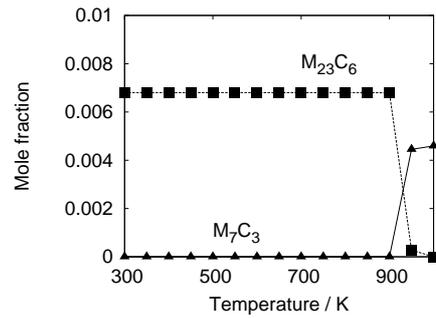
(a) $2\frac{1}{4}$Cr steel: 0.15C-0.18Si-0.63Mn-0.024Ni-2.23Cr-0.97Mo-0.2Cu-0.01Al-0.0083N wt%, 1193 K for 480 min, air cooled, tempered at 993 K for 360 min
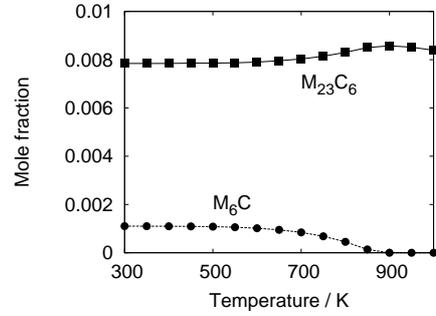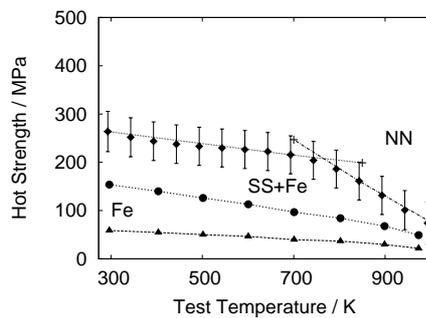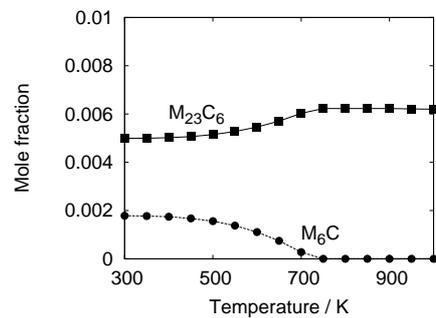
(b) $2\frac{1}{4}$Cr steel
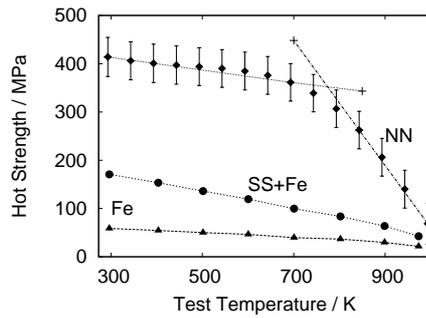
(c) 5Cr steel: 0.12C-0.33Si-0.56Mn-0.046Ni-5Cr-0.049Mo-0.05Cu-0.066Al-0.017N wt%, 1173 K for 10 min, air cooled, tempered 1023 K for 120 min

(d) 5Cr steel

(e) 9Cr1Mo steel: 0.11C-0.59Si-0.41Mn-0.1Ni-9.15Cr-1.05Mo-0.02Cu-0.011Al-0.018N wt%, 1133 K for 30 min, air cooled, tempered 1033 K for 90 min

(f) 9Cr1Mo steel

Figure 4.6: (a,c,e) Hot strength; (b,d,f) equilibrium phase fractions.

Relevance to Creep–Rupture Data

The microstructural component of strength is maintained below T_C but decreases dramatically about that temperature. Since this happens in a short duration hot–tensile tests, there should exist a similar phenomenon in creep rupture testing.
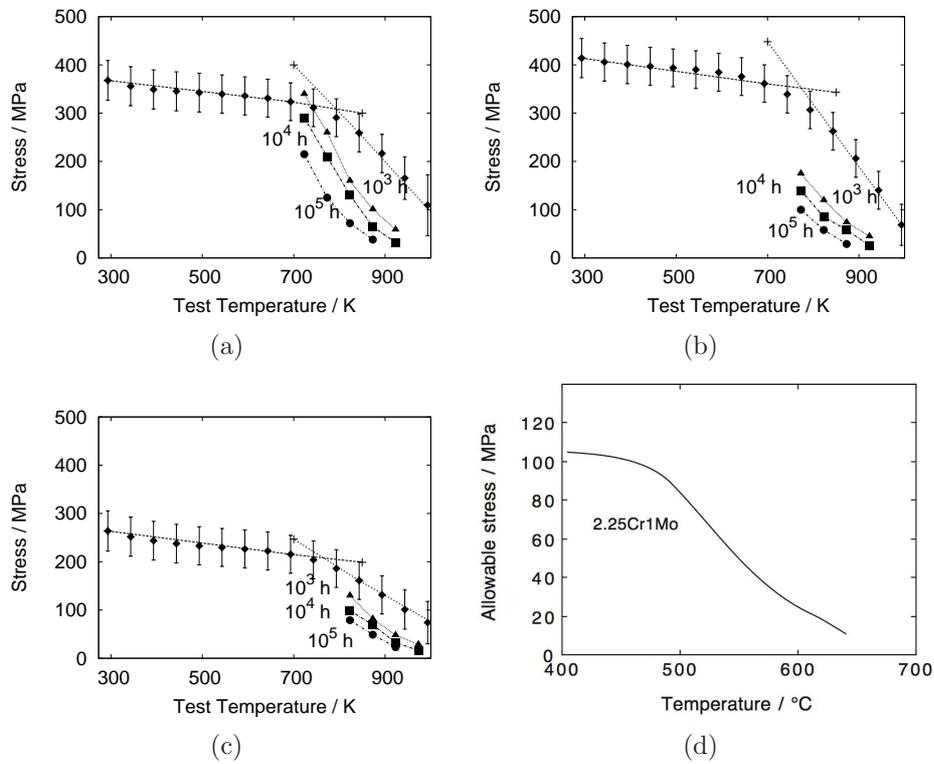

Figure 4.7: Comparison of temperature–sensitivity of creep–rupture and proof strength. (a) $2\frac{1}{4}$Cr1Mo. (b) 5Cr. (c) 9Cr1Mo. (d) The allowable stress for $2\frac{1}{4}$Cr1Mo steel [152].

Figs. 4.7a–c show creep rupture data (10^3, 10^4, 10^5 h) [141] plotted on the same graphs as the hot–strength data. It is striking that the temperature sensitivity of the rupture stress is similar to that of the proof strength for $T > T_C$. Unfortunately, low–temperature creep data are not available to make a similar comparison for $T < T_C$, but

Fig. 4.7d shows that the allowable stress in creep design varies in a manner strikingly similar to the behaviour of hot–strength as a function of temperature.

This analysis suggests that hot–strength tests could in research programmes be used as rough indicators of the temperature sensitivity of creep rupture data. This may not be too far fetched if for $T > T_C$, the mechanism remains thermally activated dislocation motion for both tensile and creep deformation.

4.3   GENETIC PROGRAMMING MODEL

The genetic programming method, used to create the models for predicting the hot–strength as a function of chemical composition and heat treatment, was described in Chapter 3. The method has been used previously in steel research for predicting the bending capability of rolled metal sheet and modelling the yield stress of carbon steel [153, 154].

In the previous section it was shown how the rate at which the strength (σ_Y) decreased with temperature (T) becomes much more pronounced beyond $T_C \simeq 800\,\mathrm{K}$, as shown in Fig. 4.7a–c. Furthermore, the slope $\partial\sigma_Y/\partial T$ for $T > T_C$ was found to be similar for hot–strength and for creep rupture data, as illustrated in Fig. 4.7a–c. This behaviour is thought to be associated with the onset of atomic mobility, such that dislocation climb becomes possible at temperatures beyond T_C.

Another useful outcome is that $|\partial\sigma_Y/\partial T|$ in the high temperature regime correlates directly with the tensile strength measured at ambient temperature, Fig. 4.8. An alloy which is strong at ambient temperature

loses more of its strength as the temperature is raised above T_C. Fig. 4.8 could in principle be used to determine $|\partial \sigma_Y / \partial T|$ for an arbitrary steel, and hence to work out the temperature–dependent strength simply by measuring the ambient temperature strength:

$$\sigma_{Y_T} = \sigma_{Y_{T_A}} + \frac{\partial \sigma_Y}{\partial T} \times (T - T_A) \qquad (4.3)$$

where T_A is the ambient temperature.

Note that for $T < T_C$, $|\partial \sigma_Y / \partial T|$ seems independent of alloy (Fig. 4.8) and corresponds to the temperature sensitivity of the solid–solution strength and that of pure iron [129].
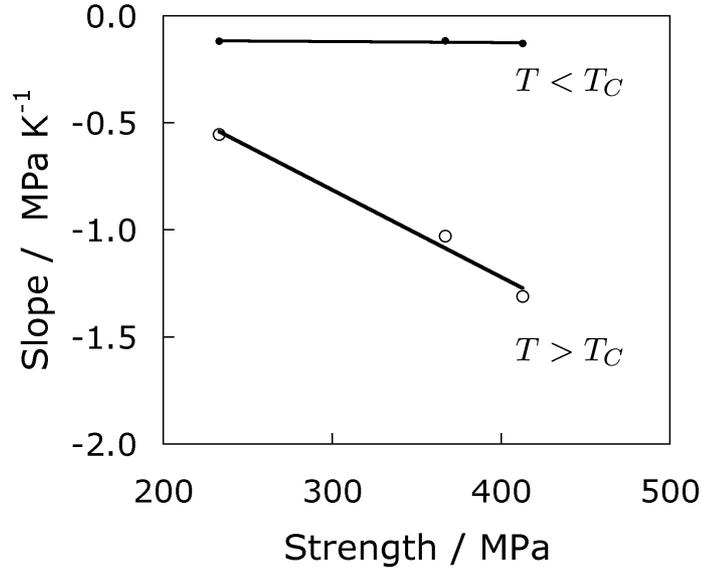

Figure 4.8: Plot of $\partial \sigma_Y / \partial T$ as a function of σ_Y measured at ambient temperature. Two cases are illustrated, for $T > T_C$ and vice versa. The alloys considered are the $2\frac{1}{4}$Cr1Mo, 5Cr and 9Cr1Mo.

Given the utility of the phenomena described above, the scope of the following section is to asses the possibility of replicating the neural network results presented in the earlier section, by an alternative

empirical method. Therefore the same data were used for the genetic programming model.

The specific training parameters for the genetic programming models are as follows: Koza crossover, a crossover rate of 0.78, a duplication rate of 0.04, point node type mutation at a rate of 0.18, tournament selection, with an initial population size of 1000 individuals, resulting in a correlation coefficient of 0.91 (if it reaches 1 it follows that the ideal solution is reached).

4.3.1 Comparison of Methods

In making these comparisons the calculations are based on the same data used for creating the neural network model. This is because the neural network methods have been shown to faithfully reproduce those data [129] – the genetic programming should therefore at least replicate those outcomes or reveal new structure within the generally verified trends.

The experimental data were in each case split randomly into two parts to create the training and test datasets, the latter being used to assess the ability of the model to predict unseen data; the test error is calculated as follows:

$$E_E = \sum_j (t_j - y_j)^2 \tag{4.4}$$

where y_j is a predicted value and t_j the target value; to calculate this error we normalised the output to be in the range ± 0.5 .

The neural network model used for the present thesis is in a Bayesian framework and hence gives a modelling uncertainty associated with each prediction [85, 142, 143, 155]. This is extremely helpful in avoiding

the usual dangers of extrapolation. For the genetic programming, the modelling uncertainty was improvised by plotting the standard deviation of three models within a committee of models. The three component models were created simply by starting the training process three times; on each occasion a different model is obtained because of the randomness–based components of the genetic programming process.

The training process for the genetic programming begins by applying only the elementary $(+ \; - \; * \; /)$, and if necessary after assessment, more complex models can be created by including other functions such as (sin, cos, tan, tanh, log, exp...). In what follows, *all functions* refer to the case where the elementary, trigonometric and arithmetic functions have all been made available to the genetic program, whereas *trigonometric functions* only trigonometric functions (sin, cos, tan, tanh, cosh and sinh) were allowed in the genetic programming. The test errors for the different approaches are listed in Table 4.2 it should be noted that in all cases the committee of models has smaller test errors than any of the individual models. The genetic programs are seen from Table 4.2 to perform less well when compared with the neural network.

Model	Test error
All functions genetic program, 3–member committee	0.611
Trigonometric functions genetic algorithm, 3–member committee	0.361
Neural network, 13–member committee	0.293

Table 4.2: Test errors for committees of models.

4.3.2 Predictions with Genetic Programming Model

The best way of assessing the performance of non–linear models is to use them to make predictions. Fig. 4.9 shows prediction for a 12Cr

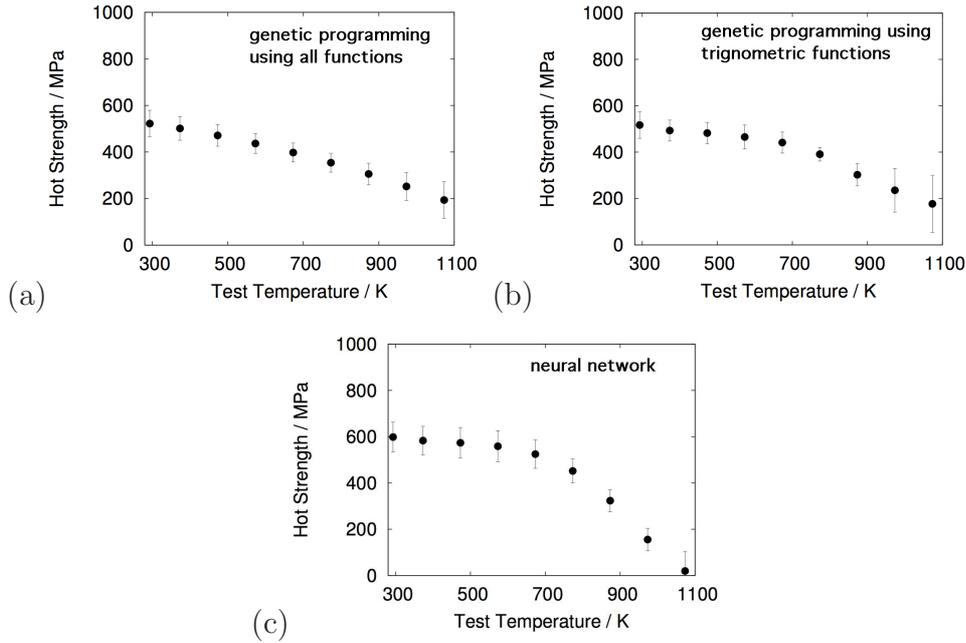steel, whereas Fig. 4.10 shows predictions for a 2.25Cr steel.

Figure 4.9: Calculations for steel containing 0.21C 0.44Si 0.62Mn 0.85Ni 11.64Cr 0.97Mo 0.06Cu 0.03Al 0.02N wt%, 1323 K for 25 min, air cooled, tempered at 913 K for 60 min. (a) Genetic programming using arithmetic, elementary and trigonometric functions. (b) Genetic programming using trigonometric functions. (c) Corresponding outcome using the neural network.

The predictions from the genetic program trained with arithmetic, elementary and trigonometric functions, failed to capture the complexity of the data. The different temperature dependency of stress on temperature on either side of T_C is not captured – instead, a rather stiff model is produced because the search for the final function was restricted to relatively simple formulae. With this limitation removed, the genetic program based on just the trigonometric functions is able to capture the stress–temperature behaviour correctly, probably because it includes hyperbolic functions which are know to be very flexible functions. Indeed, the neural network uses the hyperbolic tangent transfer
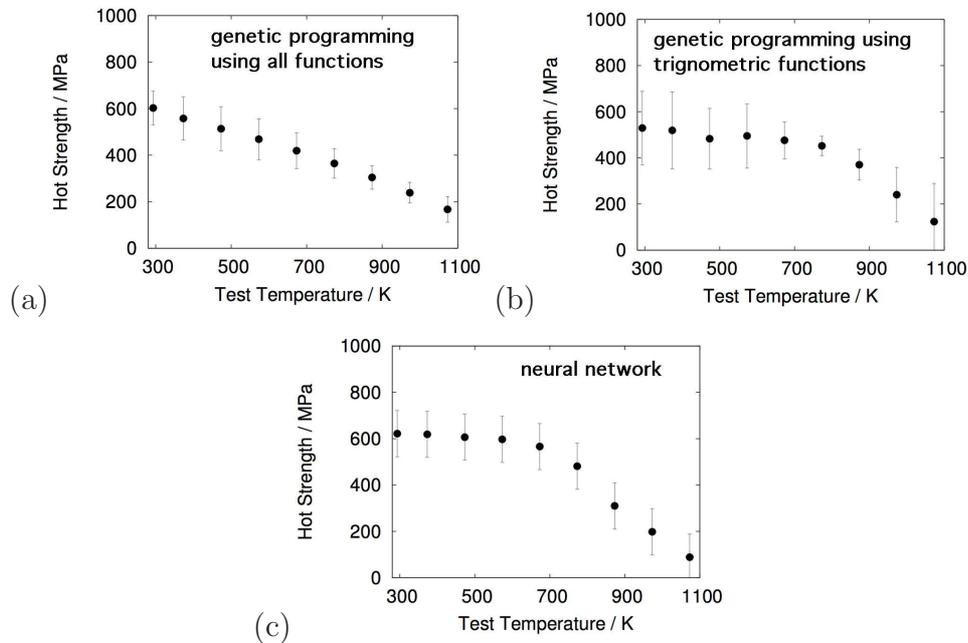
Figure 4.10: Calculations for steel containing 0.15C 0.25Si 0.61Mn 0.32Ni 2.35Cr 0.96Mo 0.03Cu 0.01Al 0.01N wt%, 1193 K for 210 min, air cooled, tempered at 988 K for 300 min. (a) Genetic programming using arithmetic, elementary and trigonometric functions. (b) Genetic programming using trigonometric functions. (c) Corresponding outcome using the neural network.

function as indicated in the previous chapter.

4.4   CONCLUSIONS

The strength of any steel can in principle be decomposed into components consisting of the strength of the respective phases that are present in the steel, or it can also be decomposed into parts that represent the strength of pure iron, solid solution strengthening, microstructure and grain size. However, the exact proportion contributed by each component to the overall strength it is not fully determined due to the myriad

of factors that determine the process. The evaluation of all these factors would require huge resources.

To estimate the hot–strength of creep–resistant ferritic steels as a function of chemical composition and heat treatment a neural network model was created. The model can successfully represent the hot–strength of steels based on chemical composition and heat treatment, it is easy to use and transparent to interpret; it does not require experiments in order to be used.

The model was combined with other observations to establish that there is a regime of temperature beyond which there is a steep decline in the microstructural contribution to strength. This decline cannot be attributed to changes in microstructure, but rather to an increased ability of dislocations to overcome obstacles with the help of thermal activation.

Creep–resistant ferritic steels are used for building steam turbines and steam pipes used in the energy sector for the generation of electricity, and they have a design life in excess of 25 years. When new alloys are invented and developed it is necessary therefore to engage in long-term testing. It was demonstrated using the mathematical modelling that in steam turbine conditions, the temperature dependence of the hot tensile-strength is essentially identical to that of creep deformation.

From a technological point of view, the similarity in $\partial \sigma_Y / \partial T$ for hot–strength and creep–rupture provides a ready method for estimating expensive creep data using ordinary tensile tests. Therefore, a short-term creep test combined with quick hot–strength data may be used to extrapolate the creep data to long times, thus reducing the need for creep testing.

For pure iron and solid solution strengthening the decrease of strength as the temperature increases is linear, however at T_C the strength of
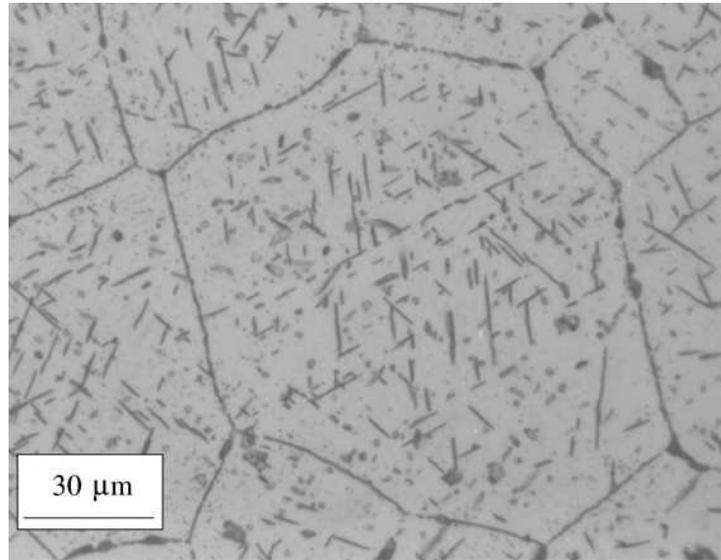
Figure 4.11: Micrograph of a steel, potentially able to withstand service temperatures in excess of 800 K [152].

the ferritic steels in taking a different slope, which is attributed to the fact that it becomes easier for dislocations to overcome obstacles by a thermally activated mechanism beyond T_C. It would be quite an achievement if a steel would be developed that can withstand service temperatures in excess of 800 K, due to the gains in the efficiency for the energy industry. This can be accomplished by strengthening the steel through solid solution alone resulting in a steel with very large grain. A possible candidate could be the one shown in Fig. 4.11 which achieves its strength due to solid solution strengthening instead of a fine microstructure.

Two empirical modelling methods were used on the same database, neural networks and genetic programming. The two methods are found to be similar in being able to create sufficiently complex non–linear functions to represent experimental information in ferritic creep–rupture resistant steels. In many respects, both methods require trials to find

the optimum overall function, but the genetic programming was more computer intensive and did not generalise as well as the highly–flexible neural network method, it also required greater intervention in the choice of functions.

Chapter 5

FATIGUE CRACK PROPAGATION

There are three steps in the failure of a component by fatigue. *Crack initiation* may occur at surface scratches caused by handling or machining, slip bands or dislocations intersecting the surface as a result of previous cyclic loading or work hardening, pores located at grain boundaries that are perpendicular to the applied stress direction, metallic and non–metallic inclusions, surface pits formed due to corrosion etc. [156–160].

Fatigue damage is described as the nucleation and growth of cracks to final failure, although the differentiation of their two stages can be ambiguous [161, 162].

After the crack is initiated and if the external stress persists the crack enters the *propagation* stage. Factors influencing the rate of growth are:

- compressive residual stresses which tend to close cracks, whilst tensile residual stresses stimulate the advance of the crack by

opening it [163];

- an increase in temperature improves toughness in ferritic steels, this in turn raises the fatigue limit and reduces the crack growth rates. These events are attributed to the decrease of tensile strength at high temperature [164, 165]. Other researchers have attributed the difference in fatigue crack growth between room temperature and 750°C to the variation in Young's modulus with temperature [166], which is due to the fact that at high temperatures the role of the interatomic bonding forces decreases;

- corrosive environments accelerate fatigue crack growth rates by some 2 to 3 orders of magnitude when compared with an air environment [167–169]. In air, fatigue-crack propagation is associated with alternating blunting and re-sharpening of the crack tip, whereas in corrosive environments cyclic cleavage may occur, by hydrogen embrittlement from electrochemical reactions associated with corrosion. The growth rates in vacuum are slower than in air because of the effect of moisture. Fatigue crack growth rates in a hydrogen environment are faster than in air because of hydrogen embrittlement [170, 171];

- a decrease in the test frequency can accelerate fatigue crack growth rates due either to hydrogen embrittlement or to an acceleration of the corrosion process at the crack tip [172–174];

- the ratio of the minimum to the maximum stress tends to proportionally increase fatigue crack growth rates by enhancing crack tip opening displacement[1] [175, 176];

- fatigue crack growth rates naturally depend on the stress intensity range ΔK, as will be seen later;

---

[1]The crack opening displacement represents the relative displacement of the two edges which form the crack.

- there are different ways of loading: *mode I* involves tension to separate crack surfaces, *mode II* is an in-plane shear mode where the crack surfaces slide over one another in a direction perpendicular to the leading edge of the crack, *mode III* is tearing and antiplane shear where the crack surfaces move relative to one another and parallel to the leading edge of the crack;

- standardized test specimens are used [177], such that the dimensions do not overly influence the crack propagation rate. The only influence comes from the thickness of the test specimens because it is important to maintain a state of plane strain in order to ensure that material properties are measured [165, 178];

- it is difficult to assess the influence of just one of the material properties on the crack propagation because they are in many cases interdependent [178, 179]. An increase in the Young's modulus however, will result in a proportional increase in the stress intensity factor [180]. The crack propagation rate is known to decrease with grain size [181, 182];

- the isolated effect of microstructure is difficult to assess. Studies on overaged and underaged 7049 aluminium alloys[2] revealed that the fatigue crack growth is influenced by the effect of microstructure and deformation mechanisms. More homogeneous slip occurs in the overaged condition, whereas localised slip leads to a crack branching tendency in the underaged condition, with a consequential increase in fatigue strength for a fine microstructure [183, 184].

Ultimate *failure* occurs when the crack cannot withstand the applied stress with fracture mechanisms consistent with ordinary tensile

---

[2]Zn 7.1, Mg 2.8, Cu 0.06, Cr 0.3, Si 0.1, Mn 0.06, Ti 0.05, Ga 0.01, Zr 0.1 in wt% resulting in a fine microstructure for the underaged condition and a coarse microstructure for the overaged condition.

deformation [185].

5.1   MODELS FOR CRACK PROPAGATION

An elegant model for fatigue crack growth was proposed by Paris *et al.*, one which is still used today. The model was the first to recognize and quantify a direct relationship between the fatigue crack growth rate and the stress intensity factor K [186]. In the crack propagation regime Paris and Erdogan discovered that the fatigue crack growth rate is described by a power law equation:

$$\frac{da}{dN} = C(\Delta K)^m \tag{5.1}$$

where a is the crack length, N represents the number of cycles, ΔK is the stress intensity range, C and m are constants depending on the material [21]. The Paris equation is extremely useful in representing the crack propagation data (Fig. 5.1) for a big variety of materials, but it has some disadvantages.

The Paris equation generally describes constant amplitude data. Elber modified the relation to include an effective stress intensity range ΔK_{eff} for variable amplitude loads. It was proposed that fatigue cracks grow only during the portion of the load cycle when the crack tip is open [187]:

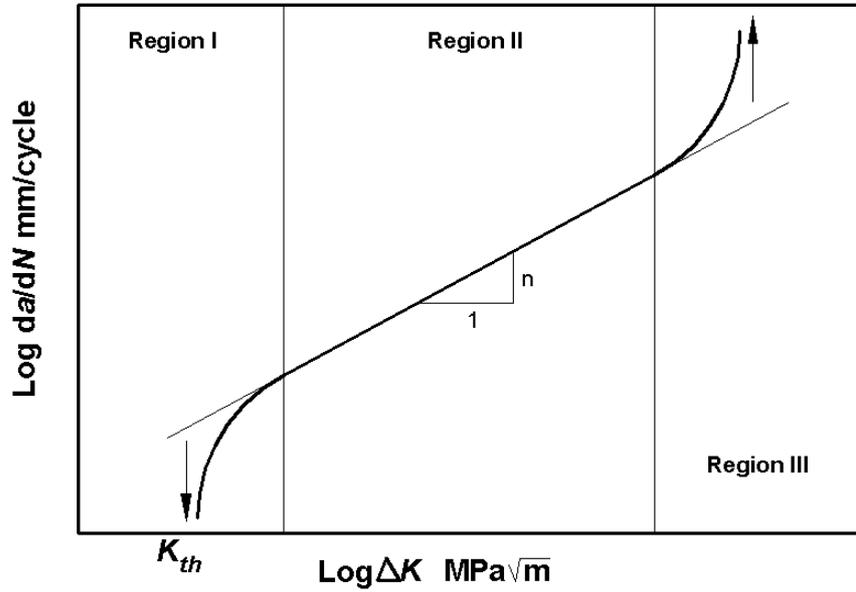$$\frac{da}{dN} = C_0(\Delta K_{eff})^m \tag{5.2}$$

where

Figure 5.1: Fatigue crack growth data.

$$\Delta K_{eff} = \begin{cases} K_{max} - K_{min} & \text{if } K_{min} > K_{open} \\ K_{max} - K_{open} & \text{if } K_{min} < K_{open} \end{cases}$$

where K_{open} is the applied stress required to open the crack, whereas K_{min} and K_{max} are the minimum respectively the maximum stress intensity factor, and:

$$C_0 = \frac{C}{0.7^m} \tag{5.3}$$

where C and m are the same constants from the Paris equation.

The Paris equation draws little influence from the material properties through the constants C and m. Duggan expressed the crack growth rate in terms of the elastic modulus, toughness, and ductility, postulating that the crack growth is critically dependent upon the

condition at the crack tip:

$$\frac{da}{dN} = \left(\frac{\pi}{32}\right)^{\frac{1}{2\eta}} \frac{1}{\eta} \left(\frac{2}{\epsilon_f E(K_c - K_{max})} \left(1 - \frac{K}{K_c}\right)\right)^{\frac{1}{\eta}} K^{\frac{2}{\eta}} \qquad (5.4)$$

where η is the so-called fatigue ductility exponent, ϵ_f is the fatigue ductility coefficient, E is the elastic modulus, K_c is the critical stress intensity factor, K_{max} is the maximum stress intensity factor and K is the stress intensity factor [188].

Ramsamooj and Shugar extended the Griffith[3] fracture criterion to include fatigue, they derived the following expression for fatigue crack growth rate:

$$\frac{da}{dN} = \frac{0.041}{EY}(K - K_{th})^2 \frac{1}{1 - \left(\frac{K_{max}}{K_{Ic}}\right)^2} \qquad (5.5)$$

where K is the stress intensity factor, $K_{I_{max}}$ is the maximum value of K_I, K_{Ic} is the fracture toughness, K_{th} is the stress intensity factor threshold, E is the elastic modulus, and Y is the yield strength [190]. The model gives satisfactory agreement with experimental data (tested against iron, aluminium and titanium alloys), but does not take account of frequency and it cannot be applied to other than mode I loading.

The Paris equation represents data only for the propagation stage of the crack growth. There have been attempts to overcome this. Khan and Paul proposed a new equation based on the Paris equation which would represent data for all three stages:

$$\frac{da}{dN} = C\left(\frac{K_{eff}}{Y}\right)^2 \left(\frac{\Delta K - \Delta K_{th}}{K_{Ic} - K_{max}}\right)^n \qquad (5.6)$$

where K_{eff} is the effective stress intensity factor, Y is the material's

---

[3]Griffith postulated that brittle fracture occurred when the energy release rate during crack growth exceeded the rate that energy was required [189].

yield strength, ΔK is the stress intensity range, ΔK_{th} is the stress intensity range threshold, K_{Ic} is the fracture toughness, K_{max} is the maximum stress intensity factor, the equation works well on 4340 steel[4], but it was not tested on other steels. It does not take account of frequency and can only be applied to mode I loading [191].

Finite element modelling was used by Socie for predicting the fatigue crack propagation of notched specimens subjected to irregular loading [192]. The model is restricted to the type of data used to create it. Furthermore finite element modelling is computationally intensive as its success depends on an appropriate mesh size, which needs to ensure that the strain gradients are not compromised.

Seed and Murphy used neural networks to find the constants m and C of the Paris equation; the results were encouraging but were limited to short cracks and only one type of steel [193]. Haque and Sudhakar applied neural networks to the fatigue crack growth behaviour of dual phase steels with various gradients of martensite; the model's behaviour on known data is positive but it is restricted to dual phase steels [194]. These models are useful but they have not been demonstrated to apply beyond their specific ranges studied.

An attempt will be made in the present work to develop a generally applicable neural network model based on physical parameters.

5.2   FATIGUE MODEL

The Bayesian framework chosen for the neural network model, described in Chapter 3, was used to express the fatigue crack growth rates of steels as a function of mechanical properties and test conditions.

---

[4]C 0.4, Cr 0.8, Mn 0.7, Mo 0.25, Ni 1.8, Si 0.2, P 0.035, S 0.04 in wt%.

5.2.1 The Variables

There exist a lot of data in the literature on the fatigue crack growth rates of metallic materials, which form the basis of the present work [195]. Steels were studied with the chemical compositions presented in Table 5.1. Traces of alloying elements (titanium, aluminium and vanadium) and residual elements (sulphur and phosphorus) were present in some of the steels, which were subjected to a maximum of three heat treatments.

Element	Minimum	Maximum
Carbon / wt%	0.1	0.8
Copper / wt%	0	0.2
Chromium / wt%	0	5
Manganese / wt%	0	2
Molybdenum / wt%	0	2
Nickel / wt%	0	2
Silicon / wt%	0	2

Table 5.1: Chemical composition range of the steels studied.

Fig. 5.2 illustrates the input variables, plotted versus the output $\frac{da}{dN}$, which were used to develop the model. Fig. 5.2 does not have the role of showing functional dependencies, it is intended just to show the distribution of the data without correlation between the different variables. The data are not distributed uniformly, but the Bayesian framework of the neural network recognises this by associating large modelling uncertainties with the sparse or noisy domains.

The properties of a steel depend on the detailed chemical composition and heat treatment, but these were not chosen as the input parameters, rather the mechanical properties which according to theory [13, 21, 186, 196] should control the crack growth rate during fatigue.
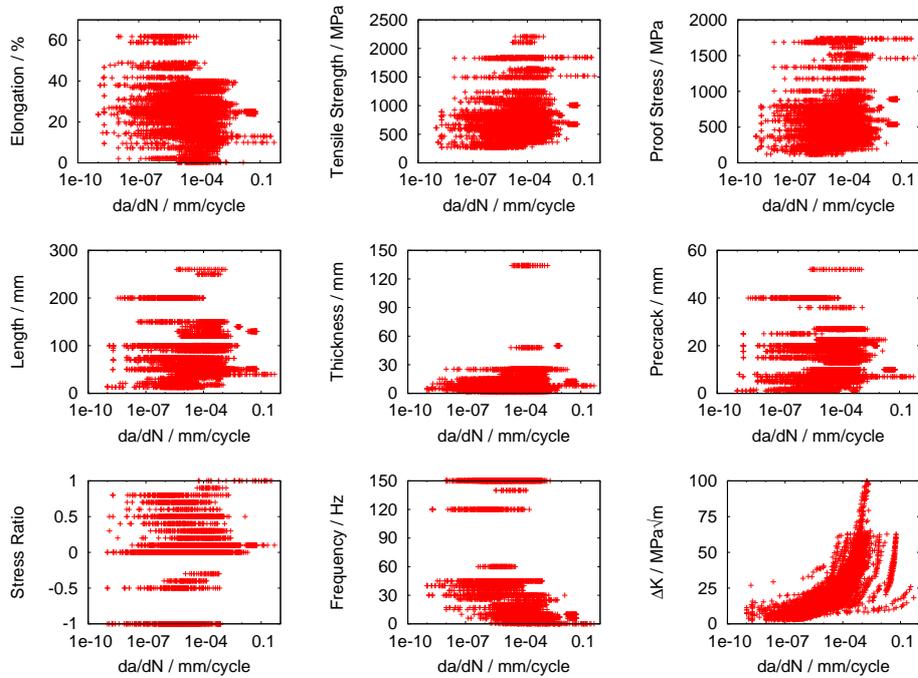
Figure 5.2: Visual illustration of the distribution of data used to create the model.

The dimensions of the test specimens and the test conditions are also important in this respect and were included as described in Table 5.2.

There were two types of loadings, an axial mode I and an in-plane bending mode II incorporated in the analysis; mode III data were not available.

5.2.2   Training the Model

The data were randomly split into two equal groups, the training and testing sets, which were normalized as described in the Chapter 3.

One hundred networks were trained, with hidden units ranging from one to twenty and five seeds in each chase. As the number of hidden

Variable	Minimum	Maximum
Elongation / %	0.2	61.8
0.2% Proof stress / MPa	121.59	1735.66
Tensile strength / MPa	270.65	2206.35
Specimen length / mm	13	260
Specimen thickness / mm	1.2	134
Pre-crack length / mm	1	52
Stress ratio	-1	1
Frequency / Hz	1	150
ΔK /	2.5	142
$\frac{da}{dN}$ / mm cycle^{-1}	9.82×10^{-10}	4.86×10^{-1}

Table 5.2: The variables and the corresponding output $\frac{da}{dN}$ the crack advance per cycle.

units increases so does the complexity and flexibility of the network, so the expected noise level from the training data decreased, with a minimum at nineteen hidden units, Fig. 5.3a.

As can be seen in Fig. 5.3b, the ability of the models to generalise on the test data does not decrease monotonically, but has an optimum at about nineteen hidden units, the log predictive error has an optimum at seventeen hidden units in Fig. 5.3c. The optimum number of models to form a committee was found to be seven as shown in Fig. 5.3d.

The final trained committee tested on the whole dataset shows reasonable agreement between the predicted and measured values; even those few data which are badly estimated are accompanied by uncertainties as shown in Fig. 5.5. There is clear improvement over the predictions shown in Fig. 5.4a and b by the best single model.

The network perceived significances are shown in Fig. 5.6. Both the mean significance and the upper and lower limits from the members of
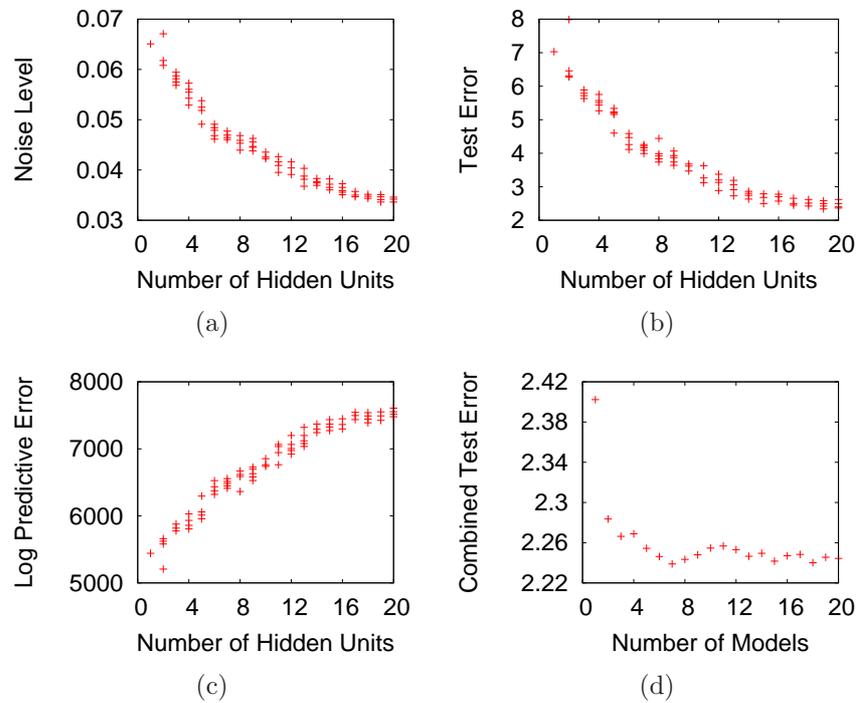
Figure 5.3: (a) Perceived level of noise for the training data (b) the test error (c) the log predictive error (d) the test error for the committees of different sizes.
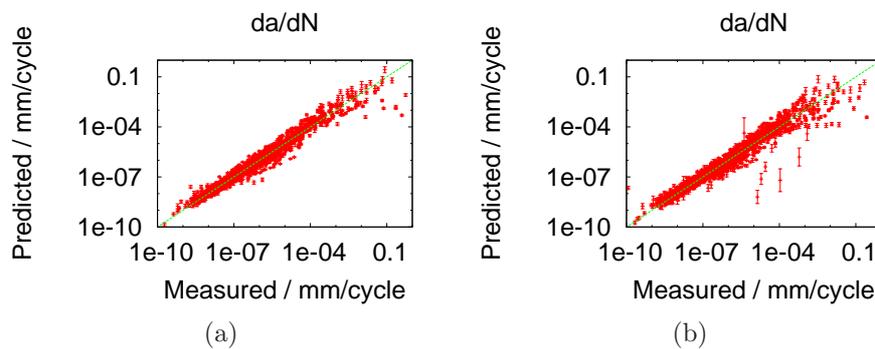

Figure 5.4: (a) The performance of the best single model on the training data, (b) The performance of the best single model on the testing data.
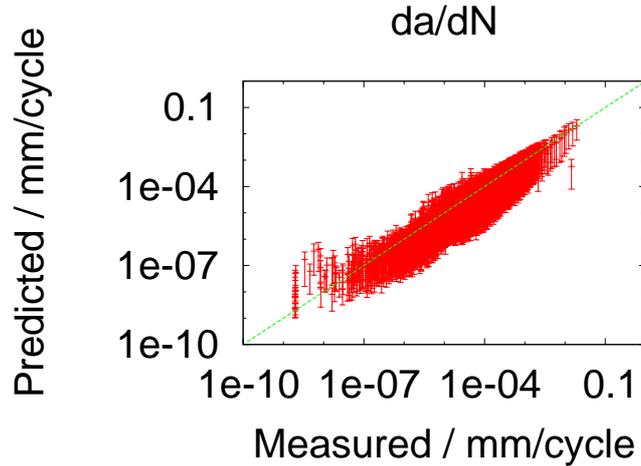
Figure 5.5: The performance of the committee of models on the whole database (training and testing data). For the committee model the standard deviation for predicted versus measured of 0.003 mm/cycle, and there were only 158 points more than three standard deviations away from their measured values out of 12807 data points.

the committee are shown.

The magnitude of the significance is a measure of the extent to which a particular input explains the variation of the output. Fig. 5.6 illustrates the significance of each input variable, as perceived by the neural network model. It can be seen that the mechanical properties (elongation, ultimate tensile strength and proof stress) have a great influence on the fatigue crack growth rates. However the input which influences the fatigue crack growth rate the most is the stress intensity range, this is normal since $\frac{da}{dN}$ is expected to be proportional with ΔK. The model also picks up an influence from specimen size. According to Griffiths and Richards, in the vast majority of the cases (because the specimens are thick plane strain condition exists) the specimen size has no significance on the fatigue crack growth [197].
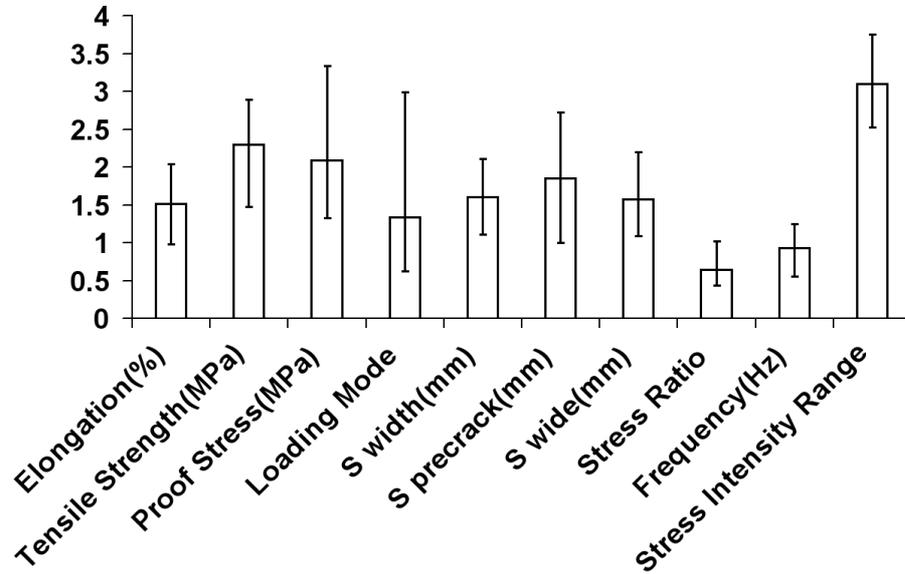
Figure 5.6: Perceived significance of the inputs in the committee model.

5.2.3 Behaviour of the Model

In order to evaluate the model behaviour a first set of estimations is illustrated in Fig. 5.7 for a typical steel having 9.5% elongation, 660 MPa 0.2% proof stress and 810 MPa tensile strength, considering the specimen length 70 mm, specimen thickness 15 mm, pre-crack size 5 mm, frequency 120 Hz, stress ratio 0 and 10 MPa$\sqrt{m}$ stress intensity factor. While each specific input shown in Fig. 5.7 is varied all the other ones are kept constant at the above mentioned values.

As it can be seen from Fig. 5.7 the input parameters have a complex influence on the crack growth rates. An increase in elongation produces an initial decrease in the crack growth rates until around 20% elongation after which there is a continuous increase. The behaviour of tensile strength is somewhat similar with that of elongation with an initial sharp decrease in the lower values followed by a sharp increase. For proof stress the behaviour is again similar with the other two mechanical

properties with an initial decrease followed by an increase. It can be concluded that a material with a good fatigue crack growth resistance will have 20 % elongation, 800 MPa proof stress and 1000 MPa tensile strength.
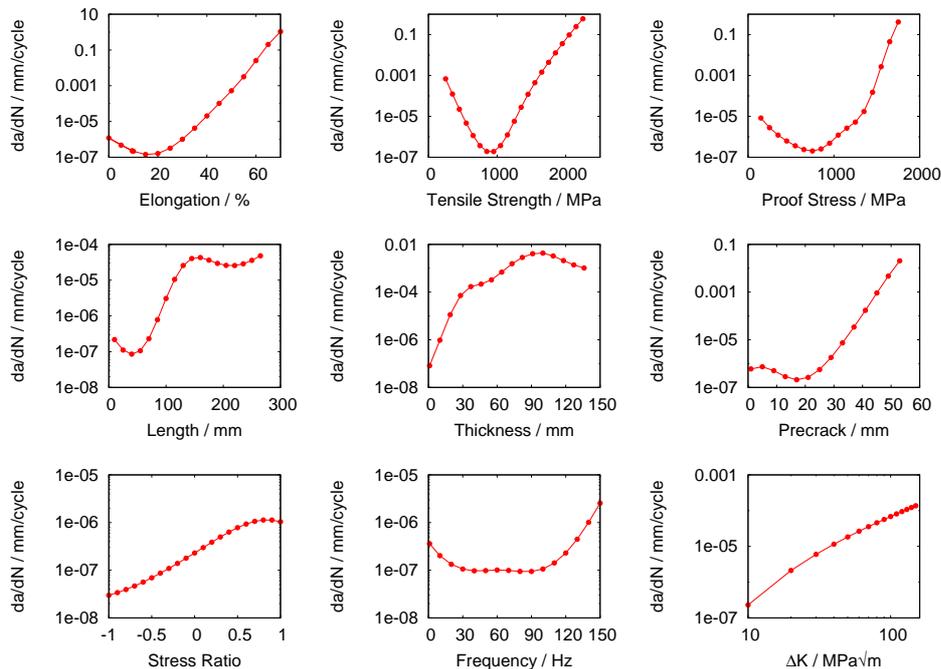

Figure 5.7: Behaviour of the model. The influence of the input parameters on the crack growt rates.

From the three specimen size inputs the only one which can be accounted for and verified against previous observations is the specimen thickness. An increase in thickness will lead to a change in the state of stress in the specimen from plane stress for thin specimens to plane stress in thick specimens resulting in higher rates in thicker sheets [198]. The change and the resulted difference in crack growth rates is due to fact that the crack growth is slower in plane stress at the same stress intensity.

The stress ratio is a very important parameter and its influence

was investigated previously and as can be seen in Fig. 5.7 decreasing the stress ratio results in smaller crack growth rates [199]. The effect of frequency is less pronounced than stress ratio with only a small difference in the fatigue crack rates between low and high frequency.

The best way of assessing a model is by making predictions. Such estimates are presented in Fig. 5.8, a bearing steel known in different countries under the names, SUJ 2[5] in Japan, AISI 52100[6] in U.S., EN 31[7] in Europe, a hard and tough steel. Its properties were not in the database used to create the model so the necessary inputs were collected from the published literature [200], including 5% elongation, 2030 MPa 0.2% proof stress, 2240 MPa tensile strength, loading mode 2, specimen length 80 mm, specimen thickness 2 mm, pre-crack size 3 mm, frequency 2 Hz and stress ratio 0. The prediction parameters were set so they would be the same as the ones reported in the experiments [200], so that a realistic assessment can be made of the performance of the model.

The measured data fall well within the range of the predictions. The model successfully captured the trends for the first and second region of the crack growth behaviour, *i.e.* crack initiation and the region which determines most of the fatigue life.

The factors which determine the crack properties of a material are numerous and amongst them, some are more important than others. It is in the author's opinion that from the materials point of view the factors which determine the fatigue crack initiation, propagation and failure are its mechanical properties. Factors such as the chemical composition, heat treatment and microstructure feature indirectly since they determine the mechanical properties. In these circumstances the model which is developed using data on steels should be applica-

---

[5]C 0.98, Mn 0.3, Cr 1.44, Si 0.24 in wt%.
[6]C 1.05, Mn 0.35, Cr 1.45, Si 0.35 in wt%.
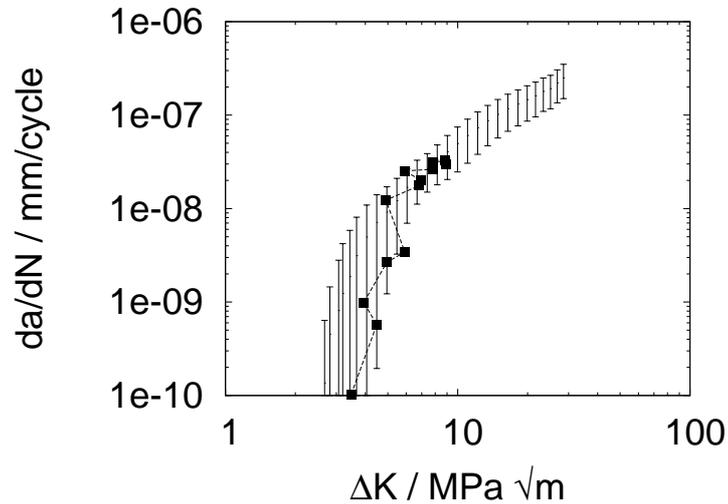[7]C 1, Mn 1.1, Cr 1.25, Si 0.25 in wt%.

Figure 5.8: Performance of the model, for a bearing steel. The points represent experimental data from [200], whereas the uncertainty range is obtained using the neural network model.

ble, without modifications, to other materials such as titanium and aluminium alloys. All that is required are the values of the relevant inputs.

Nickel–Base Superalloys

Calculations for three nickel–base superalloys (Udimet 700[8], Inconel 718[9] and Waspaloy[10]) are shown in Fig. 5.9, Fig. 5.10 and Fig. 5.11 where the uncertainty range represent the model predictions, whilst the circles represent the actual data points collected from the literature [201–205].

---

[8]C 0.06, Mn 0.1, Si 0.1, Co 16.6, Fe 0.23, Mo 4.95, Fe 1.6, Ti 3.48, Al 4.15, B 0.025, Zr 0.04, S 0.003, P 0.01, Cu 0.1 in wt%.

[9]C 0.08, Mn 0.35, Si 0.35, P 0.015, Cr 21, Co 1, Al 0.8, Mo 3.3, Ti 1.15, B 0.006, Cu 0.15, Ta 5.5 in wt%.

[10]C 0.1, Mn 0.5, Si 0.75, Cr 21, B 0.008, Fe 2, Co 15, Ti 3,25, Al 1.5, Mo 5, Zr 0.12, Cu 0.1, S 0.02 in wt%.

The inputs for the predictions in Figs. 5.9–5.11 are presented in Table 5.3–5.5 and are as follows:

| Variable	Fig. 5.9a	Fig. 5.9b
Elongation / %	5	15
0.2% Proof stress / MPa	1020	1172
Tensile strength / MPa	1520	1404
Specimen length / mm	72.5	63.5
Specimen thickness / mm	12.5	25.4
Pre-crack length / mm	12.5	18.3
Stress ratio	0.1	0.1
Frequency / Hz	40	20

Table 5.3: The corresponding inputs for the predictions in Fig. 5.9.

Variable	Fig. 5.10a	Fig. 5.10b
Elongation / %	20	20
0.2% Proof stress / MPa	1113	1113
Tensile strength / MPa	1373	1373
Specimen length / mm	50.8	31.8
Specimen thickness / mm	12.7	8.89
Pre-crack length / mm	6.4	5.3
Stress ratio	0.05	0.05
Frequency / Hz	0.667	0.667

Table 5.4: The corresponding inputs for the predictions in Fig. 5.10.

The results are fascinating since the model correctly calculates the Paris slopes for all the alloys, presented in Figs. 5.12–5.14, where the red line represent the predicted Paris slope, whilst the green line represents the actual measured Paris slope. In each case it only marginally overestimates the fatigue behaviour. The actual data lie within the uncertainty limits of the predictions.
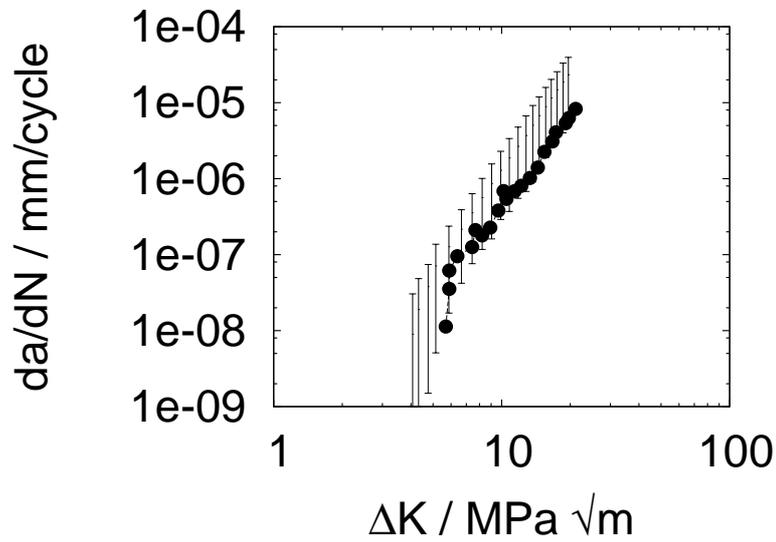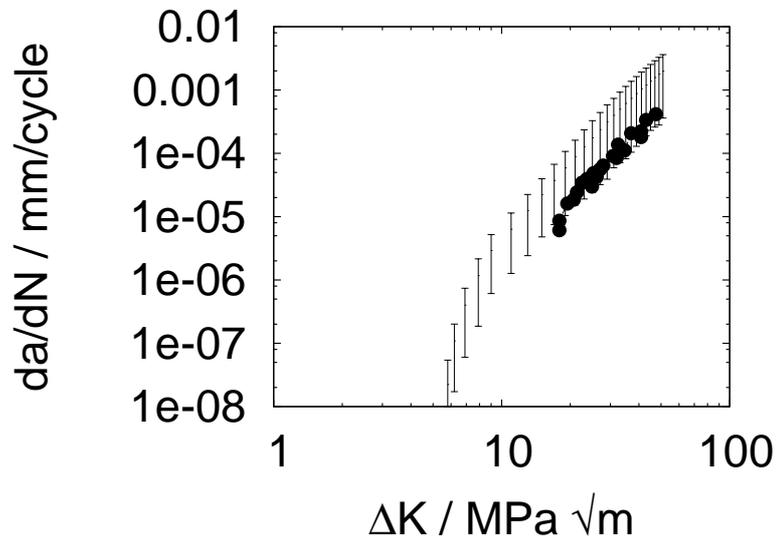
(a) Predictions for Udimet 700 nickel base superalloy [201]
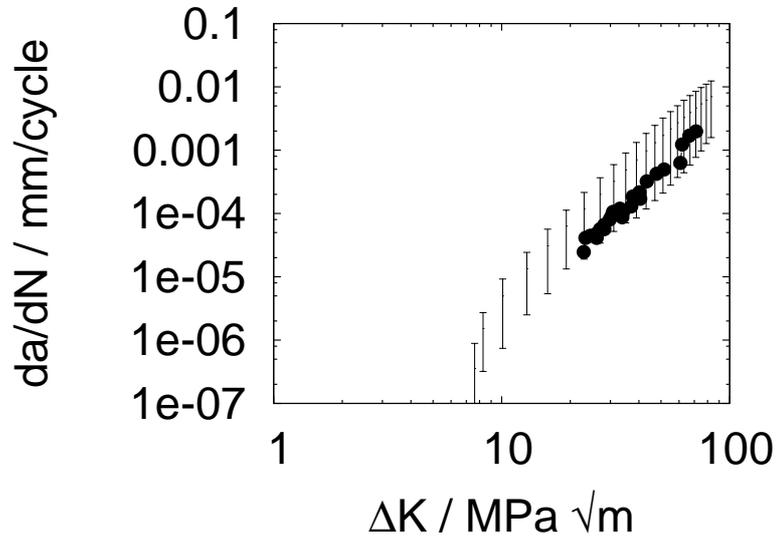
(b) Predictions for Inconel 718 nickel base superalloy [202]

Figure 5.9: Predictions for Udimet 700 and Inconel 718 nickel–base superalloys. The filled circles represent published data [201, 202] whilst the uncertainty range represent the model predictions.
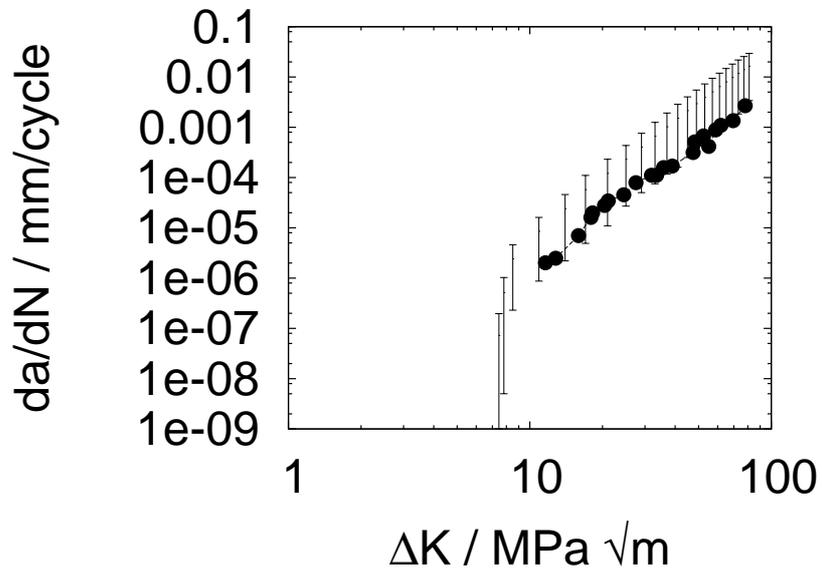
(a) Predictions for Inconel 718 nickel base superalloy [203]

(b) Predictions for Inconel 718 nickel base superalloy [203]

Figure 5.10: Predictions for Inconel 718 nickel–base superalloys. The filled circles represent published data [203] whilst the uncertainty range represent the model predictions.

(a) Predictions for Waspaloy nickel base superalloy [204]

(b) Predictions for Waspaloy nickel base superalloy [205]

Figure 5.11: Predictions for Waspaloy nickel–base superalloys. The filled circles represent published data [204, 205] whilst the uncertainty range represent the model predictions.
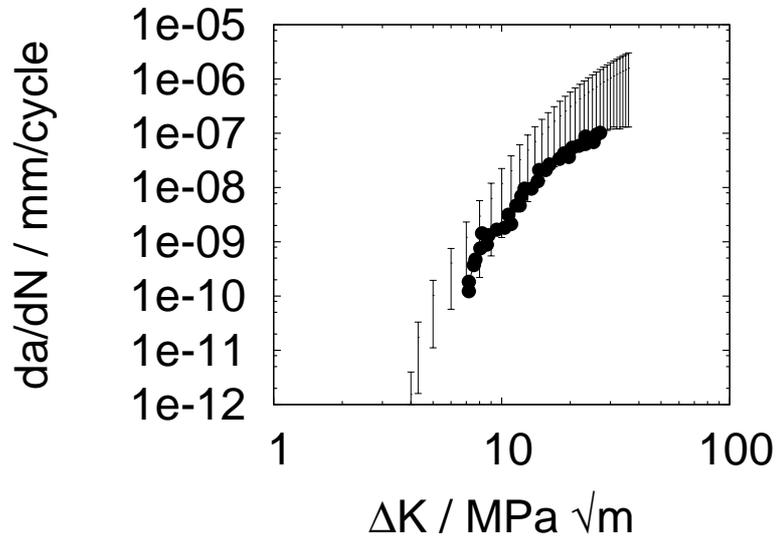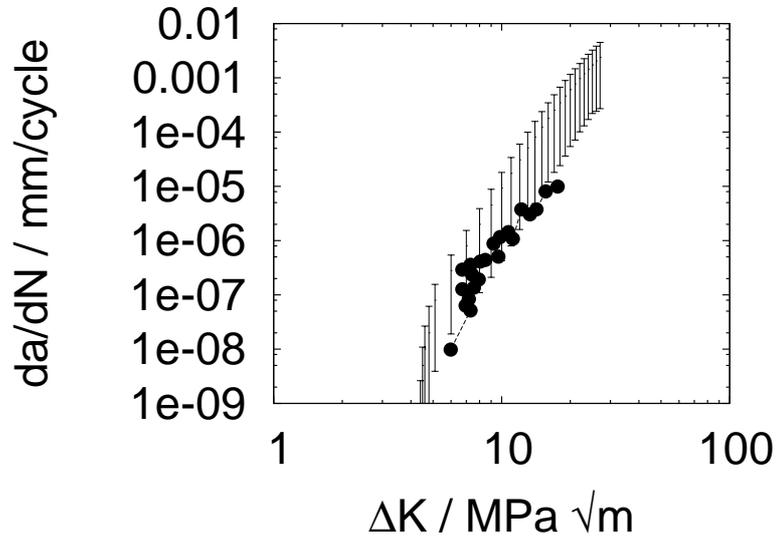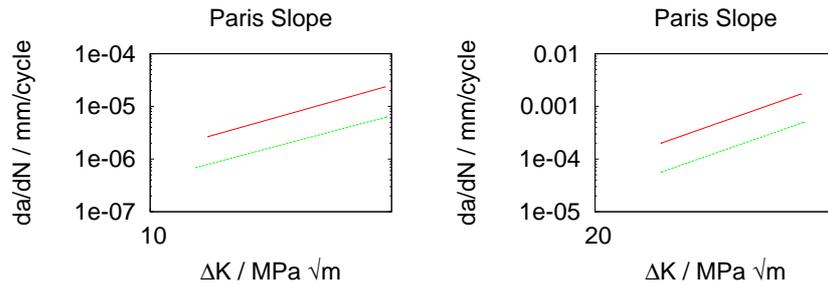
Variable	Fig. 5.11a	Fig. 5.11b
Elongation / %	27	33
0.2% Proof stress / MPa	1076	921
Tensile strength / MPa	1441	1351
Specimen length / mm	62.5	5
Specimen thickness / mm	25	3
Pre-crack length / mm	17.5	0.4
Stress ratio	0.5	0.5
Frequency / Hz	20	100

Table 5.5: The corresponding inputs for the predictions in Fig. 5.11.

(a) Corresponding slope for Fig. 5.9a (b) Corresponding slope for Fig. 5.9b

Figure 5.12: Calculated versus measured Paris slope for the nickel–base superalloys.

Titanium and Aluminium Alloys

Calculations for titanium [11] and aluminium [12] alloys were made and compared with actual data from the literature [206, 207].

Table 5.6 presents the corresponding inputs parameters for Fig. 5.15.

---

[11]Ti-6Al-4V alloy, Al 6.3, V 4.17, Fe 0.19, O 0.19, N 0.013, H 0.0035 in wt%.

[12]7075 aluminium alloy, Zn 7.2, Mg 2.8, Cu 1.7, Cr 0.06, Fe 0.3, Si 0.1, Mn 0.06, Ti 0.05, Ga 0.01 Zr 0.1 in wt%.

(a) Corresponding slope for Fig. 5.10a

(b) Corresponding slope for Fig. 5.10b
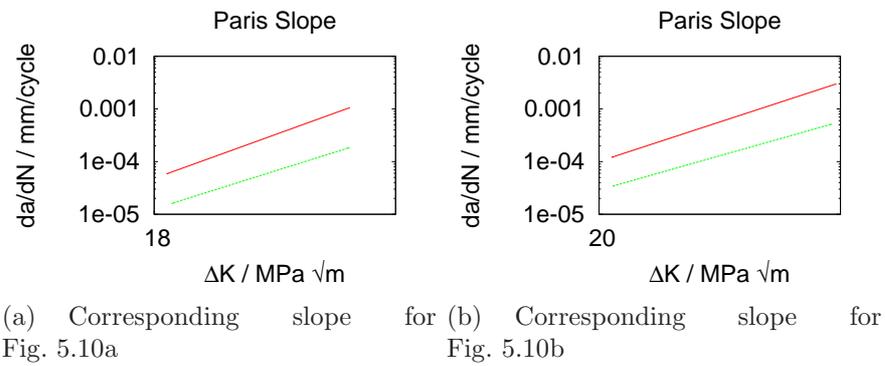
Figure 5.13: Calculated versus measured Paris slope for the nickel–base superalloys.

(a) Corresponding slope for Fig. 5.11a

(b) Corresponding slope for Fig. 5.11b
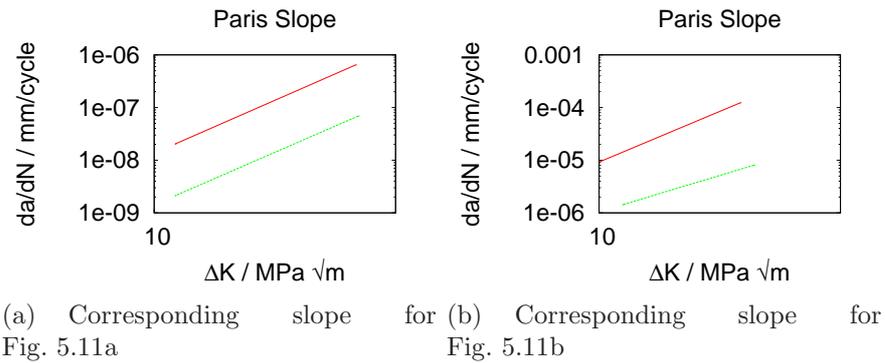
Figure 5.14: Calculated versus measured Paris slope for the nickel–base superalloys.

As was the case for the nickel–base superalloys the model nicely captured the slope for both the titanium and aluminium alloys, and it again slightly overestimates the fatigue crack growth rates.

Blind Predictions

In order to further test the model an aeroengine manufacturer was asked if they are interested in obtaining fatigue crack growth predic-

Variable	Fig. 5.15a	Fig. 5.15b
Elongation / %	14	8
0.2% Proof stress / MPa	930	524
Tensile strength / MPa	970	464
Specimen length / mm	155	155
Specimen thickness / mm	40	40
Pre-crack length / mm	9	9
Stress ratio	-1	0.5
Frequency / Hz	20	20

Table 5.6: The corresponding inputs for the predictions in Fig. 5.15.

tions on a material of their choice, with the purpose of making blind predictions. The company provided the mechanical properties from unidentified materials of their choice along with test specimen size and loading mode, in return they company received the predictions. They later revealed the actual fatigue growth rates obtained in tests together with the material type.

In Fig. 5.16 the uncertainty range represent the blind predictions made with the model whilst the filled circles represent the actual data sent back by the aeroengine manufacturer, after they consulted the predictions. Table 5.7 presents the inputs for the blind predictions.

The predictions are very good considering there were no data about the material in the database, the deviation from the actual data is small whereas the slope is nicely estimated.
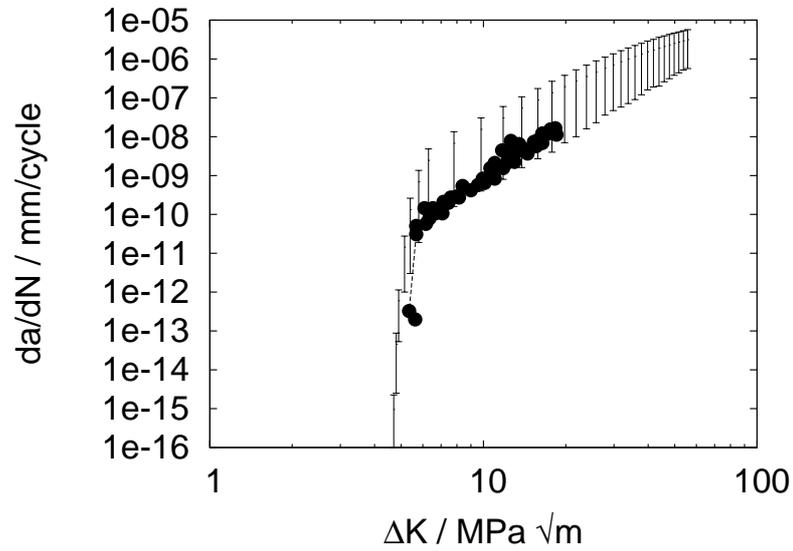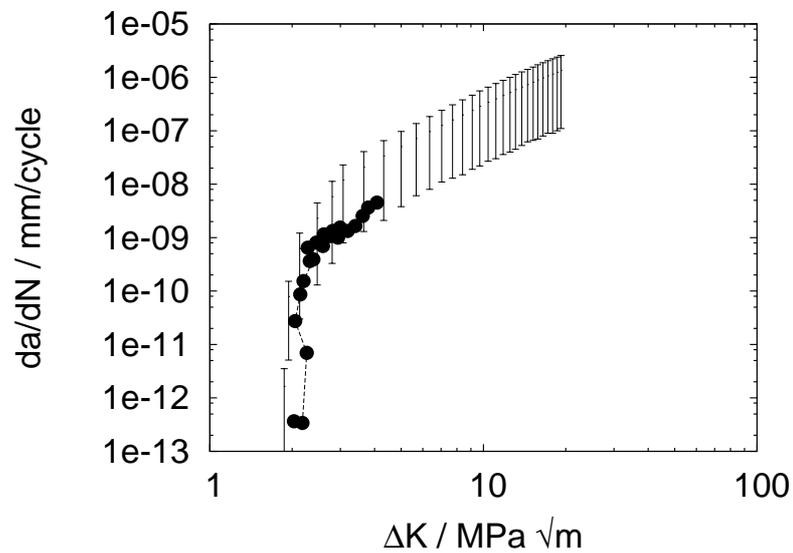
Figure 5.15: Calculation for titanium and aluminium alloys compared with actual data [206, 207], the uncertainty range represent neural network predictions whilst the filled circles represent the data.
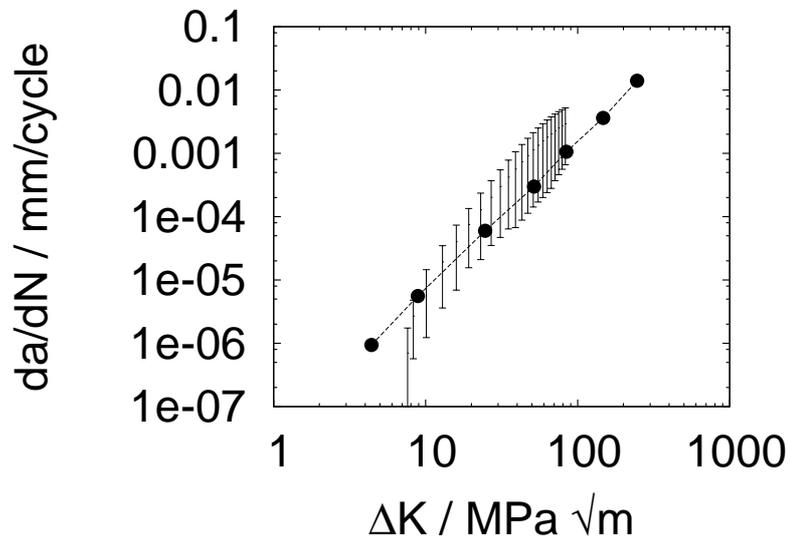
(a)
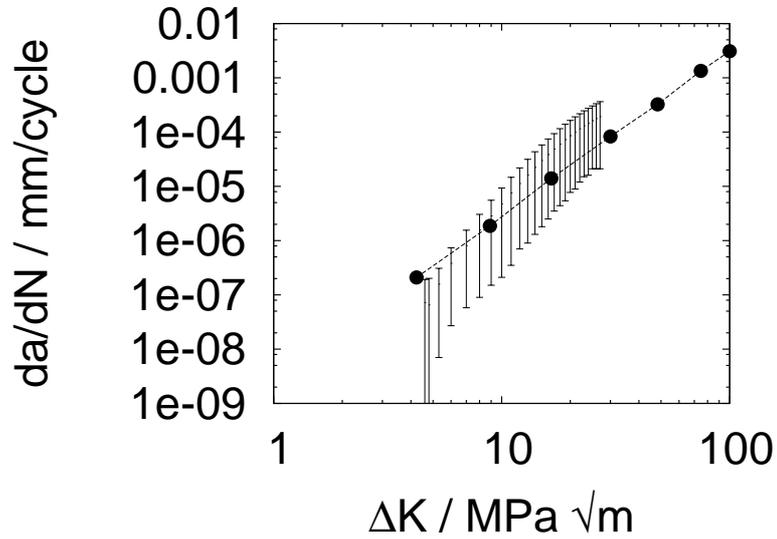
(b)

Figure 5.16: Blind predictions for Ti 6/4 forging material, having different heat trestments and chemical compositions resulting in various mechanical properties, the uncertainty range represent neural network predictions whilst the filled circles represent actual data.

Variable	Fig. 5.16a	Fig. 5.16b
Elongation / %	20	14
0.2% Proof stress / MPa	1172	940
Tensile strength / MPa	1440	998
Specimen length / mm	7	7
Specimen thickness / mm	7	7
Pre-crack length / mm	0.5	0.5
Stress ratio	0.1	0.5
Frequency / Hz	0.25	100

Table 5.7: The corresponding inputs for the predictions in Fig. 5.16.

5.3   CONCLUSIONS

A neural network model was developed, on steel data to enable the estimation of the fatigue crack growth rates as a function of material properties, test specimen size and loading conditions. The resulting model successfully predicts the fatigue crack growth rates in the first and second regimes for steels, even for data not included in the analysis, this was expected since the model was created on steel data and the number of data points available was large enough for the model to capture the interactions between the inputs and the output.

The model's ability to extrapolate was tested on bearing steel with properties which were not included in the training database and they were outside the data range for training and testing. The calculations successfully predicted the fatigue crack growth rates, when compared with data from the literature, predictions correspond with the data resulting in the same slope and behaviour.

A considerable series of calculations were conducted on nickel, titanium and aluminium alloys with the aim of establishing whether the model, with only as mechanical properties as inputs, generalises on

materials which are fundamentally different. This would be of great importance since when a new material is designed, or an old one is modified, only the mechanical properties are necessary in establishing the fatigue behaviour, with the microstructure and heat treatment implicitly included through their influence on the materials properties. The model can be exploited in the design of new metallic alloys.

Regarding nickel, titanium and aluminium alloys the model has been demonstrated to be able to reasonably estimate the fatigue crack growth rates. There is a persistent small overestimation of the fatigue crack growth rates though this is always within the predicted uncertainties.

Chapter 6

CONTACT FATIGUE

The life of a component is determined by its ability to perform its intended role; when it is no longer able to do this, it is considered to have failed. In the case of ball–bearings such as the one shown in Fig. 6.1, a failure is considered when the bearing is no longer able to perform rotations or when rotations take place with excessive vibrations and noise [9]. Complete failures happen in very few cases because the bearings are stopped when vibrations and noise are detected (continued use leads to seizure). Vibration and noise normally arise due to the formation of spalls and pits at the contact surface between the raceway and the balls [10].

The majority of such defects develop below the contact surface and their formation is influenced by:

- contact pressure increase results in a greater probability of defects appearance;

- the number of revolutions that the bearing endures relates directly to spall and pit formation;

Figure 6.1: Schematic representation of a ball–bearing, formed from an inner ring blue, an outer ring purple and the balls yellow which are kept aligned by a cage green.

- a greater frequency of revolutions correlates with a larger rate of defect formation;

- an increase in temperature leads to lubricant and seal failure, resulting in metal to metal contact between ball and raceway, thus accelerating failure;

- lubricant serves to minimise metal–to–metal contact and to transmit load. Its viscosity decreases as the temperature increases with a dramatic drop in performance when temperature increases by as little as 20°C [208];

- the seal, which keeps the lubricant in place;

- the nature of the material determines its response to fatigue.

The mechanisms leading to spall and pit formation are not fully understood. Residual stress which accumulates with each cycle, is one culprit [25, 209–212]. There are puzzling microstructural changes during service which some researchers correlate with the development of residual stress [213–217]. Others point out that the changes take place at different depths beneath the raceway, compared to where the peak residual stress occurs; the changes are instead attributed to load and temperature, the very factors responsible for residual stress development [218–221].
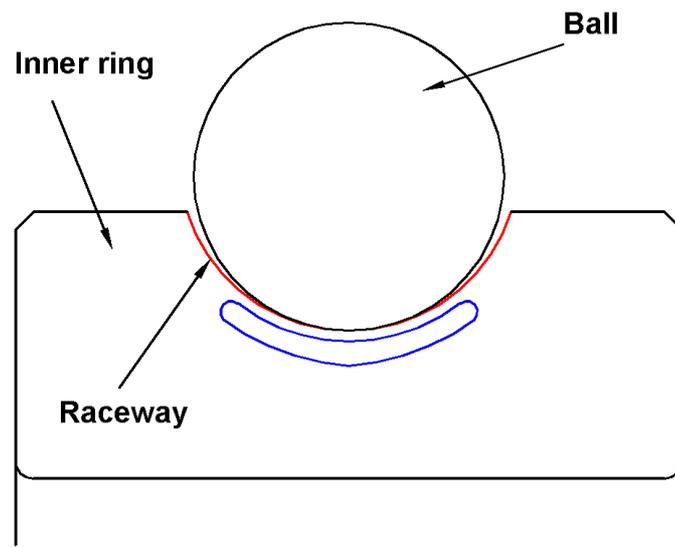
Figure 6.2: Schematic representation of the inner ring and ball in the context of a bearing. The blue area is where microstructural changes take place, this being approximately the same zone where the residual stress peaks.

The depth beneath the raceway where the maximum residual stress is recorded varies as a function of bearing geometry [211]. The depth where microstructural changes take place is not as easy to correlate with bearing geometry, load and temperature [215, 216].

The current method for estimating the life of bearings was proposed

by Lundberg and Plamgren [222], an empirical formula which only takes account of load, and is based on Weibull [223] distribution[1].

$$L_{10} = \left(\frac{D}{P}\right)^p \tag{6.1}$$

where L_{10} represents the life of the bearings when 10 % of the bearings fail before the estimated life time, D represents the bearing dynamic load *i.e.* the maximum load that the bearing can endure, P represents the equivalent load, and p is an exponent that varies according to the bearing geometry.

The Lundberg and Plamgren formula was modified, by Ioannides *et al.* [224][2], to take in account the nature of the material used and the type of lubricant.

$$L_n = a_1 a_2 a_3 \left(\frac{D}{P}\right)^p \tag{6.2}$$

where a_1 is a reliability constant that is equal to unity if the reliability level is 10 %, a_2 is the constant which takes into account the material fatigue properties and a_3 is the constant which is related to lubricant

Introducing residual stress and microstructural change in this model is difficult since the model is statistical whereas residual stress and structure are irregular and complex. For this reason, a Bayesian neural network model may be more appropriate [85, 142, 143, 155].

---

[1]The Weibull distribution (named after Waloddi Weibull) is a continuous probability distribution, with the essential property of being versatile so it can take on the characteristics of other types of distributions.

[2]The original formula needed to be modified because it was experimentally observed that it no longer predicted the life of bearings with the desired accuracy, because the influence of material and lubricant became greater.

6.1 RESIDUAL STRESS MODEL

The Bayesian neural network method used was described in Chapter 3. There are neural networks models in the literature which deal with residual stress, but none deals with the development of residual stress during the life of a bearing. The majority focus on stress due to machining or other processing stages [225–228].

6.1.1 The Variables

Experimental information on the distribution of subsurface residual stress in the ball–bearing inner ring, run under different testing conditions, were used for the development of the model. The inputs to the network are: contact pressure, hoop stress, revolutions, logarithm of revolutions, outer ring temperature and depth, presented in Table 6.1, whereas the output is the residual stress which develops in the inner ring of the bearings. The logarithm of revolutions was used in order to get a better distribution of data.

Variable	Minimum	Maximum
Outer ring temperature / ° C	73	123
Revolutions	26000	264000000
Contact pressure / MPa	2800	3300
Hoop stress / MPa	70	170
Depth /μm	0	700
Residual stress / MPa	-1110	257

Table 6.1: The variables used to develop the model.

It is emphasised that unlike linear regression analysis, the ranges stated in Table 1 and graphically shown in Fig. 6.3 are not there to define the range of applicability but just to give an impression of the distribution of the available data. This is because the inputs are not nec-

essarily uniform in all the dimensions of the input space. The Bayesian framework used makes possible the calculation of uncertainties which vary in magnitude with position in the input space, and it is these that indicate the risk of extrapolation.
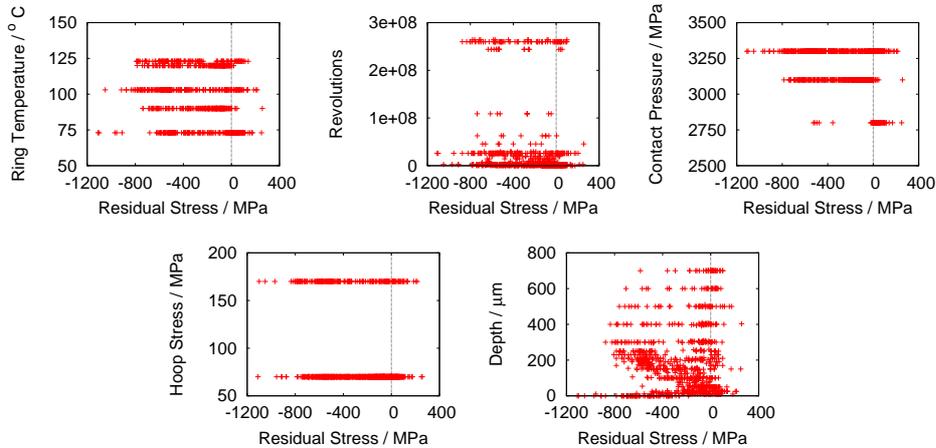
Figure 6.3: Visual illustration of the distribution of data used to create the model.

6.1.2   Training the Model

The data are split into training and testing sets, and the networks are then normalized as described in Chapter 3. The model creation trends are illustrated in Fig. 6.4.

Fig. 6.4b reflects the ability of the models to generalise on the test data, showing an optimum at about six hidden units, whereas log predictive error associates the optimum at eight hidden units (Fig. 6.4c). The optimum number of models to form a committee was found to be three as shown in Fig. 6.4d. The difference between the log predictive error and test error arises from the different formulas employed to calculate them, the performances of different models are best evaluated using the log predictive error because it penalises unexpected predic-
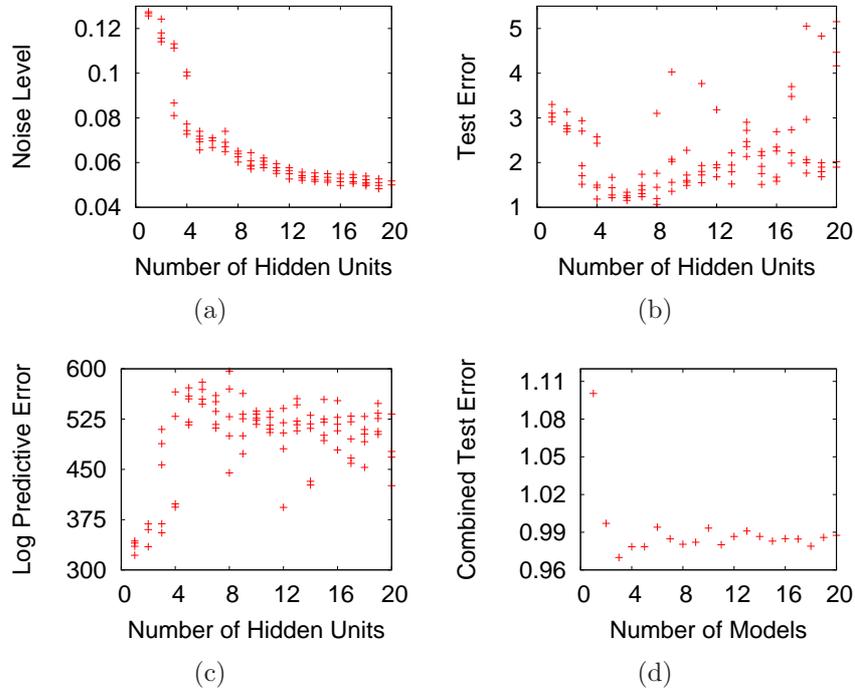
Figure 6.4: (a) Noise for the training data (b) the test error (c) the log predictive error (d) the test error for the committees of different sizes.

tions when they are accompanied by large uncertainties (the test error does not have this advantage).

The network perceived significances are shown in Fig. 6.5.

The depth where the residual stress is calculated is the most significant input to the network, entirely reasonable since the depth where the stress develops in not uniform but has a peak followed by a decrease. Contact pressure, revolutions and temperature follow in this order as the most influential parameters to the model, the order is not random as temperature depends on revolutions which in turn depends on contact pressure to develop the residual stress. The effect of the hoop stress is small relative to contact pressure and hence has a low significance. A bearing usually is subjected to a hoop stress of 50–200 MPa whereas
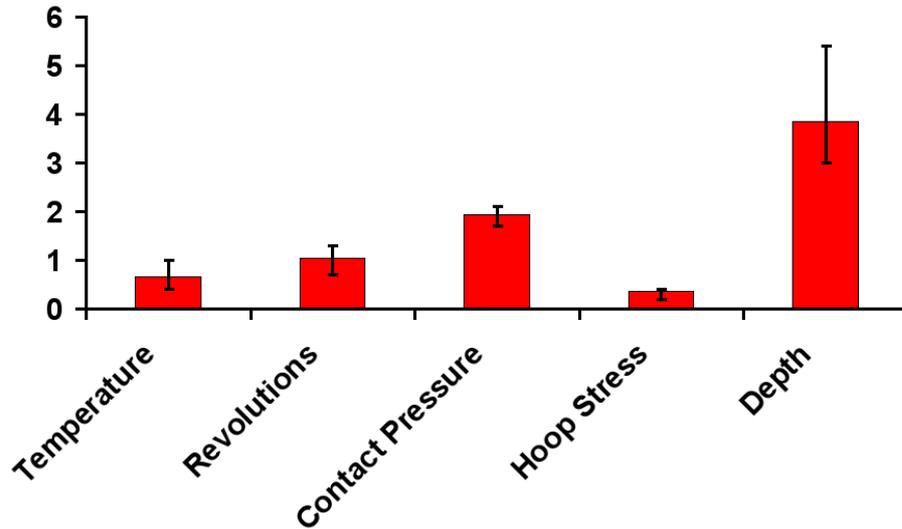
Figure 6.5: Model perceived significance for the committee.

contact pressure ranges between 1000 up to 5000 MPa [208].

The final trained committee tested on the whole dataset shows reasonable agreement between the predicted and measured values; even those few data which are badly estimated are accompanied by large uncertainties as shown in Fig. 6.6. There is clear improvement over the predictions shown in Fig. 6.7a and Fig. 6.7b by the best single model, which justifies the use of a committee of models instead of the best model.

6.1.3   Model's Behaviour

Fig. 6.8 shows calculation for the distribution of residual stress, in the radial direction. Predictions are made for depth ranging from the surface to a distance of 1000 μm; all the other input parameters are kept constant at the following values: contact pressure 3300 MPa, 2.6×10^8 revolutions, 123°C ring temperature and hoop stress of 170 MPa.
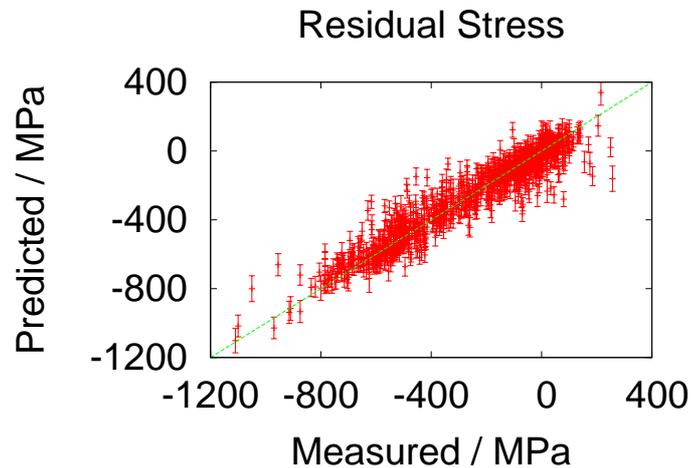
Figure 6.6: The performance of the committee of models on the whole database. The standard deviation is 98 MPa, and there were 47 points more than three standard deviations away from their measured values, out of a total of 922 data points.
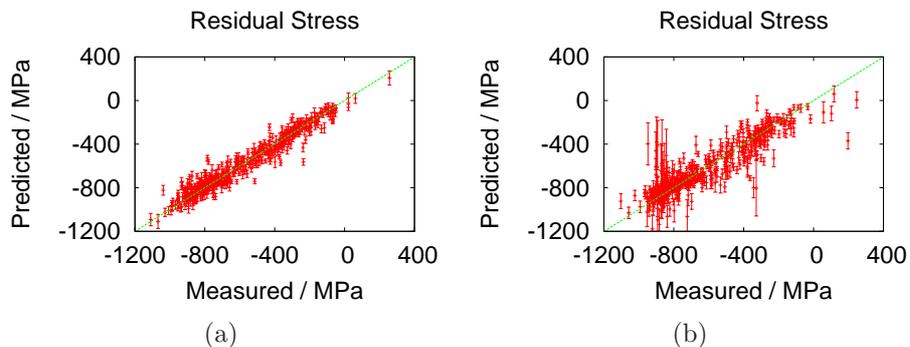
Figure 6.7: (a) The performance of the best single model on the training data; (b) the performance of the best single model on the testing data.

The calculations presented in Fig. 6.8 captures the transition, of residual stress, from compression to tension and back to compression, then beginning to level at around 0 MPa after 800 μm depth. A second characteristic is the peak at around 50 μm in the tensile region
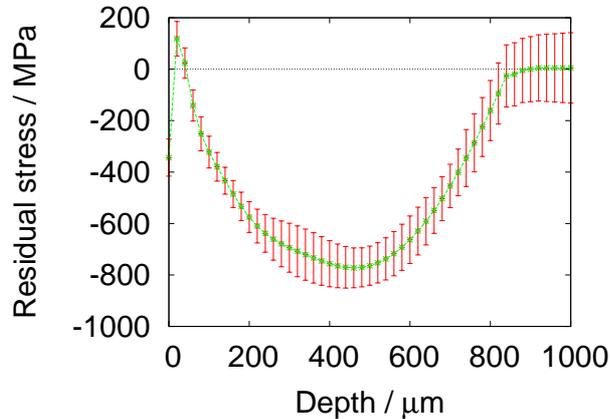
Figure 6.8: Residual stress distribution in the radial direction beneath the raceway. Input parameters contact pressure 3300 MPa, revolutions 2.6×10^8, ring temperature 123°C and hoop stress 170 MPa.

and a minimum at around 400 μm in the compression region. This double transition with a levelling around 0 MPa has been observed experimentally [25, 209–212, 215, 216]. The transition is due to a combination of rolling and sliding present simultaneously during the life of bearings. The curve is composed of two different ones which if overlaid results in the residual stress shown in Fig. 6.8. The first one results from sliding and creates a tensile effect which looses its strength quickly as the depth increases and the second one results from rolling has a stronger compressive effect and it takes deeper depths to lose its strength (Fig. 6.9).

Calculations for contact pressure are presented in Fig. 6.10, for contact pressure ranging from 2000 to 5000 MPa at three different assumed depths in the bearing raceway (the surface, 50 and 400 μm depth) with all the other inputs kept constant.

The predictions for the surface (Fig. 6.10a) are as expected since the residual stress is not influenced by contact pressure, but determined mostly by previous machining and heat treatment operations.
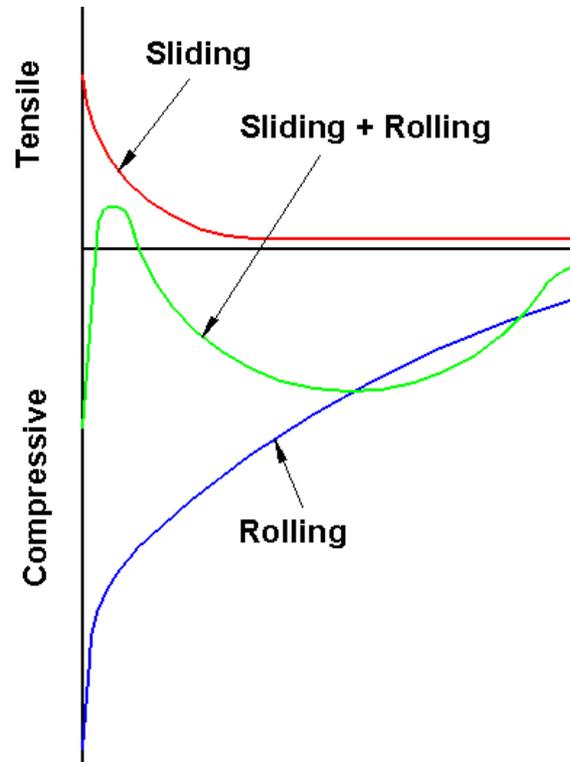
Figure 6.9: Schematic representation of how the complex development beneath the raceway of bearings takes place.

In Figs. 6.10b and 6.10c the magnitude of residual stress increases (in the tensile and respective compressive region) proportionally with the contact pressure, again expected [211, 221].

Hoop stress development is presented in Fig. 6.11 for values between 50 and 300 MPa, at the same depths as contact pressure whilst the rest of the parameters were contact pressure 300 MPa, revolutions 2.6×10^8 and ring temperature 123°C. The influence of hoop stress is not very significant as it is not an active parameter but is a bearing assembly parameter and secondly because the value itself is small, and seldom surpasses 400 MPa [208].

The influence of temperature on residual stress is shown in Fig. 6.12
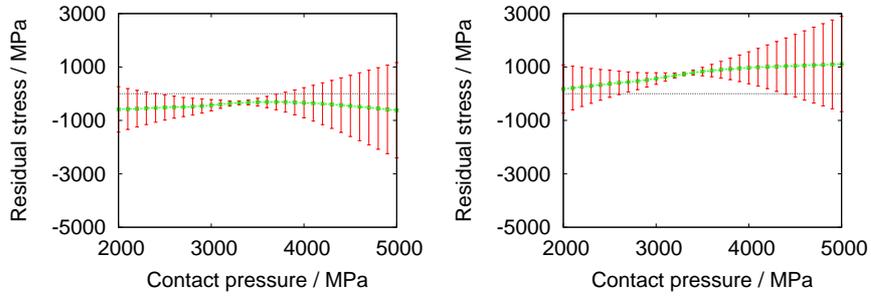
(a) Contact pressure at the surface (b) Contact pressure at 50 μm depth
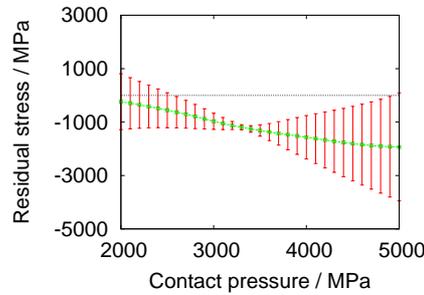
(c) Contact pressure at 400 μm depth

Figure 6.10: The values of the input parameters kept constant are ring temperature 123°C, hoop stress 170 MPa and 2.6×10^8 revolutions.

over the range 50 to 200°C with the other parameters locked at 3300 MPa contact pressure, 70 MPa hoop stress and 2.6×10^8 revolution. At all the different depths as the outer ring temperature is increased the accumulation of residual stress begins with a small decrease followed by an increase. The influence of increasing the temperature is not big by itself but it facilities kinetic processes in the steel which result in residual stress development.

The effect of increasing the number of revolutions is revealed in Fig. 6.13 at the same depth as the previous figures and for 3300 MPa contact pressure, 170 hoop stress and 123°C outer ring temperature. The predictions are as expected since the magnitude of the residual stress increases with the number of revolutions. This has been ob-
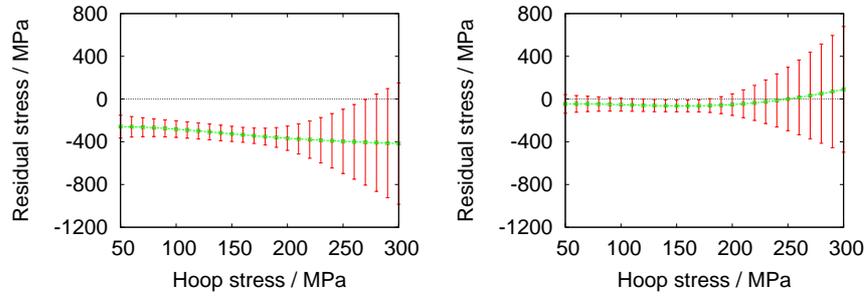
(a) Hoop stress development at the surface.

(b) Hoop stress at 50 μm depth.
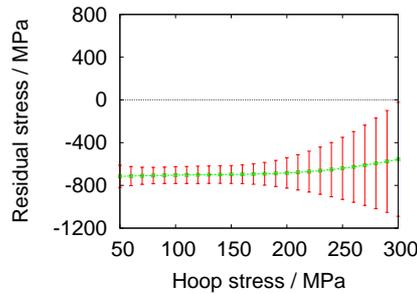
(c) Hoop stress at 400 μm depth.

Figure 6.11: The fixed values for the input parameters: 3300 MPa contact pressure, 2.6×10^8 revolutions and 123°C ring temperature.

served before and is due to the fact that with each new revolution, the bearing accumulates a tiny amount of residual stress [211, 221].

6.2 CONCLUSIONS

The model was developed to estimate the residual stress development below the raceway of ball–bearings as a function of contact pressure, hoop stress, revolutions, outer ring temperature and depth. The model accurately expressed the influence that depth has on the development of residual stress, followed by contact pressure revolution and temperature. The uncertainties are larger when extrapolating beyond
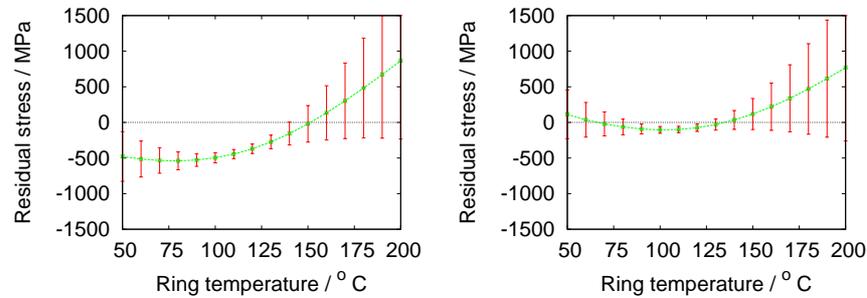
(a) Ring temperature influence on residual stress at the surface.

(b) Ring temperature influence at 50 μm depth.
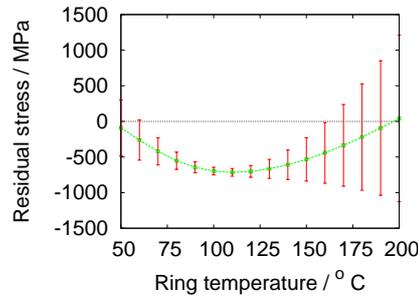
(c) Ring temperature influence at 400 μm depth.

Figure 6.12: Values of the fixed input parameters: contact pressure 3300 MPa, hoop stress 170, revolutions 2.64×10^8.

populated regions of the input space, to eliminate this inconvenience more data are required for those particular areas.

The model's performance was tested on all the input parameters and the calculation are considered successful, especially for depth where the complex shape of the curve was captured and when compared to experimental data is very similar [211, 221]. The influence of contact pressure, revolutions, temperature and hoop stress was expected as an increase in each of them will lead to residual stress accumulation.
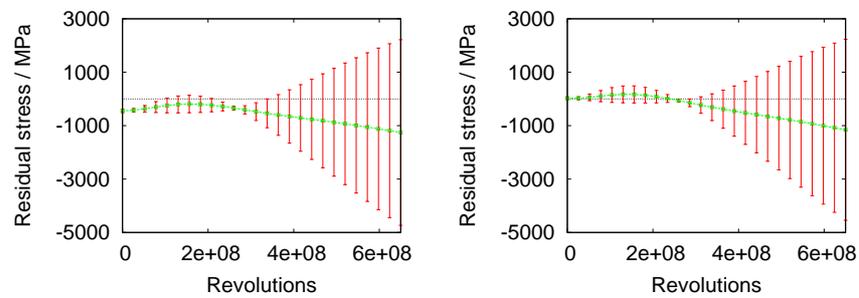
(a) Revolutions effect at the surface on residual stress

(b) Revolutions at 50 μm depth
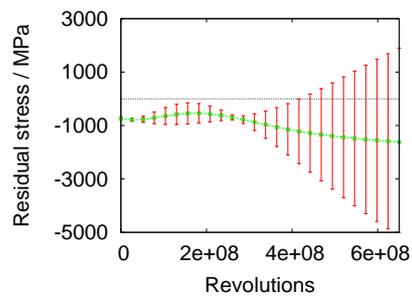
(c) Revolutions at 400 μm depth

Figure 6.13: Input parameters values contact pressure 3300 MPa, hoop stress 170 MPa, outer ring temperature 123°C.

Chapter 7

CONCLUSIONS AND FUTURE WORK

In this thesis considerable progress has been made towards the quantitative estimation of complicated properties such as hot–strength, fatigue crack growth rates and residual stress. The work could be seen as the beginning of an adventure to similarly cover all other properties of interest and to lead eventually to a suite of methods useful in design. The exciting outcome is that the models successfully captured known relationships and revealed new trends, thus making it clear that well designed models based on information theory can be exploited even in regimes where extrapolation is required.

One of the most important outcomes, which will lead to shortening testing times and thus cost savings, is the remarkable similitude between hot–strength and creep strength, making it possible to estimate the long time (10^5 hours) creep strength of ferritic steels from the hot–strength corroborated with a short time (10^3 hours) creep test. This is due to the fact that both hot–strength, which is quantified through short term tests and creep strength, which is quantified through pro-

117

longed testing, have their thermal dependence influenced by the same factors *i.e.*, climb mechanisms.

The model which describes the hot–strength as a function of chemical composition and heat treatment is a ready–to–use tool that enables the estimation of hot–strength for ferritic steels. This is of value as it reduces the need for testing since it is possible to estimate the hot–strength during the development of new steels.

The energy sector relies heavily on ferritic steels pipes for the transport of steam at high temperatures, and for the manufacture of boilers. The operating temperature of these components is limited by the capability of the steel to perform its intended function. An operating temperature of 650 °C is desired but not currently possible. The model findings lead to the conclusion that a different concept is required in which the steel should not be strengthened by fine grains, but exclusively by solid solution strengthening.

The hot–strength model could be enhanced as follows:

- the possibility that creep rupture strain and hot–ductility can be correlated should be investigated;

- from a fundamental point of view conducting creep tests for temperatures below 700 K would be useful for further validation.

The first appreciable result of the fatigue model is its ability to calculate the fatigue crack growth rates for steels. Another exciting outcome is that the fatigue model created for steels seems to properly represent the behaviour of other metallic materials such as alloys of titanium, aluminium and nickel. This is the first time that such a procedure has been attempted and the results are encouraging. This is probably a reflection of the fact that fatigue has been expressed in terms

of simpler mechanical properties, rather than factors such as composition and heat treatment, which determine the properties. The latter dependence will de different for different materials. It is not beyond the imagination that creep properties may be similarly expressed, using inputs consisting of the melting temperature, activation energy for diffusion mechanical properties, without incorporating microstructure in an explicit manner.

Given that the fatigue behaviour for metallic materials can also be estimated through the basic mechanical properties one might consider that even the toughness, an important design parameter, should be possible to formulate via elementary properties.

The fatigue crack growth rate model can be subjected to the following improvements:

- introducing Young's modulus as an input, which might resolve the overestimation of fatigue rates for nickel, titanium and aluminium alloys;

- as some of the applications to which a material is subjected are at high temperatures, introducing elevated temperature data to enable the calculation of fatigue rates in circumstances where creep and fatigue act together.

The accumulation of residual stress in ball bearings is the mechanism which produces failure in the bearing if the running conditions are optimum. Given that the running conditions for bearings are usually known the model created that enables the prediction of residual stress is of significant interest as potentially it might reduce the need for residual stress measurement techniques. It is limited to ball bearings however is easily usable and applicable, it captures the residual stress profiles at various depths below the raceway of bearings just from temperature, number of cycles and applied stresses.

The following steps could be followed to enhance the residual stress model

- lubrication and impurities size is one of the major factors influencing the usefulness of bearings, their introduction will transform the model into a more complex one;

- bearing geometry should be introduced as input parameter as this will enable the model to be applicable to other types of bearings such as taper and roller bearings;

- as residual stress is a factor influencing the life of bearings; it is not difficult with the right data to create a model which represents the influence of residual stress on the overall life of bearings.

Chapter 8

Appendix

8.1   Computer Programs

Program MAP STEEL HOT STRENGTH

Provenance of Source Code

R.C. Dimitriu,
Phase Transformations Group,
Department of Materials Science and Metallurgy,
University of Cambridge,
Cambridge, CB2 3QZ, U.K.
E-mail: rcd34@cam.ac.uk

The neural network program was produced by:

David MacKay,
Cavendish Laboratory,
University of Cambridge,
Madingley Road,
Cambridge, CB3 0HE, U.K.

Added to MAP: April 2007

Purpose

This package allows prediction of hot–strength of ferritic steels, as a function of chemical composition and heat treatment.

Specification

Language: FORTRAN / C
Product form: Source code / Executable files
Operating System: Linux & Windows 95/98/200/XP

Description

MAP NEURAL HOT STRENGTH contains a suite of programs which enable the user to estimate the hot–strength of ferritic steels as a function of chemical composition and heat treatment conditions. It makes use of a neural network program called *generate44*, which was developed by David MacKay and is part of the *bigback5* program. The network was trained using a large database of experimental results [141]. 20 different models are provided, which differ from each other by the number of hidden units and by the value of the seed used when training the network. It was found that a more accurate result could be obtained by averaging the results from all the models [129]. This suite of programs calculates the results of each model and then combines them, by averaging, to produce a committee result and error estimate, as described by MacKay (page 387 of reference [86]). The source code for the neural network program can be downloaded from David MacKay's website; the executable files only are available from MAP. Also provided are FORTRAN programs (as source code) for normalising the input data, averaging the results from the neural network program and unnormalising the final output file, along with other files necessary for running the program.

Programs are available which run on a Linux, and on a PC under Windows 95/98/2000/XP. A set of program and data files are provided for the model, which calculate the hot–strength of ferritic steels in NN. The files for unix and Linux are included in a directory called NN. This directory contains the following files and subdirectories:

README

A text file containing step-by-step instructions for running the program, including a list of input variables.

MINMAX
A text file containing the minimum and maximum limits of each input and output variable. This file is used to normalise and unnormalise the input and output data.

test.dat
An input text file containing the input variables used for predictions.

model.gen
This is a unix shell file containing the command steps required to run the module. It can be executed by typing csh model.gen at the command prompt. This shell file compiles and runs all the programs necessary for normalising the input data, executing the network for each model, unnormalising the output data and combining the results of each model to produce the final committee result.

HOT STR.exe
This executable program for the PC correspond to the unix command file model.gen.

no of lines.ex
This executable file reads the information of number of data from keyboard input and creates

no of rows.dat
file, this file is used by spec.ex/spec.exe to create spec.t1.

spec.ex/spec.exe
This executable file reads the information in no of rows.dat and creates a file called spec.t1.

spec.t1
A dynamic file, created by spec.ex/spec.exe, which contains information about the module and the number of data items being supplied. It is read by the program generate44/generate55.exe.

norm test.in
This is a text file which contains the normalised input variables. It is generated by the program normtest.for in subdirectory s.

generate44 / generate55
This is the executable file for the neural network program. generate44 runs on unix and generate55 on the PC. It reads the normalised input data file, norm test.in, and uses the weight files in subdirectory c. The results are written to the temporary output file out.

ot, out, res, sen
These files are created by generate44 and can be deleted.

Result
Contains the final un-normalised committee results for the predicted hot–strength.

SUBDIRECTORY s

no of lines.c
The source code for program no of lines.ex.

spec.c
The source code for program spec.ex.

normtest.for
Program to normalise the data in test.dat and produce the normalised input file norm test.in. It makes use of information read in from no of rows.dat and committee.dat.

gencom.for
This program uses the information in committee.dat and combines the predictions from the individual models, in subdirectory outprdt, to obtain an averaged value (committee prediction). The output (in normalised form) is written to com.dat.

treatout.for
Program to un-normalise the committee results in com.dat and write the output predictions to unnorm com. This file is then renamed Result.

committee.dat
A text file containing the number of models to be used to form the committee result and the number of input variables. It is read by gencom.for, normtest.for and treatout.for.

SUBDIRECTORY c

w*f
The weights files for the different models.

*.lu
Files containing information for calculating the size of the error bars for the different models.

c*
Files containing information about the perceived significance value for each model.

R*
Files containing values for the noise, test error and log predictive error for each model.

SUBDIRECTORY d

outran.x
A normalised output file which was created during the building of the model. It is accessed by generate44 via spec.t1.

SUBDIRECTORY outprdt

out1, out2 etc.
The normalised output files for each model.

com.dat
The normalised output file containing the committee results. It is generated by gencom.for.

Detailed instructions on the use of the program are given in the README files. Further information about this suite of programs can be obtained from reference [129].

Parameters

Input parameters
The input variables for the model are listed in the README file. The maximum and minimum values for each variable are given in the file MINMAX.

Output parameters
These program gives the hot–strength of ferritic steels. The correspond-

ing output files is called Result.dat or Result.

Error Indicators
None.

Accuracy
A full calculation of the error bars is presented in reference [129].

Further Comments
None.

Example

1. Program text
Complete program.

2. Program data
See sample data file: test.dat.

3. Program results
See sample output file: Result or Result.dat.

Auxiliary Routines
None

Keywords
neural network, hot–strength, creep

Program MAP STEEL-NI-TI-AL CRACK

Provenance of Source Code

R.C. Dimitriu,
Phase Transformations Group,
Department of Materials Science and Metallurgy,
University of Cambridge,
Cambridge, CB2 3QZ, U.K.
E-mail: rcd34@cam.ac.uk

The neural network program was produced by:

David MacKay,
Cavendish Laboratory,
University of Cambridge,
Madingley Road,
Cambridge, CB3 0HE, U.K.

Purpose

This package allows the estimation of fatigue crack growth rates of
steels, nickel, titanium and aluminium alloys as a function of material's
mechanical properties, test specimen size, loading mode, test conditions
and stress intensity factor.

Specification

Language: FORTRAN / C
Product form: Source code / Executable files
Operating System: Linux & Windows 95/98/200/XP

Description

MAP STEEL-NI-TI-AL CRACK contains a program that enables the
user to calculate the fatigue crack growth rates in steels, nickel, tita-
nium, aluminium alloys and any other metallic material having as in-
put parameters the material's mechanical properties (elongation, proof
stress and tensile strength), test specimen size, loading mode, test con-

ditions and stress intensity factor. It uses a neural network program called *generate44*, developed by David MacKay and is part of the *bigback5* program. The network was trained using a very large database of experimental results [195]. 100 networks were trained, which differ from each other by the number of hidden units (ranging from 1 to 20) and by the value of the seeds (ranging from 1 to 5) used when training the network. It was found that a more accurate result could be obtained by averaging the results from all the models [129]. The programs calculates the results for each model and then combines them, by averaging, to produce a committee result and error estimate, as described by MacKay (page 387 of reference [86]). The source code for the neural network program can be downloaded from David MacKay's website; the executable files only are available from MAP. Also provided are FORTRAN programs (as source code) for normalising the input data, averaging the results from the neural network program and unnormalising the final output file, along with other files necessary for running the program.

Programs are available which run on a Linux, and on a PC under Windows 95/98/2000/XP. A set of program and data files are provided for the model, which calculate the crack growth rates in metallic materials. The files for unix and Linux are included in a directory called NN. This directory contains the following files and subdirectories:

README
A text file containing step-by-step instructions for running the program, including a list of input variables.

MINMAX
A text file containing the minimum and maximum limits of each input and output variable. This file is used to normalise and unnormalise the input and output data.

test.dat
An input text file containing the input variables used for predictions.

model.gen
This is a unix shell file containing the command steps required to run the module. It can be executed by typing csh model.gen at the command prompt. This shell file compiles and runs all the programs necessary for normalising the input data, executing the network for each

model, unnormalising the output data and combining the results of each model to produce the final committee result.

model.exe
This executable program for the PC correspond to the unix command file model.gen.

no of lines.ex
This executable file reads the information of number of data from keyboard input and creates

no of rows.dat
file, this file is used by spec.ex/spec.exe to create spec.t1.

spec.ex/spec.exe
This executable file reads the information in no of rows.dat and creates a file called spec.t1.

spec.t1
A dynamic file, created by spec.ex/spec.exe, which contains information about the module and the number of data items being supplied. It is read by the program generate44/generate55.exe.

norm test.in
This is a text file which contains the normalised input variables. It is generated by the program normtest.for in subdirectory s.

generate44 / generate55
This is the executable file for the neural network program. generate44 runs on unix and generate55 on the PC. It reads the normalised input data file, norm test.in, and uses the weight files in subdirectory c. The results are written to the temporary output file out.

ot, out, res, sen
These files are created by generate44 and can be deleted.

Result
Contains the final un-normalised committee results for the predicted fatigue crack growth rates.

SUBDIRECTORY s

no of lines.c
The source code for program no of lines.ex.

spec.c
The source code for program spec.ex.

normtest.for
Program to normalise the data in test.dat and produce the normalised input file norm test.in. It makes use of information read in from no of rows.dat and committee.dat.

gencom.for
This program uses the information in committee.dat and combines the predictions from the individual models, in subdirectory outprdt, to obtain an averaged value (committee prediction). The output (in normalised form) is written to com.dat.

treatout.for
Program to un-normalise the committee results in com.dat and write the output predictions to unnorm com. This file is then renamed Result.

committee.dat
A text file containing the number of models to be used to form the committee result and the number of input variables. It is read by gencom.for, normtest.for and treatout.for.

SUBDIRECTORY c

w*f
The weights files for the different models.

*.lu
Files containing information for calculating the size of the error bars for the different models.

c*
Files containing information about the perceived significance value for each model.

R*
Files containing values for the noise, test error and log predictive error for each model.

SUBDIRECTORY d

outran.x
A normalised output file which was created during the building of the model. It is accessed by generate44 via spec.t1.

SUBDIRECTORY outprdt

out1, out2 etc.
The normalised output files for each model.

com.dat
The normalised output file containing the committee results. It is generated by gencom.for.

Detailed instructions on the use of the program are given in the README files. Further information about this suite of programs can be obtained from reference [129].

Parameters

Input parameters
The input variables for the model are listed in the README file. The maximum and minimum values for each variable are given in the file MINMAX.

Output parameters
These program gives the fatigue crack growth rates for steel and other metallic materials. The corresponding output files is called Result.dat or Result.

Error Indicators
None.

Accuracy
A full calculation of the error bars is presented in reference [129].

Further Comments
None.

Example

1. Program text

Complete program.

2. Program data
See sample data file: test.dat.

3. Program results
See sample output file: Result or Result.dat.

Auxiliary Routines
None

Keywords
neural network, fatigue crack growth rates, steel, nickel base superalloy, titanium alloy, aluminium alloy

Program MAP STEEL RES STRESS

Provenance of Source Code

R.C. Dimitriu,
Phase Transformations Group,
Department of Materials Science and Metallurgy,
University of Cambridge,
Cambridge, CB2 3QZ, U.K.
E-mail: rcd34@cam.ac.uk

The neural network program was produced by:

David MacKay,
Cavendish Laboratory,
University of Cambridge,
Madingley Road,
Cambridge, CB3 0HE, U.K.

Purpose

This package allows the calculation of the residual stress developed during rolling within a steel bearing as a function of temperature, revolutions, contact pressure, hoop stress and the depth within the bearing were the residual stress develops.

Specification

Language: FORTRAN / C
Product form: Source code / Executable files
Operating System: Linux & Windows 95/98/200/XP

Description

MAP STEEL RES STRESS contains a program that facilitates the calculation of residual stress development within the raceway of a ball–bearing as a function of temperature, revolutions, contact pressure, hoop stress and depth. It uses a neural network program called *generate44*, developed by David MacKay and is part of the *bigback5* program.

The network was trained using a large database of experimental results provided by SKF the ball–bearing manufacturing company. 100 networks were trained, with one to twenty hidden units and one to five seeds. It was found that a more accurate result could be obtained by averaging the results from all the models [129]. The programs calculates the results for each model and then combines them, by averaging, to produce a committee result and error estimate, as described by MacKay (page 387 of reference [86]). The source code for the neural network program can be downloaded from David MacKay's website; the executable files only are available from MAP. Also provided are FORTRAN programs (as source code) for normalising the input data, averaging the results from the neural network program and unnormalising the final output file, along with other files necessary for running the program.

Programs are available which run on a Linux, and on a PC under Windows 95/98/2000/XP. A set of program and data files are provided for the model, which calculate the residual stress development in bearings. The files for unix and Linux are included in a directory called NN. This directory contains the following files and subdirectories:

README
A text file containing step-by-step instructions for running the program, including a list of input variables.

MINMAX
A text file containing the minimum and maximum limits of each input and output variable. This file is used to normalise and unnormalise the input and output data.

test.dat
An input text file containing the input variables used for predictions.

model.gen
This is a unix shell file containing the command steps required to run the module. It can be executed by typing csh model.gen at the command prompt. This shell file compiles and runs all the programs necessary for normalising the input data, executing the network for each model, unnormalising the output data and combining the results of each model to produce the final committee result.

model.exe
This executable program for the PC correspond to the unix command file model.gen.

no of lines.ex
This executable file reads the information of number of data from keyboard input and creates

no of rows.dat
file, this file is used by spec.ex/spec.exe to create spec.t1.

spec.ex/spec.exe
This executable file reads the information in no of rows.dat and creates a file called spec.t1.

spec.t1
A dynamic file, created by spec.ex/spec.exe, which contains information about the module and the number of data items being supplied. It is read by the program generate44/generate55.exe.

norm test.in
This is a text file which contains the normalised input variables. It is generated by the program normtest.for in subdirectory s.

generate44 / generate55
This is the executable file for the neural network program. generate44 runs on unix and generate55 on the PC. It reads the normalised input data file, norm test.in, and uses the weight files in subdirectory c. The results are written to the temporary output file out.

ot, out, res, sen
These files are created by generate44 and can be deleted.

Result
Contains the final un-normalised committee results for the predicted residual stress.

SUBDIRECTORY s

no of lines.c
The source code for program no of lines.ex.

spec.c
The source code for program spec.ex.

normtest.for
Program to normalise the data in test.dat and produce the normalised input file norm test.in. It makes use of information read in from no of rows.dat and committee.dat.

gencom.for
This program uses the information in committee.dat and combines the predictions from the individual models, in subdirectory outprdt, to obtain an averaged value (committee prediction). The output (in normalised form) is written to com.dat.

treatout.for
Program to un-normalise the committee results in com.dat and write the output predictions to unnorm com. This file is then renamed Result.

committee.dat
A text file containing the number of models to be used to form the committee result and the number of input variables. It is read by gencom.for, normtest.for and treatout.for.

SUBDIRECTORY c

w*f
The weights files for the different models.

*.lu
Files containing information for calculating the size of the error bars for the different models.

c*
Files containing information about the perceived significance value for each model.

R*
Files containing values for the noise, test error and log predictive error for each model.

SUBDIRECTORY d

outran.x
A normalised output file which was created during the building of the model. It is accessed by generate44 via spec.t1.

SUBDIRECTORY outprdt

out1, out2 etc.
The normalised output files for each model.

com.dat
The normalised output file containing the committee results. It is generated by gencom.for.

Detailed instructions on the use of the program are given in the README files. Further information about this suite of programs can be obtained from reference [129].

Parameters

Input parameters
The input variables for the model are listed in the README file. The maximum and minimum values for each variable are given in the file MINMAX.

Output parameters
These program gives the residual stress for ball bearings. The corresponding output files is called Result.dat or Result.

Error Indicators
None.

Accuracy
A full calculation of the error bars is presented in reference [129].

Further Comments
None.

Example

1. Program text
Complete program.

2. Program data

See sample data file: test.dat.

3. Program results
See sample output file: Result or Result.dat.

Auxiliary Routines
None

Keywords
neural network, residual stress, steel, bearings

Bibliography

[1] T. Sourmail, H. K. D. H. Bhadeshia, and D. J. C. MacKay. Neural network model of creep strength of austenitic stainles steel. *Materials Science and Technology*, 18:655–663, 2002.

[2] M. A. Yescas-Gonzalez. *Modelling the Microstructure and Mechanical Properties of Austempered Ductile Irons*. PhD thesis, University of Cambridge, 2001.

[3] J. D. Robson and H. K. D. H. Bhadeshia. Modelling precipitation sequences in power plant steels: Part I, kinetic theory. *Materials Science and Technology*, 13:631–639, 1997.

[4] J. D. Robson and H. K. D. H. Bhadeshia. Modelling precipitation sequences in power plant steels, Part II: Application of kinetic theory. *Materials Science and Technology*, 28A:640–644, 1997.

[5] R. Kemp, G. A. Cottrell, H. K. D. H. Bhadeshia, G. R. Odette, T. Yamamoto, and H. Kishimoto. Neural-network analysis of irradiation hardening in low-activation steels. *Journal of Nuclear Materials*, 348(3):311–328, 2006.

[6] T. Sourmail and C. Garcia-Mateo. A model for predicting the martensite–start temperature of steels. *Computational Materials Science*, 34:213–218, 2005.

[7] J. H. Pak, J. H. Jang, H. K. D. H. Bhadeshia, and L. Karlsson. Optimization of neural network for charpy toughness of steel welds. *Materials and Manufacturing Processes*, 24:16–21, 2009.

[8] H. K. D. H. Bhadeshia. Modelling of recrystallisation in mechanically alloyed materials. *Materials Science and Engineering A*, A223:91–98, 1997.

[9] H. E. Boyer, editor. *Metals Handbook*, volume 19, Fatigue and Fracture, chapter Contact Fatigue, pages 331–336,361–362. American Society for Metals, December 1996.

[10] H. E. Boyer, editor. *Metals Handbook*, volume 11, Failure Analysis and Prevention, chapter Failures of Rolling-Element Bearings, pages 500–505. American Society for Metals, November 2002.

[11] R. I. Stephens, A. Fatemi, R. R. Stephens, and H. O. Fuchs. *Metal Fatigue in Engineering*. Jon Wiley & Sons, New York, 2nd Edition, 2000.

[12] A. Wohler. *Uber die Festigkeitsversuche mit Eisen and Stahl*. Zeitschrift fur Bauwesen, 1870.

[13] J. F. Knott. *Fundamentals of Fracture Mechanics*, chapter 9, pages 234–241. Butterworths, London, 1981.

[14] ASTM E 647-95a. Standard test method for measurement of fatigue crack growth rates. *Annual Book of ASTM Standards*, 13:577–613, 1999.

[15] E. T. Wessel. Practical fracture mechanics for structural steel. *Symposium on Fracture Toughness Concepts for Weldable Structural Steel, Risley, England*, pages H1–44, 1969.

[16] A. R. Jack and A. T. Price. The initiation of fatigue cracks from notches in mild steel plates. *Int. J. Fract. Mech*, 6:401–409, 1970.

[17] R. O. Ritchie, G. G. Garrett, and J. F. Knott. Crack growth monitoring: Optimisation of the electrical potential technique using an analogue method. *Int. J. Fract. Mech.*, 7:462–467, 1971.

[18] R. B. Waterhouse. Fretting fatigue. *International Materials Reviews*, 37:77–79, 1992.

[19] Y. Murakami, T. Nomoto, and T. Ueda. Factors influencing the mechanism of superlong fatigue failure in steels. *Fatigue and Fracture of Engineering Materials and Structures*, 22:581–590, 1999.

[20] K. Sugino, K. Miyamoto, and M. Nagumo. Failure process of bearing steel in rolling contact fatigue. *Trans. of the Iron and Steel Institute of Japan*, 11:9–18, 1971.

[21] P. C. Paris and F. Erdogan. A critical analysis of crack propagation laws. *J. Basic Eng.(Trans. ASME)*, 85:528–534, 1963.

[22] P. C. Paris. The fracture mechanics approach to fatigue. *Proc. 10th Sagamore Conf.*, pages 107–132, 1965.

[23] W. E. Poole, R. L. Thom, E. V. Zarestky, W. J. Anderson, and E. N. Bamberger. *Materials and Processing*. E.V. Zarestky, 1992.

[24] S. Cretu and N. G. Popinceanu. The influence of residual stress induced by plastic deformation on rolling contact fatigue. *Wear*, 105:153–170, 1985.

[25] M. Kaneta, H. Yatsuzuka, and Y. Murakami. Mechanism of crack growth in lubricated rolling / sliding contact. *Trans. Am. Soc. Lubr. Eng.*, 28(3):407–414, 1985.

[26] R. F. Cook and G. M. Pharr. Direct observation and analysis of indentation cracking in glasses and ceramics. *J. Am. Ceram. Soc.*, 73:787–817, 1990.

[27] J. P. Withers and H. K. D. H. Bhadeshia. Residual stress. Part II - nature and origins. *Mater. Sci. Tech. Ser.*, 17:366–375, 2001.

[28] J. P. Withers and H. K. D. H. Bhadeshia. Residual stress. Part I - measurement techniques. *Mater. Sci. Tech. Ser.*, 17:355–365, 2001.

[29] H. Muro, T. Tsushima, and M. Nagafuchi. Initiation and propagation of surface cracks in rolling fatigue of high hardness steel. *Wear*, 35:261–282, 1975.

[30] J. J. Bush, W. L. Grube, and G. H. Robinson. Microstructural and residual stress changes in hardened steel due to rolling contact. *ASM Trans.*, 54:390–412, 1961.

[31] D. P. Koistinen and R. E. Marburger. Internal stresses induced by phase transformations. *Michigan 1985 Conference Proceedings on Internal Stresses*, pages 110–119, 1958.

[32] E. D. Doyle, Y. C. Wong, and M. I. Ripley. Residual stress evaluation in martensitic stainless steel as a function of gas quenching pressure using thermal neutrons. *Physica B*, 865:897–899, 2006.

[33] Y. B. Guo and M. E. Barkey. Modelling of rolling contact fatigue for hard machined components with process-induced residual stress. *Int. J. Fatigue*, 26:605–613, 2003.

[34] M. Liu, J. Takagi, and A. Tsukuda. Effect of tool nose radius and tool wear on residual stress distribution in hard turning of bearing steel. *J. Mater. Process. Tech.*, 150(3):234–241, 2004.

[35] T. G. Dawson and T. R. Kurfess. Quantification of tool wear using white light interferometry and three-dimensional computational metrology. *Int. J. Mach. Tool Manu.*, 45:591–596, 2005.

[36] Y. Matsumoto, D. Madaga, D. W. Hoeppner, and T. Y. Kim. Effect of machinning process on the fatigue strength of hardened AISI 4340 steel. *J. Eng. Ind.-T ASME*, 113:154–159, 1991.

[37] S. O. A. El-Helieby and G. W. Rowe. A quantitative comparison between residual stress and fracture properties of surface-ground bearing steel. *Wear*, 58:155–172, 1979.

[38] T. Barski. Rolling contact fatigue and the significance of residual stresses. *Proceedings of the Fourth International Conference on Residual Stresses*, pages 913–915, 1994.

[39] Y. Matsumoto, M. M. Barash, and C. R. Liu. Effect of hardness on the surface integrity of ansi 4340 steel. *J. Eng. Ind.-T ASME*, 3:299–305, 1986.

[40] A. M. Abrao and D. K. Aspinwall. The surface integrity of turned and ground hardened bearing steels. *Wear*, 196:279–284, 1996.

[41] D. W. Reichard, R. J. Parker, and E. V. Zaretsky. Residual stress and subsurface hardness changes induced during rolling contact. *NASA Technical Note D-4456*, 1968.

[42] R. L. Scott, R. K. Kepple, and M. H. Miller. The effect of processing induced near-surface residual stress on ball bearing fatigue. *Rolling Contact Phenomena*, 1:301–316, 1962.

[43] A. J. Gentle, A. F. Jordan, and A. D. Martin. Phase transformations in high-carbon, high-hardness steels under contact loads. *J. Eng. Ind.-T ASME*, 36:329–340, 1965.

[44] E. Kula and V. Weiss. *Residual Stress and Relaxation.* Plenum Pub. Corp., 1982.

[45] E. Orowan. Causes and effects of internal stresses. *Michigan 1985 Conference Proceedings on Internal Stresses*, pages 59–80, 1958.

[46] C. O. Ruud. *Residual Stress Measurements.* Klihn Dana and Madlin Dana, 2000.

[47] J. Lu. *Handbook of Mesurement of Residual Stress.* Fairmont Printing & Publishing, 1996.

[48] B. Duncan. *The economic history of steel making 1867–1939.* Cambridge University Press, U.K., 1961.

[49] J. Gernet. *A History of Chinese Civilisation.* Cambridge University Press, U.K., 1982.

[50] H. Hasegaw. *The Steel Industry in Japan.* Routledge, New York, 1996.

[51] W. Hume-Rothery, R. W. Smallman, and C. W. Haworth. *The Structure of Metals and Alloys.* The Institute of Metals, London.

[52] W. Hume-Rothery. *Atomic Theory for Students of Metallurgy.* The Institute of Metals, London.

[53] T. B. Massalski. *Physical Metallurgy, 2nd ed.* North-Holland, 1970.

[54] W. C. Leslie. *The Physical Metallurgy of Steels.* McGraw Hill, New York, 1982.

[55] R. W. K. Honeycombe and H. K. D. H. Bhadeshia. *Steels: Microstructure and Properties, 2nd edition.* Butterworths–Hienemann, London, 1995.

[56] G. Beranger, G. Henry, and G. Sanz. *The Book of Steel.* Intercept Limited, U.K., 1996.

[57] F. B. Pickering. *Physical Metallurgy and the Design of Steels.* Applied Science Publishers, Essex, U. K., 1978.

[58] O. H. Whyatt and D. D. Hughs. *Metals Ceramics and Polymers.* Cambridge University Press, Cambride.

[59] W. D. Callister Jr. *Materials Science and Engineering: An Introduction.* Wiley, 2003.

[60] http://www.doitpoms.ac.uk/tlplib/dislocations/index.php.

[61] D. Hull and D. J. Bacon. *Introduction to Dislocations 4th Ed.* Butterworth Heinemann, 2001.

[62] E. O. Hall. *Proc. Phys Soc. Series B*, 64:747–761, 1951.

[63] N. J. Petch. *J. Iron Steel Inst.*, 174:25–39, 1953.

[64] R. E. Reed-Hill. *Physical Metallurgy Principles.* PWS-Kent Publishing Company, Boston.

[65] F. G. Caballero, H. K. D. H. Bhadeshia, K. J. A. Mawella, D. G. Jones, and P. Brown. Design of novel high–strength bainitic steels: Part I. *Materials Science and Technology*, 17:512–516, 2001.

[66] F. G. Caballero, H. K. D. H. Bhadeshia, K. J. A. Mawella, D. G. Jones, and P. Brown. Design of novel high–strength bainitic steels: Part II. *Materials Science and Technology*, 17:517–522, 2001.

[67] S. Downey, P.N. Kalua, and K. Hanb. The effect of heat treatment on the microstructure stability of modified 316LN stainless steel. *Materials Science and Engineering: A*, 480(1–2):96–100, 2008.

[68] S. Berbenni, V. Faviera, and M. Berveillera. Micromacro modelling of the effects of the grain size distribution on the plastic flow stress of heterogeneous materials. *Computational Materials Science*, 39(1):96–105, 2007.

[69] R. Song, D. Ponge, D. Raabe, J. G. Speer, and D. K. Matlock. Overview of processing, microstructure and mechanical properties of ultrafine grained BCC steels. *Materials Science and Engineering: A*, 441(1–2):1–17, 2006.

[70] F. Louchet, J. Weiss, and T. Richeton. Hall-Petch law revisited in terms of collective dislocation dynamics. *Physical Review Letters*, 97(7):075504(1–4), 2006.

[71] V. Bata and E. V. Pereloma. An alternative physical explanation of the HallPetch relation. *Acta Materialia*, 52(3):657–665, 2004.

[72] J. A. Wert. Comments on "An alternative physical explanation of the Hall–Petch relation". *Acta Materialia*, 50(12):1487–1490, 2004.

[73] L. M. Brown and R. K. Ham. *Strengthening Methods in Crystals.* Elsevier.

[74] J. W. Christian. *The Theory of Transformations in Metals and Alloys.* Pergamon Press, 1965.

[75] G. E. Dieter. *Mechanical Metallurgy.* McGraw–Hill Book Company, 1988.

[76] E. N. da C Andrade and B. Chalmers. *Proc. R Soc. London*, 138A:348–353, 1923.

[77] R. L. Fulmann, R. P. Carreker, and J. C. Fisher. *Trans AIME*, 197:657–659, 1953.

[78] E. N. da C Andrade. *Proc. R Soc. London*, 90A:329–342, 1914.

[79] S. Chatterjee, M. Murugananth, and H. K. D. H. Bhadeshia. δ–TRIP steel. *Materials Science and Technology*, 23:in press, 2007.

[80] E. Keehan, L. Karlsson, and H.-O. Andrén. Influence of C, Mn and Ni on strong steel weld metals: Part I, effect of nickel. *Science and Technology of Welding and Joining*, 11:1–8, 2006.

[81] E. Keehan, L. Karlsson, H.-O. Andrén, and H. K. D. H. Bhadeshia. Influence of C, Mn and Ni on strong steel weld metals: Part II, increased impact toughness. *Science and Technology of Welding and Joining*, 11:9–18, 2006.

[82] E. Keehan, L. Karlsson, H.-O. Andrén, and H. K. D. H. Bhadeshia. Influence of C, Mn and Ni on strong steel weld metals: Part III, increased strength. *Science and Technology of Welding and Joining*, 11:19–24., 2006.

[83] M. Murugananth and H. K. D. H. Bhadeshia. Components of the creep strength of welds. In H. Cerjak and H. K. D. H. Bhadeshia, editors, *Mathematical Modelling of Weld Phenomena 6*, pages 243–260, London, U.K., 2002. Maney.

[84] D. Cole, C. Martin-Moran, A. Sheard, H. K. D. H. Bhadeshia, and D. J. C. MacKay. Modelling the creep rupture strength of ferritic steel welds. *Science and Technology of Welding and Joining*, 5:81–89, 2000.

[85] H. K. D. H. Bhadeshia. Neural networks in materials science. *ISIJ International*, 39:966–979, 1999.

[86] D. J. C. MacKay. *Mathematical Modelling of Weld Phenomena 3*. The Institute of Materials, 1997.

[87] H. Fuji, D. J. C. MacKay, and H. K. D. H. Bhadeshia. Bayesian neural network analysis of fatigue crack growth rate in nickel base superalloys. *ISIJ Int.*, 36(11):1373–1382, 1996.

[88] R. C. Eberhart and R. W. Dobbins. *Neural Network PC Tools A practical guide*. Academic Press Inc., 1990.

[89] D. J. C. MacKay. *Neural Computation 2*. The Institute of Materials, 1992.

[90] R. Kemp. *Alloy Design for a Fusion Power Plant*. PhD thesis, University of Cambridge, 2006.

[91] A. D. Luca, L. M. Ricciard, and R. Vasudeva. Random matrices and neural networks. *Kybernetik*, 6(5):163–170, 1970.

[92] D. J. C. MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.

[93] C. M. Bishop. *Neural Networks for Pattern Recognition*. Claerdon Press Oxford, 1995.

[94] R. C. Dimitriu. Unpublished work.

[95] M. A. Yescas, H. K. D. H. Bhadeshia, and D. J. C. MacKay. Estimation of the amount of retained austenite in austempered ductile irons using neural networks. *Mater. Sci. Eng.*, 2000.

[96] Y. Shen. *The Rate of Creep Deformation*. PhD thesis, University of Cambridge, 2003.

[97] F. Masuyama and H. K. D. H. Bhadeshia. Creep strength of high ferritic steels designed using neural networks and phase stability calculations. *Fifth International Conference on Advances in Materials Technology for Fossil Power Plants*, pages 4B–01, 2000.

[98] J. R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection.* MIT Press, 1992.

[99] J. R. Koza. *Genetic Programming 2: Automatic Discovery of Reusable Programs.* MIT Press, 1994.

[100] T. Brajlih, I. Drstvensek, and B. Valentan. Improving the accuracy of rapid prototyping procedures by genetic programming. *5th International Conference of the Danube–Adria Association for Automation Manufacturing*, pages 113–116, 2006.

[101] Y. Shimizu, K. Tsuji, and M. Nomura. Optimal disassembly sequence generation using a genetic programming. *International Journal of Production Research*, 45:4537–4554, 2007.

[102] U. Beldek and K. Leblebicioglu. Strategy creation, decomposition and distribution in particle navigation. *Information Sciences*, 177(3):755–770, 2007.

[103] J. K. Estrada-Gil, J. C. Fernandez-Lopez, and E. Hernandez-Lemus. A genetic programming decision tree induction method to find epistatic effects in common complex diseases. *Bioinformatics*, 23(13):I167–I174, 2007.

[104] P. A. Whigham and P. F. Crapper. Modelling rainfall-runoff using genetic programming. *Mathematical and Computer Modelling*, 33(6–7):707–721, 2001.

[105] D. Agnelli, A. Bollini, and L. Lombardi. Image classification: An evolutionary approach. *Pattern Recognition Letters*, 23(1–3):303–309, 2002.

[106] L. Gritz and J. K. Hahn. Genetic programming for articulated figure motion. *Journal of Visualisation and Computer Animation*, 6(3):129–142, 1995.

[107] N. Chakraborti. Genetic algorithms in materials design and processing. *International Materials Reviews*, 49(3-4):246–260, 2004.

[108] W. Banzhaf, P. Nordin, R. E. Keller, and F. D. Francone. *Genetic Programming - An Introduction; On the Automatic Evolution of Computer Programs and its Applications.* Morgan Kaufmann, 1998.

[109] D. L. Nash and G. W. Rogers. Herd sire portfolio selection: A comparison of rounded linear and integer programming. *Journal of Dairy Science*, 78(11):2486–2495, 1995.

[110] W. Loonen, P. S. Heuberger, A. H. Bakema, and P. Schot. Application of a genetic algorithm to minimise agricultural nitrogen deposition in nature reserves. *Agricultural Systems*, 88(2–3):360–375, 2006.

[111] A. A. Motsinger-Reif, S. M. Dudek, L. W. Hahn, and M. D. Ritchie. Comparison of approaches for machine-learning optimisation of neural networks for detecting gene-gene interactions in genetic epidemiology. *Gentic Epidemiology*, 32(4):325–340, 2008.

[112] C. Fillon and A. Bartoli. A divide and conquer strategy for improving efficiency and probability of success in genetic programming. *Proceedings of the 9th European Conference on Genetic Programming*, pages 13–23, 2006.

[113] T. Sawa. Exact bias and mean square error of K-class estimators. *Econometrica*, 38(4):6–13, 1970.

[114] C. C. Kilgus and W. C. Gore. Root mean square error in encoded digital telemetry. *IEE Transactions on Communications*, 20(3):315–320, 1972.

[115] C. Laquerbe, P. Floquet, S. Domenech, and L. Pibouleau. Development of separation sequences using a genetic algorithm. *Rairo-Recherche Operationnelle-Operations Research*, 31(4):375–397, 1997.

[116] B. Naudts and L. Kallel. A comparison of predictive measures of problem difficulty in evolutionary algorithms. *Evolutionary Computation, IEE Transactions*, 4(1):1–15, 2000.

[117] J. P. Yang and C. K. Soh. Structural optimisation by genetic algorithms with tournament selection. *Journal of Computing in Civil Engineering*, 11(3):195–200, 1997.

[118] S. M. Bhandarkar, Y. Q. Zhang, and W. D. Potter. An edge-detection technique using genetic algorithm-based optimisation. *Pattern Recognition*, 27(99):1159–1180, 1994.

[119] K. Chellapilla. Evolving computer programs without sub-tree crossover. *Evolutionary Computation, IEE Transactions*, 1(3):209–216, 1997.

[120] H. Majeed and C. Ryan. A less destructive context-aware crossover operator for GP. *Proceedings of the 9th European Conference on Genetic Programming*, pages 36–48, 2006.

[121] M. E. Lopez-Medina, S. Vazquez-Montiel, and J. Herrera-Vazquez. Optimisation of optical systems using genetic algorithms: a comparison among different implementations of the algorithm. *Conference Proceedings RIAO/OPTILAS 2007*, pages 642–647, 2008,.

[122] J. Couchet, D. Manrique, J. Rios, and A. Rodriguez-Paton. Crossover and mutation operators for grammar–guided genetic programming. *Soft Computing*, 11(10):943–955, 2007.

[123] L. M. Schmitt and S. Droste. Convergence to global optima for genetic programming systems with dynamically scaled operators. *Genetic and Evolutionary Computation Conference; GECCO 2006*, pages 879–886, 2006.

[124] F. Herrera, M. Lozano, and L. J. Verdegay. Dynamic and heuristic fuzzy connectives-based crossover operators for controlling the diversity and convergence of real–coded genetic algorithms. *International Journal of Intelligent Systems*, 11(12):1013–1040, 1996.

[125] H. K. D. H. Bhadeshia, R. C. Dimitriu, S. Forsik, J. H. Pak, and J. H. Ryu. On performance of neural networks in materials science. *Materials Science and Technology*, 2008.

[126] K. J. Irvine, T. Gladman, and F. B. Pickering. The strength of austenitic stainless steels. *J. of the Iron and Steel Inst.*, 207:1017–1028, 1969.

[127] Y. Tomita and K. Okabayashi. *Metall. Trans. A*, 18A:115–129, 1987.

[128] C. Y. Young and H. K. D. H. Bhadeshia. Strength of mixtures of bainite and martensite. *Material Science and Technology*, 10:209–214, 1994.

[129] R. C. Dimitriu and H. K. D. H. Bhadeshia. Hot–strength of creep–resistant ferritic steels and relationship to creep–rupture data. *Materials Science and Technology*, 23:1127–1131, 2007.

[130] S. H. Nam and S. Y. Oh. Real-time dynamic visual tracking using PSD sensors and extended trapezoidal motion planning. *Applied Intelligence*, 10(1):53–70, 1999.

[131] T. K. Meng and C. Butler. Solving multiple response optimisation problems using adaptive neural networks. *International Journal of Advanced Manufacturing Technology*, 13(9):666–675, 1997.

[132] H. K. D. H. Bhadeshia, D. J. C. MacKay, and L.-E. Svensson. The impact toughness of C–Mn steel arc–welds – a Bayesian neural network analysis. *Materials Science and Technology*, 11:1046–1051, 1995.

[133] T. Cool, H. K. D. H. Bhadeshia, and D. J. C. MacKay. The yield and ultimate tensile strength of steel weld metals. *Materials Science and Engineering A*, 223A:186–200, 1997.

[134] B. Chan, M. Bibby, and N. Holtz. Predicting 800 to 500 degrees C weld cooling times by using backpropagation neural networks. *Transactions of the Canadian Society for Mechanical Engineering*, 20(1):75–85, 1996.

[135] J. M. Schooling, M. Brown, and P. A. S. Reed. An example of the use of neural computing techniques in materials science - the modelling of fatigue thresholds in Ni-base superalloys. *Materials Science and Engineering A*, 260(1–2):222–239, 1999.

[136] V. Narayan, R. Abad, B. Lopez, H. K. D. H. Bhadeshia, and D. J. C. MacKay. Estimation of hot torsion stress strain curves in iron alloys using a neural network analysis. *ISIJ International*, 39:999–1005, 1999.

[137] P. D. Hodgson, L. X. Kong, and C. H. J. Davies. The prediction of the hot strength in steels with an integrated phenomenological and artificial neural network model. *Journal of Materials Processing Technology*, 87:131–138, 1999.

[138] L. X. Kong and P. D. Hodgson. Predicting flow strength of austenitic steels with an IPANN model using different training strategies. *Advances in Engineering Software*, 31:945–954, 2000.

[139] L. X. Kong, P. D. Hodgson, and D. C. Collinson. Extrapolative prediction of the hot strength of austenitic steels with a combined constitutive and ANN model. *Journal of Materials Processing Technology*, 102:84–89, 2000.

[140] Z. Sterjovski, D. Nolan, K.R. Carpenter, D.P. Dunne, and J. Norrish. Artificial neural networks for modelling the mechanical properties of steels in various applications. *Journal of Materials Processing Technology*, 170:536–544, 2005.

[141] Y. Kojchi, I. Hiroshi, T. Hideo, Y. Masayoshi, K. Osamu, K. Kiyoshi, and K. Kazuhio. NRIM data sheets 1b, 3b, 8b, 11b, 12b, 17b, 18b, 19b, 20b, 21b. Technical report, National Research Institute for Metals, Tokyo, Japan, 1994.

[142] D. J. C. MacKay. Bayesian interpolation. *Neural Computation*, 4:415–447, 1992.

[143] D. J. C. MacKay. Practical Bayesian framework of backpropagation networks. *Neural Computation*, 4:448–472, 1992.

[144] J. H. Pak. *Modelling the Impact Toughness of Weld Metals and the Formation of Coalesced Bainite*. PhD thesis, Pohang University of Science and Technology, 2008.

[145] MTDATA. Software, National Physical Laboratory, Teddington, U.K., 2006.

[146] A. A. B. Sugden and H. K. D. H. Bhadeshia. A model for the strength of the as-deposited regions of steel weld metals. *Metallurgical Transactions A*, 19A:1597–1602, 1988.

[147] D. Kalish, S. A. Kulin, and M. Cohen. Bainitic structures and thermomechanical treatments applied to steel. *Journal of Metals*, 17:157–164, 1965.

[148] H. K. D. H. Bhadeshia and A. R. Waugh. Bainite: An atom probe study of the incomplete reaction phenomenon. *Acta Metallurgica*, 30:775–784, 1982.

[149] M. Peet, S. S. Babu, M. K. Miller, and H. K. D. H. Bhadeshia. Three–dimensional atom probe analysis of carbon distribution in low–temperature bainite. *Scripta Materialia*, 50:1277–1281, 2004.

[150] F. G. Caballero, M. K. Miller, S. S. Babu, and C. Garcia-Mateo. Atomic scale observations of bainite transformation in a high carbon high silicon steel. *Acta Materialia*, 55:381–390, 2007.

[151] W. C. Leslie. Iron and its dilute substitutional solid solutions. *Metallurgical Transactions A*, 3:5–23, 1972.

[152] N. Komai, F. Masuyama, I. Ishihara, T. Yokoyama, Y. Yamadera, H. Okada, K. Miyata, and Y. Sawaragi. Development and application of 2.25Cr-1.6W (HCM2S steel large diameter and thick section pipe. In *Advanced Heat Resistant Steels for Power Generation*, London, 1999. IOM Communications.

[153] M. Kovacic, P. Uratnik, M. Berzocnik, and R. Turk. Prediction of the bending capability of rolled metal sheet by genetic programming. *Materials and Manufacturing Processes*, 22:634–640, 2007.

[154] M. Kovacic, M. Berzocnik, and R. Turk. Modelling the hot yield stress curves for carbon silicon steel by genetic programming. *Materials and Manufacturing Processes*, 20:543–551, 2005.

[155] D. J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.

[156] F. Lefebvre and I. Sinclair. Micromechanical aspects of fatigue in a MIG welded aluminium airframe alloy Part II. Short fatigue crack behaviour. *Materials Science and Engineering A*, 407(1–2):265–272, 2005.

[157] P. S. Song and Y. L. Shieh. Stop drilling procedure for fatigue life improvement. *International Journal of Fatigue*, 26(12):1333–1339, 2004.

[158] E. Kerscher, K. H. Langa, and D. Lohea. Increasing the fatigue limit of a high-strength bearing steel by thermomechanical treatment. *Materials Science and Engineering A*, 483–484:415–417, 2008.

[159] P. Lazzarina, P. Livierib, F. Bertoa, and M. Zappalortoa. Local strain energy density and fatigue strength of welded joints under uniaxial and multiaxial loading. *Engineering Fracture Mechanics*, 75(7):1875–1889, 2008.

[160] K. J. Miller. The behaviour of short fatigue cracks and their initiation. Part II-A general summary. *Fatigue Fract. Engng. Mater. Struct*, 10:93–113, 1987.

[161] T. Ghidini and C. D. Donne. Fatigue crack propagation assessment based on residual stresses obtained through cut-compliance technique. *Fatigue & Fracture of Engineering Materials & Structures*, 30(3):214–222, 2007.

[162] C. X. Li, Y. Sunb, and T. Bella. Factors influencing fretting fatigue properties of plasma-nitrided low alloy steel. *Materials Science and Engineering A*, 292(1):18–25, 2000.

[163] W. Geary and J. E. King. Residual stress effects during near threshold fatigue crack growth. *Int. J. Fatigue*, 9(1):11–16, 1987.

[164] K. Makhlouf and J. W. Jones. Near-threshold fatigue crack growth behaviour of a ferritic stainless steel at elevated temperatures. *Int. J. Fatigue*, 14(2):97–104, 1992.

[165] H. K. Yoon, S. P. Leea, B. H. Mina, S. W. Kimb, Y. Katohc, and A. Kohyamac. Fatigue life and fatigue crack propagation behaviour of JLF-1 steel. *Fusion Engineering and Design*, 61–62:677–682, 2002.

[166] Y. Mutoh, S. J. Zhu, T. Hansson, S. Kurai, and Y. Mizuhara. Effect of microstructure on fatigue crack growth in TiAl intermetalics at elevated temperatures. *Materials Science & Engineering A*, A323:62–69, 2002.

[167] K. Makhlouf, H. Sidhom, I. Triguia, and C. Braham. Corrosion fatigue crack propagation of a duplex stainless steel X6 CrNiMoCu25-6 in air and in artificial sea water. *Int. J. Fatigue*, 25(2):167–179, 2003.

[168] H. Masuda, S. Matsuoka, S. Nishijima, and M. Shimodaria. The role of corrosion in fatigue crack propagation for structural alloys in 3% NaCl solution. *Corrosion Science*, 28(5):433–447, 1988.

[169] V. Schroedera, C. J. Gilberta, and R. O. Ritchie. A comparison of the mechanisms of fatigue-crack propagation behaviour in a Zr-based bulk amorphous metal in air and an aqueous chloride solution. *Materials Science and Engineering A*, 317(1–2):145–152, 2001.

[170] M. A. Bassidi, J. Masounave, and J.I. Dickson. *Proceeding of the Conference on Duplex Stainless Steels. St. Louis, 2528 October*, pages 445–453, 1982.

[171] A. T. Stewart. The influence of environment and stress ratio on fatigue crack growth at near threshold stress intensities in low alloy steel. *Engng. Fracture Mech.*, 13:463–478, 1980.

[172] J. D. Atkinson and J. E. Forrest. Factors influencing the rate of growth of fatigue cracks in RVP steels exposed to a simulated PWR primary water environment. *Corrosion Science*, 25(8–9):607–631, 1985.

[173] T. Yokobori and K. Sato. The effect of frequency on fatigue crack propagation data and striation spacing in 2024-T3 aluminium alloy and in SM-50 steel. *Engng. Fracture Mech.*, 8:81–95, 1976.

[174] K. Nakasa and H. Takai. Effect of temperature on the frequency dependency of crack propagation velocity in delayed failure under cyclic load. *Engng. Fracture Mech.*, 10:783–793, 1978.

[175] A. Ohta and E. Sasaki. Influence of stress ratio on the threshold level for fatigue crack propagation in high strength steels. *Engng. Fracture Mech.*, 9(2):307–315, 1977.

[176] W. O. Soboyejo and J. F. Knott. Effects of stress ratio on fatigue crack propagation in Q1N (HY80) pressure vessel steel. *Int. J. of Fatigue*, 12(5):403–407, 1990.

[177] ASTM E647 08. Standard test method for measurement of fatigue crack growth rates. *ASTM International*, West Conshohocken, PA, 2008.

[178] K. Tanaka and Y. Akiniwa. Fatigue crack propagation behaviour derived from S–N data in very high cycle regime. *Fatigue & Fracture of Engineering Materials & Structures*, 25(8–9):775–784, 2002.

[179] J. R. Weeks, B. Vyas, and H. S. Isaacs. Environmental factors influencing stress corrosion cracking in boiling water reactors. *Corrosion Science*, 25(8–9):775–768, 1985.

[180] A. Ohta, N. Suzuki, and T. Mawari. Effect of Young's modulus on basic crack propagation properties near fatigue threshold. *Int. J. Fatigue*, 4:224–226, 1992.

[181] M. D. Chapetti, H. Miyata, T. Tagawa, T. Miyata, and A. Fujioka. Fatigue crack propagation behaviour in ultra-fine grained low carbon steel. *Int. J. of Fatigue*, 27(3):235–243, 2005.

[182] G. Chai. Fatigue behaviour of duplex stainless steels in the very high cycle regime. *Int. J. Fatigue*, 28(11):1611–1617, 2006.

[183] M. da Fonte, F. Romerio, M de Freitas, S. E. S. Tschegg, and E. K. Tsehegg. The effect of microstructure and environment on fatigue crack growth in 7049 aluminium alloy at negative stress ratios. *Int. J. of Fatigue*, 25:1209–1216, 2003.

[184] K. Nakajima, K. Terao, and T. Miyata. The effect of microstructure on fatigue crack propagation of $\alpha + \beta$ titanium alloys In-situ observation of short fatigue crack growth. *Materials Science & Engineering A*, A243:176–181, 1998.

[185] H. Mughrabi. Specific features and mechanisms of fatigue in the ultrahigh-cycle regime. *International Journal of Fatigue*, 28(11):1501–1508, 2006.

[186] P. C. Paris, M. Gomez, and W. E. Anderson. A rational analytical theory of fatigue. *Trend Engng*, 13:9–14, 1961.

[187] W. Elber. The significance of fatigue crack closure. *ASTM STP*, 486:230–242, 1971.

[188] T. V. Duggan. A theory for fatigue crack propagation. *Engineering Fracture Mechanics*, 9:735–747, 1977.

[189] A. A. Griffith. The phenomena of rupture and flow in solids. *Phil. Trans. Roy. Soc. Lond.*, A221:163–198, 1920.

[190] D. V. Ramsamooj and T. A. Shugar. Model prediction of fatigue crack propagation in metal alloys in laboratory air. *Int. J. of Fatigue*, 23:S287–S300, 2001.

[191] A. S. Khan and T.K. Paul. A new model for fatigue crack propagation in 4340 steel. *Int. J. of Plasticity*, 10(8):957–972, 1994.

[192] D. F. Socie. Prediction of fatigue crack growth in notched members under variable amplitude loading histories. *Engineering Fracture Mechanics*, 9:849–865, 1977.

[193] G. M. Seed and G. S. Murphy. The applicability of neural networks in modelling the growth of short fatigue cracks. *Fatigue & Fracture of Engineering Materials & Structures*, 21:183–190, 1998.

[194] M. E. Haque and K. V. Sudhakar. Artificial neural network based prediction model for fatigue crack growth in dual phase steel. *Fatigue & Fracture of Engineering Materials & Structures*, 23:63–68, 2001.

[195] M. Jono, Y. Iono, A. Ohta, S. Kawai, H. Kobayashi, K. Komai, T. Sakai, K. Hirakawa, and K. Mori. Data book on fatigue crack growth rates of metallic materials, Vol. 1 and Vol. 2. Technical report, The Society of Materials Science, Japan, Kyoto, Japan, 1983.

[196] S. Suresh. *Fatigue of Materials*. Cambridge University Press, 2008.

[197] J. R. Griffiths and C. E. Richards. Fatigue testing. *Mater. Sci. Eng R*, 11:305–315, 1973.

[198] D. Broek and J. Schijve. Fatigue crack growth; effect of sheet thickness. *Aircraft Engineering*, 38(11):31–33, 1966.

[199] E. K. Walker. Effects of environments and complex load history on fatigue life. *ASTM STP*, 462:1–14, 1970.

[200] A. Otsuka, Y. Fujii, and K. Maeda. A new testing method to obtain Mode II fatigue crack growth characteristics of hard materials. *Fatigue Fract Engng Mater Struct*, 27:203–212, 2004.

[201] M. A. Hicks and J. E. King. Temperature effects on fatigue thresholds and structure sensitive crack growth in a nickel-base superalloy. *Int. J Fatigue*, 5:67–74, 1983.

[202] R. L. Tobler. Low temperature effects on the fracture behaviour of a nickel base superalloy. *Cryogenics*, 1:669–674, 1976.

[203] L. A. James and W. J. Mills. Effect of heat treatment and heat to heat variations in the fatigue crack growth response of alloy 718. *Engineering Fracture Mechanics*, 22(5):797–817, 1985.

[204] J. Telesman and L. J. Ghosn. The unusual near-threshold fatigue crack growth behaviour of a single crystal superalloy and the resolved shear stress as the crack driving force. *Engineering Fracture Mechanics*, 34(5/6):1183–1196, 1989.

[205] J. C. Healy, L. Grabowski, and C. J. Beevers. Short fatigue crack growth in a nickel base superalloy at room and elevated temperature. *Int J. Fatigue*, 2:1183–1196, 1991.

[206] S. S. Tschegg. Fatigue crack growth and thresholds at ultrasonic frequencies. *Int J. Fatigue*, 28:1456–1464, 2006.

[207] T. L. MacKay. Fatigue crack propagation rate at low ΔK of two aluminium sheet alloys, 2024-T3 and 7075-T6. *Engineering Fracture Mechanics*, 11:753–761, 1979.

[208] SKF. *General Catalogue*, June 2003.

[209] A. P. Voskamp. *Microstructural Changes During Rolling Contact Fatigue*. PhD thesis, University of Delft, 1996.

[210] L. Coelho, A. Dias, H. P. Lieurade, and H. Maitournam. Experimental and numerical rolling contact fatigue study on the 32Cr-MoV13 steel. *Fatigue Fract. Eng. Mater. Struct.*, 27(9):811–823, 2004.

[211] T. A. Harris. *Rolling Bearing Analysis*. Chichester Wiley, New York, 2001.

[212] A. C. Batista, D. Morao, J. C. Le Flour, and J. L. Lebrun. Residual stress and damage evolution during pitting tests of automotive gears. *Proceedings of the Fourth International Conference on Residual Stresses*, pages 1060–1070, 1994.

[213] H. H. Munro and N. Tsushima. Microstructural and microhardness and residual stress changes due to rolling contact. *Wear*, 15:309–330, 1970.

[214] S. Way. Pitting due to rolling contact. *ASME J. Appl. Mech.*, 2:A49–A58., 1935.

[215] A. P. Voskamp, R. Osterlund, P. Becker, and O. Vingsbo. Gradual changes in residual stress and microstructure during contact fatigue in ball bearings. *Met. Tehnol.*, 7:14–21, 1980.

[216] R. Osterlund and O. Vingsbo. Phase changes in fatigued ball bearings. *Metall. and Mat. Trans. A*, 11(5):701–707, 1980.

[217] W. E. Poole, R. L. Thom, E. V. Zarestky, W. J. Anderson, and E. N. Bamberger. *Materials and Processing*. E.V. Zarestky, 1992.

[218] R. K. Kepple and R. L. Mattson. Rolling element fatigue and macroresidual stress. *J. Lub. Tech*, 92(1):76–82, 1970.

[219] R. L. Scott, R. K. Kepple, and M. H. Miller. The effect of processing induced near-surface residual stress on ball bearing fatigue. *Roll. Cont. Phenom.*, 2:303–306, 1962.

[220] I. V. Rivero and C. O. Ruud. Deviation of residual stress patterns in 52100 bearings steel due to inherent microstructural transformations after rolling contact. *Mat. Charact.*, 53:381–393, 2004.

[221] B. Bhushan. *Introduction to Tribology*. J. Wiley & Sons, New York, 2002.

[222] G. Lundberg and A. Plamgren. Dynamic capacity of rolling bearings. *Acta Ploytechnica Scandinavica*, 85:1–67, 1947.

[223] W. Weibull. A statistical theory of the strength of materials. *Acta Ploytechnica Scandinavica*, 77:1–34, 1939.

[224] E. Ioannides, G.Bergling, and A. Gabelli. An analytical formulation for the life of rolling bearings. *Acta Ploytechnica Scandinavica*, 137:1–80, 1999.

[225] D. Umbrello, G. Ambrogio, L. Filice, and R. Shivpuri. A hybrid finite element method-artificial neural network approach for predicting residual stresses and the optimal cutting conditions during hard turning of AISI 52100 bearing steel. *Materials & Design*, 29:873–883, 2008.

[226] R. Shivpuri, H. Jiang, and D. Umbrello. Investigation of cutting conditions and cutting edge preparations for enhanced compressive subsurface residual stress in the hard turning of bearing steel. *Journal of Materials Processing Technology*, 171(2):180–187, 2006.

[227] Y. B. Guo and M. E. Barkey. FE-simulation of the effects of machining-induced residual stress profile on rolling contact of hard machined components. *Int. J. of Mechanical Sciences*, 46(3):371–388, 2004.

[228] M. H. El-Axir. A method of modelling residual stress distribution in turning for different materials. *Wear*, 42(9):1055–1063, 2002.