

Aki Sorsa

PREDICTION OF MATERIAL
PROPERTIES BASED ON NON-
DESTRUCTIVE BARKHAUSEN
NOISE MEASUREMENT

UNIVERSITY OF OULU GRADUATE SCHOOL;
UNIVERSITY OF OULU,
FACULTY OF TECHNOLOGY,
DEPARTMENT OF PROCESS AND ENVIRONMENTAL ENGINEERING



ACTA UNIVERSITATIS OULUENSIS
C Technica 442

AKI SORSA

**PREDICTION OF MATERIAL
PROPERTIES BASED ON NON-
DESTRUCTIVE BARKHAUSEN
NOISE MEASUREMENT**

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Technology and Natural Sciences of the University of Oulu for public defence in Kuusamonsali (Auditorium YB210), Linnanmaa, on 1 February 2013, at 12 noon

UNIVERSITY OF OULU, OULU 2013

Copyright © 2013
Acta Univ. Oul. C 442, 2013

Supervised by
Professor Kauko Leiviskä

Reviewed by
Doctor Brian Shaw
Professor Mikko Kolehmainen

ISBN 978-952-62-0067-5 (Paperback)
ISBN 978-952-62-0068-2 (PDF)

ISSN 0355-3213 (Printed)
ISSN 1796-2226 (Online)

Cover Design
Raimo Ahonen

JUVENES PRINT
TAMPERE 2013

Sorsa, Aki, Prediction of material properties based on non-destructive Barkhausen noise measurement.

University of Oulu Graduate School; University of Oulu, Faculty of Technology, Department of Process and Environmental Engineering, P.O. Box 4300, FI-90014 University of Oulu, Finland
Acta Univ. Oul. C 442, 2013
Oulu, Finland

Abstract

Barkhausen noise measurement is an intriguing non-destructive testing method suitable for ferromagnetic materials. It is based on the stochastic movements of magnetic domain walls when the tested sample is placed in an external varying magnetic field. Barkhausen noise is typically utilised so that some features are calculated from the signal and then compared with the studied material properties. Typical features are, for example, the root-mean-square value (RMS), peak height, width and position. Better utilisation of the method, however, requires quantitative predictions of material properties.

The aim of this thesis is to study and select a suitable methodology for the quantitative prediction of material properties based on Barkhausen noise measurement. The prediction considered is divided into four steps: feature generation, feature selection, model identification and model validation. In feature generation, a large set of features is calculated with different mathematical procedures. This feature set is explored in the feature selection step to find the most significant features in terms of predictions. A model with the selected features is identified and some independent data are usually used for model validation.

This thesis presents the developed procedures required in feature generation and the results of the studies using different feature selection strategies and modelling techniques. The studied feature selection methods are forward selection, simulated annealing and genetic algorithms. In addition, two-step algorithms are investigated where a pre-selection step is used before the actual selection. The modelling techniques used are multivariable linear regression, partial least squares regression, principal component regression and artificial neural networks. The studies also consider the use and effect of different objective functions.

The results of the studies show that the proposed modelling scheme can be used for the prediction task. The models identified mainly include reasonable terms and the prediction accuracy is fairly good considering the challenge. However, the application of Barkhausen noise measurement is very case-dependent and thus conflicts may occur. Furthermore, the changes in unmeasured material properties may lead to the unexpected behaviour of some features. The results show that linear models are adequate for capturing the major interactions between material properties and Barkhausen noise but indicate that the use of neural networks would lead to better model performance. The results also show that genetic algorithms give better selection results but at the expense of the computational cost.

Keywords: Barkhausen noise, feature selection, modelling, non-destructive testing, residual stress

Sorsa, Aki, Materiaaliominaisuuksien ennustaminen materiaalia rikkomattoman Barkhausen-kohina-mittauksen avulla.

Oulun yliopiston tutkijakoulu; Oulun yliopisto, Teknillinen tiedekunta, Prosessi- ja ympäristötekniikan osasto, PL 4300, 90014 Oulun yliopisto

Acta Univ. Oul. C 442, 2013

Oulu

Tiivistelmä

Barkhausen-kohina-mittaus on ferromagneettisille materiaaleille soveltuva materiaalia rikkomaton testausmenetelmä. Mittaus perustuu magneettisten alueiden välisten rajapintojen stokastisiin liikkeisiin, kun testattava kappale asetetaan vaihtuvaan magneettikenttään. Tyypillisesti Barkhausen-kohina-mittaussignaalista lasketaan piirteitä, joita sitten verrataan tutkittaviin materiaaliominaisuuksiin. Usein käytettyjä piirteitä ovat signaalin keskineliön neliöjuuri (RMS-arvo) sekä piikin korkeus, leveys ja paikka. Menetelmää voidaan soveltaa paremmin, jos tutkittavia materiaaliominaisuuksia voidaan ennustaa kvantitatiivisesti.

Tämän tutkimuksen tavoitteena on tutkia ja valita menetelmiä, jotka soveltuvat materiaaliominaisuuksien kvantitatiiviseen ennustamiseen Barkhausen-kohina-mittauksen perusteella. Ennustusmallit luodaan neljässä vaiheessa: piirteiden laskenta, piirteiden valinta, mallin identifiointi ja mallin validointi. Piirteiden laskennassa yhdistellään erilaisia matemaattisia laskutoimituksia, joista tuloksena saadaan suuri joukko erilaisia piirteitä. Tästä joukosta valitaan ennustukseen soveltuvimmat piirteiden valinta -vaiheessa. Tämän jälkeen ennustusmalli identifioidaan ja viimeisessä vaiheessa sen toimivuus todennetaan riippumattomalla testausaineistolla.

Väitöskirjassa esitetään piirteiden laskentaan kehitettyjä algoritmeja sekä mallinnustuloksia käytettäessä erilaisia piirteiden valintamenetelmiä ja mallinnustekniikoita. Tutkitut valintamenetelmät ovat eteenpäin valinta, taaksepäin eliminointi, simuloitu jäähtyminen ja geneettiset algoritmit. Väitöskirjassa esitellään myös kaksivaiheisia valintamenettelyjä, joissa ennen varsinaista piirteiden valintaa suoritetaan esivalinta. Käytetyt mallinnustekniikat ovat monimuuttujaregressio, osittainen pienimmän neliösumman regressio, pääkomponenttiregressio ja neuroverkot. Tarasteluissa huomioidaan myös erilaisten kustannusfunktioiden vaikutukset.

Esitetyt tulokset osoittavat, että käytetyt menetelmät soveltuvat materiaaliominaisuuksien kvantitatiiviseen ennustamiseen. Identifioidut mallit sisältävät pääasiassa perusteltavia termejä ja mallinnustarkkuus on tyydyttävä. Barkhausen-kohina-mittaus on kuitenkin erittäin tapauskohtainen ja täten ristiriitoja kirjallisuuden kanssa voidaan joskus havaita. Näihin ristiriitoihin vaikuttavat myös ei-mitattavat muutokset materiaaliominaisuuksissa. Esitetyt tulokset osoittavat, että lineaariset mallit kykenevät ennustamaan suurimmat vuorovaikutukset materiaaliominaisuuksien ja Barkhausen-kohinan välillä. Tulokset kuitenkin viittaavat siihen, että neuroverkoilla päästäisiin vielä parempiin mallinnustuloksiin. Tulokset osoittavat myös, että geneettiset algoritmit toimivat piirteiden valinnassa paremmin kuin muut tutkitut menetelmät.

Asiasanat: ainetta rikkomaton testaus, Barkhausen-kohina, jännösännitys, mallinnus, piirteiden valinta

Preface

This work was carried out at the Control Engineering Laboratory, Department of Process and Environmental Engineering, University of Oulu during the years 2008–2012. During these years two research projects were completed. INTELBARK-project started in 2008 and was funded by Academy of Finland. That project was followed by NOVEBARK funded by TEKES.

First, I want to thank my supervisor Prof. Kauko Leiviskä for guidance, support and comments during the work. I also want to thank the reviewers of the thesis, Dr. Brian Shaw from Newcastle University and Prof. Mikko Kolehmainen from University of Eastern Finland.

Both of the research projects mentioned above were accomplished in cooperation with the Department of Materials Science in Tampere University of Technology and industrial partners Stresstech Oy, Moventas Wind Oy, Takoma Gears Oy and Katsa Oy. I want to thank the industrial partners for the challenging research task and especially the people at Stresstech Oy for their interest and comments on my work. I am most grateful to Suvi Santa-aho from Tampere University of Technology for her effort on carrying out the Barkhausen noise measurements, numerous fruitful discussions and friendship during the research work. I also want to express my gratitude to all my colleagues in Control Engineering Laboratory, especially to Jari Ruuska.

I want to express my deepest gratitude to Elina for her endless support and patience. I am also grateful to my parents, Ulla and Ilkka, and my siblings, Janne, Leena and Päivi, for their support and influence during my whole life. My son, Aleks, has taught me how to set the priorities correctly and is greatly appreciated for just being around.

List of abbreviations and symbols

ANN	Artificial neural networks
BN	Barkhausen noise
BPNN	Backpropagation neural network
CCD	Central composite experimental design
CPCR	Correlation-based principal component regression
CV	Subscript referring to cross-validation
FWHM	Full width at half-maximum
LOO	Leave-one-out
LMO	Leave-multiple-out
MLR	Multivariable linear regression
MSEP	Mean of the squared error of prediction
NDT	Non-destructive testing
PCA	Principal component analysis
PCR	Principal component regression
PLS	Partial least squares
PLSR	Partial least squares regression
PNN	Probabilistic neural network
RBFNN	Radial basis function neural network
RMS	Root mean square
RMSEP	Root mean squared error of prediction
SPA	Successive projection algorithm
SSEP	Sum of the squared error of prediction
a	Fitting parameter
B	Magnetic flux
\mathbf{b}, b	Vector of regression coefficients, regression coefficient
β	Constant in simulated annealing
B_r	Remanence
c	Constant in simulated annealing / Scaling parameter
c_h	Centre of the radial basis function
δ	Uniform random number in simulated annealing
$\mathbf{E}, E, \mathbf{e}, e$	Error matrix, system energy in simulated annealing, error vector, error
\mathbf{F}	The weight matrix in radial basis function neural networks
f	Activation function / Fitting function
H_v	Vicker's hardness
H	Entropy / Applied magnetic field
h	Peak height
H_c	Coercivity
J	Objective function value
J_{bsf}	Best objective function value found

J_{cons}	Conservative guess of the minimum objective function value
J_{cur}	Objective function value of the current solution
J_{new}	Objective function value of the new solution
k	Parameter in k-fold cross-validation / Slope / Boltzmann's constant in simulated annealing
λ	Penalty constant
M	Number of input variables
m	Selected number of input variables
N	Number of data points
N_{bits}	Number of bits
N_{CV}	Number of data points in internal validation data set
N_{gen}	Number of generations
N_{pop}	Population size
\mathbf{P}, P	Loading matrix in PCR and PLSR / Number of output variables
p	Probability / Peak position
p_c	Crossover probability
p_m	Mutation probability
s	Spread of the normal distribution
σ	Residual stress
σ_h	Spread of the radial basis function
$\mathbf{T}, T, \mathbf{t}$	Score matrix in PCR and PLSR, temperature in simulated annealing, score vector
\mathbf{W}, w	Matrix of network weights, weighting coefficient / Peak width
$\mathbf{X}, \mathbf{x}, x$	Input matrix, input vector, input variable
$\mathbf{Y}, \mathbf{y}, y$	Output matrix, output vector, output variable

Contents

Abstract	
Tiivistelmä	
Preface	7
List of abbreviations and symbols	9
Contents	11
1 Introduction	13
1.1 Background	13
1.2 Scope of the work	14
1.3 Research problem.....	15
1.4 Research contribution	16
1.5 Author's contribution	16
1.6 Structure of the thesis.....	17
2 Barkhausen phenomenon and its use in non-destructive testing	19
2.1 The origin of the Barkhausen effect.....	19
2.2 Barkhausen noise measurement	20
2.3 Barkhausen noise signal analysis	21
2.4 Material characterization with Barkhausen noise	23
2.4.1 Microstructure	24
2.4.2 Hardness	25
2.4.3 Stress state	26
2.5 Barkhausen noise applications	27
2.5.1 Deformation detection	27
2.5.2 Fatigue monitoring	29
2.5.3 Case depth evaluation.....	30
2.5.4 Grinding burn detection.....	32
3 Variable selection and model identification	33
3.1 Filter methods	34
3.2 Mathematical modelling techniques.....	36
3.2.1 Multivariable linear regression.....	36
3.2.2 Principal component regression	37
3.2.3 Partial least squares regression	40
3.2.4 Artificial neural networks.....	41
3.3 Optimal search engines	44
3.4 Deterministic search engines.....	44
3.5 Stochastic search engines.....	46

3.5.1	Simulated annealing	46
3.5.2	Genetic algorithms.....	50
3.6	Objective functions	59
3.6.1	Resampling methods	60
3.6.2	Criteria used.....	61
3.7	Functions used in the studies.....	62
4	Experimental set up	65
4.1	Studied materials.....	65
4.2	Measurements	65
4.3	Data sets	66
5	Results	67
5.1	Modelling approach	67
5.2	Feature generation.....	68
5.2.1	Features calculated directly from the BN signal.....	68
5.2.2	Features of the BN profile	70
5.3	Feature selection.....	72
5.3.1	Manual selection.....	72
5.3.2	Forward selection and backward elimination	74
5.3.3	Simulated annealing	78
5.3.4	Genetic algorithms.....	78
5.3.5	Two-step selection strategies	80
5.4	Modelling techniques	86
5.5	Interpretation and validity of the results.....	89
5.6	Summary of the results.....	90
6	Discussion and future work	93
6.1	Feature generation.....	93
6.2	Feature selection procedures	94
6.3	Modelling techniques and model accuracy	97
6.4	Objective functions	98
6.5	The significance of the selected features.....	100
6.6	Meeting the objectives of the research.....	103
7	Conclusions	105
	References	109

1 Introduction

1.1 Background

The term “non-destructive testing” (NDT) is used for material testing methods that can be applied without compromising the usefulness of the material. Tests can be applied to materials, parts, assemblies or structures. Most non-destructive methods are indirect, producing an estimation of the quality, strength or serviceability of the tested object. (McMaster 1963)

Destructive methods can also be applied to material testing. It is, however, obvious that the tested object is lost in such a testing scheme and thus cannot be applied to objects that are later used in service. Destructive tests can be used for producing estimations of the objects’ performance in service. However, the tested objects and the testing conditions must be very similar to the objects and conditions in service to guarantee that the estimations are reliable. (McMaster 1963)

Non-destructive tests can be applied to objects that are later used in service. Thus it is possible to distinguish between valid and defective objects and remove or reprocess the latter. However, the usability of the selected non-destructive testing method depends on the agreement between the non-destructively measured property and the property determining the serviceability of the object. In order to use a non-destructive testing method in a certain application, this agreement must be verified. Thus, the applicability of the method must be proven for the material used, as well as the method of processing and service application. (McMaster 1963)

Non-destructive methods offer many advantages in industry. These advantages mainly fall into four categories: increased productivity, increased serviceability, safety and identification of materials. The reasons for increased productivity are, for example, the prevention of wasted material and manpower, better and more uniform quality, lowered operating and production costs, and process monitoring and control. Increased serviceability is obtained by preventing the malfunctioning and breakdown of equipment while the safety contribution is associated with the prevention of accidents. The last category is the identification of materials, which refers to the benefits obtained from identifying the differences in chemical composition and the physical and metallurgical properties of materials. (McGonnagle 1963)

Even though non-destructive methods seem more attractive, destructive methods also have some benefits. The results are usually quantitative while only qualitative results are often obtained with non-destructive methods. The agreement between the measured property and the serviceability controlling property is direct and thus there is no need to verify it. (McMaster 1963)

There is a wide variety of non-destructive testing methods utilizing different physical phenomena. Test methods can be divided into visual, pressure and leak, penetrant, thermal, radiography, acoustic, magnetic, electrical and electrostatic, electromagnetic induction and miscellaneous. (McGonnagle 1963)

Barkhausen noise (BN) measurement is an electromagnetic testing method that is suitable for ferromagnetic materials (Cullity 1972). It has been shown to be sensitive to different material properties such as microstructure, composition, residual stress and hardness. The measurement is intriguing because it is fast and cheap and requires only rather simple equipment. The studies found in the literature (for example Stewart *et al.* 2003, Augustyniak *et al.* 2010) provide more or less only qualitative results, where a studied material property is compared with some feature that is obtained from the BN measurement signal. The quantitative prediction of material properties is desirable in order to utilise the measurements in NDT applications better. The potential applications are, for example, material characterisation and deformation, fatigue, case-depth and grinding burn evaluation.

1.2 Scope of the work

This work aims at investigating and selecting suitable methodology for predicting material properties based on Barkhausen noise measurement. This subject has not been that well studied in the literature and thus it is of great interest. The prediction starts from feature generation, where information is extracted from the measurement signal using different mathematical procedures. This step is crucial in considering the performance of the models and thus an efficient and adequate feature generation step is one of the main goals. Another main goal is the efficient identification of prediction models. In this task, feature selection is essential and thus studied thoroughly in this thesis. The modelling technique deployed naturally affects the predictions and therefore different modelling techniques are also tested.

This study concentrates on the analysis and utilisation of Barkhausen noise signals. Considerations about the improvement of the measurement system are beyond the scope of this work. In addition, the fundamental understanding of the

phenomenon and development of a mechanistic model (even though important topics) are not investigated here.

1.3 Research problem

This thesis discusses the quantitative prediction of material properties based on BN measurement. The problem is basically two-fold. The first problem is related to the Barkhausen phenomenon and the challenges when applying it to NDT. The second problem is related to model development.

The Barkhausen phenomenon has a stochastic nature due to the irregular and stochastic movement of magnetic domain walls within the material and the complex interactions between material properties and BN that exist. Furthermore, the influence of unmeasured material properties leads to unexplained changes in BN. When applying BN measurement to NDT applications, another problem is the case-specificity of the measurement. Depending on the material and application, different properties of the measurement signal are significant. Thus, automatic procedures are needed to find models that can predict the studied material properties.

Before the problem of finding a model can even be tackled, information must be extracted in a suitable form from the measurement signals. Thus, the first step in model development is feature generation, where different mathematical procedures are used to produce features with different information content. The outcome of feature generation is a large set of features to be used in predicting material properties.

The first task in model identification is the selection of features because irrelevant features in the model will deteriorate its performance. Selection is affected by the model structure, which also needs to be selected. The model structure should map the interactions between the selected input variables and the explained output variable. However, too complex model structures may lead to overfitted models that cannot be used in future predictions. Linear models are simple and able to capture the major interactions. Obviously, they are not capable of mapping nonlinear interactions and thus may not lead to the best possible model performance. The future use of the model must also be guaranteed in the validation step.

1.4 Research contribution

The main contribution of this thesis is to enhance knowledge and experience about suitable methodology for the quantitative prediction of material properties based on Barkhausen noise measurement. Feature generation and selection procedures are tested and compared, and also suitable modelling techniques are studied. The potential for using the obtained results in industrial on-line or off-line applications is considered. Even though only the prediction of material properties is studied here, the methodology can also be quite easily converted to the needs of quality control where, for example, the detection of defective components or identification of the hardened surface layer thickness of components is required.

1.5 Author's contribution

The research work for this thesis was carried out in two projects in close cooperation with the Laboratory of Materials Science at Tampere University of Technology. The present author was the researcher responsible for the development of the prediction models of material properties and thus the main author of the publications in that field. The publications that form the basis of this thesis are (Sorsa & Leiviskä 2009a), (Sorsa & Leiviskä 2009b), (Sorsa & Leiviskä 2010), (Sorsa & Leiviskä 2011a), (Sorsa & Leiviskä 2011b), (Sorsa & Leiviskä 2011c), (Sorsa *et al.* 2008a), (Sorsa *et al.* 2008b), (Sorsa *et al.* 2010b), (Sorsa *et al.* 2012a) and (Sorsa *et al.* 2012b). For these publications, the present author carried out all the methodological research work concerning modelling while Suvi Santa-aho, M. Sc., was responsible for material handling and data acquisition. Prof. Kauko Leiviskä was the supervisor of the research work while the other authors provided advice and commented on the manuscripts. Section 6 also gives some results that are yet unpublished. The present author is the main contributor of these studies as well. The projects also produced a lot of publications concerning the basic understanding of the Barkhausen phenomenon and its relationship to material properties. Publications in this field concerning the Barkhausen phenomenon are relevant and are thus referred to in Section 2. However, the present author is not the main contributor of these publications.

1.6 Structure of the thesis

The first section introduces the topic with background information and the scope of the research. Section 2 concentrates on the Barkhausen phenomenon and its measurement. This section also presents a brief literature review about the possible applications of Barkhausen noise measurement. Section 3 presents the methodology used in variable selection. The methodology includes different modelling techniques, search engines and objective functions. In Section 4, the experimental set up is given. The studies and their results are presented in Section 5 while discussion and future considerations are given in Section 6. Section 7 summarises the thesis.

2 Barkhausen phenomenon and its use in non-destructive testing

In this section, the theoretical background behind Barkhausen noise and its measurement are briefly presented. It also includes a short literature review on the use of Barkhausen noise in different NDT applications.

2.1 The origin of the Barkhausen effect

The magnetic properties of materials are due to the moving electrons in atoms, creating atomic dipole moments. In ferromagnetic materials, these dipole moments are oriented parallel within small magnetic domains separated by domain walls. When no external magnetic field is present, the directions of these magnetic domains are random and thus they cancel each other out and the overall magnetization of the piece is zero. (Cullity 1972)

When the piece is placed in an increasing external magnetic field, the domains start to change their sizes so that the domains with a direction close to the applied field increase their size. Thus, the magnetization of the piece increases. When the external field is increased enough, the magnetization of the piece saturates to a maximum value (B_S). In this case, there is only one magnetic domain within the material with a direction parallel to the applied field. (Cullity 1972, Jiles 2000, Durin & Zapperi 2006)

Magnetic hysteresis is the difference between the magnetizing and demagnetizing paths of the material. When the applied field strength is decreased after the saturating magnetization is reached, the magnetization of the material starts to decrease. However, decreased magnetization does not follow the same curve as increased magnetization. When the applied field strength is reduced to zero, the piece still remains magnetized. The remaining magnetization is called remanence (B_R). When the external field is now increased in the opposite direction, the magnetization of the material further decreases. The opposite external field needed to obtain zero magnetization is called coercivity (H_C). Again, the magnetization saturates when the field strength is further increased. Decreasing the external field, reversing it and increasing it again leads to a hysteresis loop as shown in Fig. 1. (Anderson *et al.* 1990)

Even though the applied external magnetic field changes smoothly and the hysteresis loop shown in Fig. 1 seems smooth, the domain wall motions are jerky and discontinuous. This is due to what are termed pinning sites within the

material. The domain walls get trapped behind these sites and when the external field is strong enough, they abruptly and irreversibly break away from the pinning sites. These sudden jumps lead to abrupt changes in the magnetization of the material. Detecting these changes leads to the noise-like signal called the Barkhausen noise. (Cullity 1972, Jiles 2000, Durin & Zapperi 2006)

The pinning sites are small inhomogeneities within the material. They may be, for example, grain boundaries, small volumes of second phase material or inclusions. Also, residual microstresses affect the domain wall motions and thus the Barkhausen noise. (Cullity 1972, Jiles 2000)

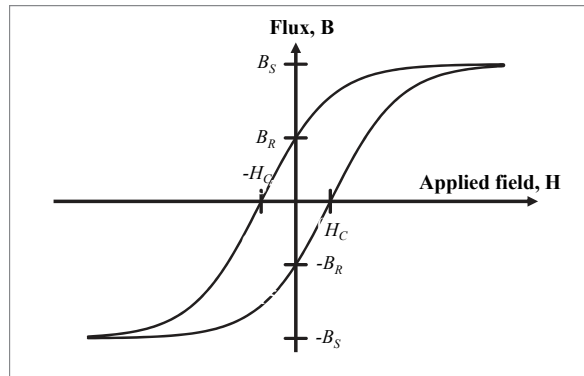


Fig. 1. The hysteresis loop.

2.2 Barkhausen noise measurement

The Barkhausen noise measuring principle is simple. An alternating current is fed to the magnetizing yoke to generate a changing magnetic field for the repeated magnetizing and demagnetizing cycles of the specimen. The magnetic field that is induced in the specimen is captured with a pick-up coil, amplified and filtered to obtain the Barkhausen noise signal. A diagram of a typical Barkhausen noise-measuring instrument is shown in Fig. 2.

Barkhausen noise measurement only gives information from the near-surface region of the material. The penetration depth of the measurement depends on the magnetization frequency used. With higher frequencies, the penetration depth of the measurement is lower (Jiles 2000, Moorthy *et al.* 2005, Wilson *et al.* 2009). A rough division into high and low frequency approaches is given in (Moorthy *et al.*

2005). In the high frequency analysis, the magnetizing frequency higher than 10 Hz is used while low frequency approaches use frequencies lower than 1 Hz.

The external magnetic field typically changes according to triangular or sinusoidal functions. Even though Durin & Zapperi (2006) suggest that triangular waveform should be preferred, both are used in the literature.

For many applications, Barkhausen noise measurement has certain advantages. The measurement is fast and only rather simple equipment is needed (Lindgren & Lepistö 2002). It is also possible to measure components with complex geometries (Moorthy *et al.* 2005, Desvaux *et al.* 2004).

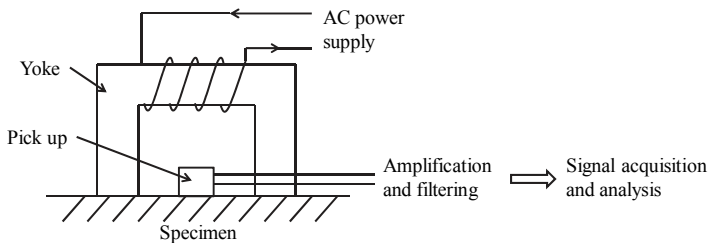


Fig. 2. The diagram of a Barkhausen noise instrument; drawn based on (Lindgren 2003) and (Santa-aho 2008).

2.3 Barkhausen noise signal analysis

A typical Barkhausen noise signal is shown in Fig. 3. Usually, some feature is calculated from the signal and then compared with studied material properties. The root-mean-squared (RMS) value of the signal is the most usual one. It has been used, for example, in (Lindgren & Lepistö 2002) and (O’Sullivan *et al.* 2003). It is given by:

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} . \quad (1)$$

Above, N is the number of data points and \mathbf{x} is the data vector whose i :th entry is x_i . Another usual feature is Barkhausen noise energy, which is obtained by integrating the squared signal over one Barkhausen noise envelope (Kim *et al.* 1992). The envelope includes measurements from half of one magnetization loop (one burst in Fig. 3). Later, Barkhausen noise energy was also calculated for the whole measurement signal (Gauthier *et al.* 1998).

A more recent approach considers Barkhausen noise as a function of the applied magnetic field. Such an approach is used, for example, in (Moorthy *et al.* 1998), (Saquet *et al.* 1999) and (Stewart *et al.* 2004). Typically, the RMS value of the measurement signal is taken followed by a moving average filtering (Blaow *et al.* 2006). Through such a procedure, what is termed the Barkhausen noise profile is obtained. A typical profile is shown in Fig. 4. The profile is calculated for ascending and descending applied magnetic field directions separately. These profiles should be mirror images of each other (Blaow *et al.* 2006) and thus their average can be used. The profile height, width and position (typically called peak height, width and position) are then determined and compared with the material properties. It is also possible to use specific points of the profile to explain changes in material properties (Wilson *et al.* 2009). Also, the maximum slope of the profile has been used (Ng *et al.* 2001).

A problem with Barkhausen noise measurement is the robustness of the calculated features. It is proposed in (Stewart *et al.* 2004) that more robust features are obtained if certain functions are fitted to the profile and the fitting parameters are used as features. The Lorentzian function was used in (Stewart *et al.* 2004) and the exponentially modified Gaussian distribution in (Blaow *et al.* 2006).

The features that can be used are not limited to those mentioned above. It is quite common to convert the data into a frequency domain and then analyse the obtained spectra. Frequency domain analysis has been used, for example, in (Yamaura *et al.* 2001) and (Davut & Gür 2007). Yamaura *et al.* (2001) further divided the spectra into frequency ranges and used the ratios of those as features. In addition, pulse height distribution (Davut & Gür 2007) and statistical values such as skewness (Stewart *et al.* 2004) have been used.

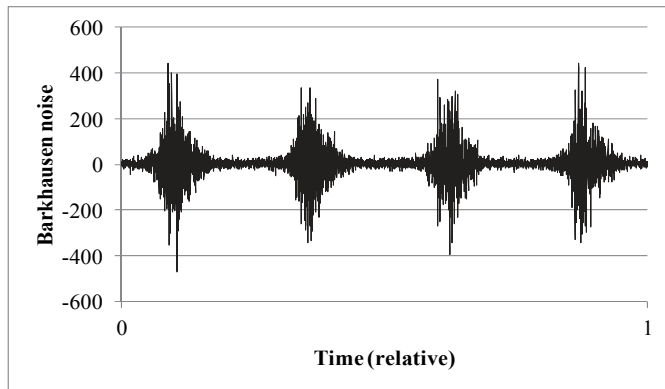


Fig. 3. A typical Barkhausen noise signal with 4 bursts.

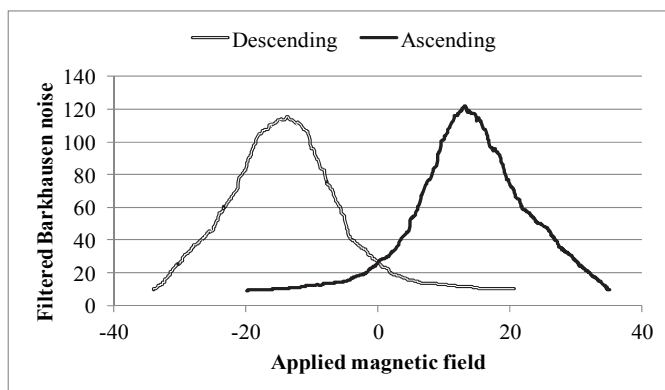


Fig. 4. A typical Barkhausen noise profile.

2.4 Material characterization with Barkhausen noise

Barkhausen noise has been shown to be very sensitive to different material properties. The microstructure, residual stress and hardness of the material have all been studied with Barkhausen noise. The following is a more detailed review of the studies found in the literature.

2.4.1 Microstructure

The microstructure of the material has an influence on the physical and magnetic properties of materials. The term “microstructure” includes information about the different phases present in the material. The number, distribution, volume fraction, shape and size of the phases all define the microstructure of the material. (Anderson *et al.* 1990)

In the literature, many microstructural properties of materials have been studied and compared to Barkhausen noise measurements. Such properties include the grain size (Ranjan *et al.* 1987, Moorthy *et al.* 1997a, Gaterier-Rothea *et al.* 1998, Yamaura *et al.* 2001, Anglada-Rivera *et al.* 2001, Kim & Kwon 2003), ferrite, pearlite and martensite phases (Saquet *et al.* 1999, Moorthy *et al.* 2000, D’Amato *et al.* 2003, Koo *et al.* 2003, Kaplan *et al.* 2007, Kleber *et al.* 2008) and the carbon content (Gatelier-Rothea *et al.* 1998, Ng *et al.* 2001, Koo *et al.* 2003, Capó-Sánchez *et al.* 2004, Pérez-Benitez *et al.* 2005). The physical background for the influence of the grain size is the increased number of magnetic domains and domain walls with smaller grain sizes (Jiles 2000). Basically, this means that more Barkhausen jumps occur but with a smaller amplitude (Jiles 2000). However, the increased number of domain walls leads to increased Barkhausen activity. In (Moorthy *et al.* 1997a), the feature corresponding to the amount of Barkhausen activity was the peak height while Gatelier-Rother *et al.* (1998) and Anglada-Rivera *et al.* (2001) used the maximum Barkhausen noise amplitude. In (Moorthy *et al.* 1997a), it was also noticed that the peak position shifted to lower applied field strengths (decreased) when increasing the grain size. This is also expected based on the theoretical analysis given in Jiles (2000). In (Yamaura *et al.* 2001), only the Barkhausen noise measurement spectra were studied. Their significant finding was that the ratio between high and low frequency components decreased with increased grain size. They also found out that grain misorientation also has a significant influence on the measurement.

In the literature, quite controversial results are presented for the influence of the carbon content. It seems that generally an increased carbon content leads to increased Barkhausen activity (Ng *et al.* 2001, Capó-Sánchez *et al.* 2004). Ranjan *et al.* (1987) explains this by the fact that an increased carbon content leads to a decreased grain size, which further leads to increased Barkhausen activity as presented above. Indeed, Ng *et al.* (2001) studied the relationship between the carbon content and three features of the Barkhausen noise signal (the rising slope and the FWHM value of the profile, and the RMS value) and found that all the

features increased with an increasing carbon content. Capó-Sánchez *et al.* (2004) also observed a region with similar results but with higher carbon content where the maximum Barkhausen amplitude saturates and starts to decrease. Negative or positive correlation between Barkhausen activity and carbon content was also reported in (Koo *et al.* 2003) with different materials. The conclusion was that the relationship is unclear.

Ferrite, pearlite and martensite phases have been studied with Barkhausen noise. Moorthy *et al.* (2000) and Kleber *et al.* (2008) reported that the amount of ferrite may be detected on the basis of Barkhausen noise measurement. The martensite content was analysed in (Saquet *et al.* 1999) and (Kaplan *et al.* 2007) while Koo *et al.* (2003) analysed the pearlite content. Good correlations were reported especially in (Kleber *et al.* 2008), (Kaplan *et al.* 2007) and (Koo *et al.* 2003). Typically, a feature corresponding to Barkhausen activity (the RMS value or peak height) was used but also the peak position was sensitive to microstructure changes.

There are also studies where changes in microstructure have been analysed. Tempering of material leads to martensite decomposition, which changes the microstructure. These changes have been studied in (Saquet *et al.* 1999), (Moorthy *et al.* 2000) and (Davut & Gür 2007). Microstructure characterization of hardened and strained samples has been reported in (Blaow *et al.* 2006) and (O'Sullivan *et al.* 2004), respectively, while Moorthy *et al.* (1997b) analysed the microstructure changes of heat-affected zones in welding.

2.4.2 Hardness

Increased hardness is related to an increased number of pinning sites (dislocation density). This impedes the movements of dislocations and domain walls. Consequently, increased hardness is expected to lead to lower Barkhausen activity. (Cullity 1972, Sipahi 1994)

This has also been observed in studies and reported in the literature. The RMS value of the signal has been found to decrease with increasing hardness (Sipahi 1994, O'Sullivan *et al.* 2004, Santa-aho *et al.* 2009). The same observation has been made about the relationship between hardness and peak height (Moorthy *et al.* 1997b, Kaplan *et al.* 2007). Furthermore, a linear relationship between the inverse of Barkhausen noise activity (peak height or RMS) and hardness has been reported in (Moorthy *et al.* 1997b) and (O'Sullivan *et al.* 2004).

It is expected that the coercivity of the material increases with hardness (Kinser *et al.* 2005) due to the higher dislocation density. For example, O'Sullivan *et al.* (2004) reported an exponential relationship between coercivity and the inverse of Barkhausen noise. However, Mészáros & Szabó (2005) found that the coercivity value could not be used in hardness evaluation for duplex stainless steels. Instead they showed that the peak position could be used for the task. Good correlations between the peak position and hardness have also been reported in (Davut *et al.* 2004). Despite the results in (Mészáros & Szabó 2005), coercivity and the peak position are generally reported to be closely related to each other (Stewart *et al.* 2004, Davut & Gür 2007).

2.4.3 Stress state

The measure of the forces acting within a solid piece of material is called stress. Stress is defined as the force per unit area and thus its unit is expressed as N/m^2 (Pa). A division into applied and residual stress can be made. The stress due to external forces is referred to as the applied stress while the stress remaining in material after all external forces are removed is the residual stress. Residual stresses can be further divided based on their scale. Macrostress varies over large distances while microstress varies over the scale of a grain. (McGonnagle 1963, Cullity 1972, Withers & Bhadeshia 2001a, Withers & Bhadeshia 2001b)

Residual stresses arise from differences between neighbouring regions, parts or phases within the material. They may be caused, for example, by inhomogeneous plastic deformation or thermal gradients during processing. Basically, all manufactured components hold residual stresses due to processing. Residual stresses may be detrimental, leading to unexplained structural failure when combined with applied stress. However, deliberate compressive stress may improve the fatigue resistance of a component. (Withers & Bhadeshia 2001a, Withers & Bhadeshia 2001b)

A lot of results have been reported in the literature where Barkhausen noise has been used in evaluating the stress state of a material. It is typically observed that tensile stress increases Barkhausen activity and compressive stress decreases it (Jagadish *et al.* 1990, Lindgren & Lepistö 2002, Santa-aho *et al.* 2009). Mierczak *et al.* (2011), for example, reported a linear relationship between the reciprocal of the peak height and applied stress. The same linear behaviour is also reported in (Santa-aho *et al.* 2012a) but it is further noticed that the slope depends on hardness. However, the relationship shows saturating behaviour with both

tension and compression. Furthermore, it has been reported that when increasing the tensile stress after saturation is reached, Barkhausen activity starts to decrease (Anglada-Rivera *et al.* 2001, Stewart *et al.* 2004, Pérez-Benitez *et al.* 2005, Blaow *et al.* 2007).

Typically, changes in the stress state have been explained using features corresponding to Barkhausen noise activity. These features are the RMS value (Gatelier-Rothea *et al.* 1998, Lindgren & Lepistö 2002), the Barkhausen noise energy (Gauthier *et al.* 1998, Iordache *et al.* 2003) and the peak height (Stewart *et al.* 2004, Moorthy *et al.* 2004b, Blaow *et al.* 2007, Mierczak *et al.* 2011). Other features that are reported to correlate with the stress state are the peak position and width (Stewart *et al.* 2004, Blaow *et al.* 2007) and skewness (Stewart *et al.* 2004).

The prediction of residual stress is challenging because the microstructure, composition and other unmeasured properties also affect Barkhausen noise. The influences of different material properties cannot be distinguished from the signal. Also, it has been reported that the direction of the measurement has an influence on the measurement and thus has to be taken into consideration when applying Barkhausen noise measurement. To overcome these difficulties, the measurement device must be calibrated for each material in order to obtain appropriate stress state evaluations. (Gauthier *et al.* 1998, Lindgren & Lepistö 2002, Blaow *et al.* 2007)

2.5 Barkhausen noise applications

Barkhausen noise can be applied to certain NDT applications. Such applications are material deformation, fatigue, case-depth and grinding burn evaluation. The next section gives a more detailed description of these applications.

2.5.1 Deformation detection

The behaviour of material when under an applied stress varies depending on the material. The shape of the material may change or even a fracture may occur. A graphical representation of the behaviour of a material under the applied stress is called a stress-strain curve. Stress refers to the applied stress expressed as force per unit cross-sectional area of specimen (N/m^2), while strain refers to the change in the material dimension (length). Often the strain is expressed as a ratio of the change in the length and the initial length and thus it becomes a dimensionless

value (known as engineering strain). (Anderson *et al.* 1990) A typical stress-strain curve is shown in Fig. 5.

The stress-strain curve can be divided into two regions: elastic and plastic. In the elastic region, the strain increases linearly with the stress. If the stress is released within the elastic region, the material returns to its original shape. When the stress exceeds the yield strength of the material, permanent changes are caused to the material shape. Also, the rate of change decreases due to the work hardening of the material. When the applied stress exceeds the tensile strength of the material, a neck begins to develop somewhere along the specimen. After the neck is formed, any further strain is localized in the neck leading finally to the fracture of the specimen. (Anderson *et al.* 1990)

The description of material deformation under applied stress given above is rather rough. In fact, the situation is not that simple (Moorthy 1999b, Lindgren 2003). The plastic deformation of the material begins at stresses lower than the yield strength and the deformation can be divided into four stages: perfectly elastic, microyielding, macroyielding and progressive plastic deformation (Moorthy *et al.* 1999b).

Barkhausen noise is able to separate the different stages of deformation (Moorthy *et al.* 1999b, Iordache *et al.* 2003). It has been found that Barkhausen noise activity (Barkhausen noise energy and the peak height) increases significantly with tensile elastic strain (Stefanita *et al.* 2000b, Blaow *et al.* 2004, Iordache *et al.* 2003). The initiation of microyielding has been reported to decrease Barkhausen noise amplitude (Moorthy *et al.* 1999b).

It is reported in the literature that the effect of plastic deformation is anisotropic (Krause *et al.* 1995, Dhar *et al.* 2001, Lindgren & Lepistö 2001). Typically, a decrease in Barkhausen activity is reported with increased strain in the loading direction (Birkett *et al.* 1989, Lindgren & Lepistö 2001) while a decrease is noticed in the transverse direction (Lindgren & Lepistö 2001). However, some studies have shown that with early plastic deformation, Barkhausen activity has increased (Hwang & Kim 1988, Stefanita *et al.* 2000a, Dhar *et al.* 2001, Iordache *et al.* 2003). In (Dhar *et al.* 2001), an initial increase in Barkhausen noise energy was followed by a decrease and finally saturation.

Overall, the influence of plastic deformation on Barkhausen noise is very complicated and also not that well understood (Stefanita *et al.* 2000a, Lindgren 2003). Nevertheless, Barkhausen noise has been recognized as a potential method for detecting the degree of deformation within a material (Moorthy *et al.* 1999b, Dhar *et al.* 2001, Kleber & Vincent 2004). One reported advantage of Barkhausen

noise compared to other methods is that, according to the results obtained, the different stages of deformation can be detected and distinguished (Moorthy *et al.* 1999b, Iordache 2003). This information can be useful in evaluating the condition of parts and structures in service.

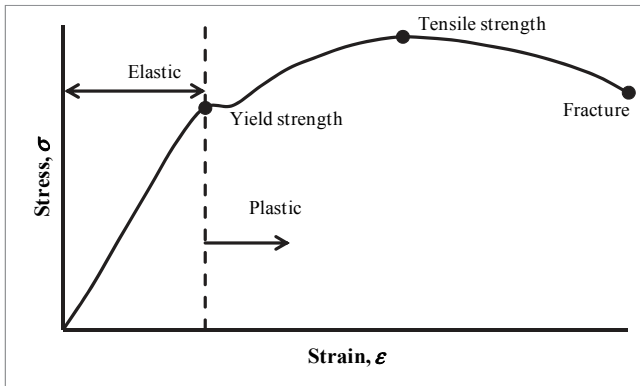


Fig. 5. A typical stress-strain curve. Redrawn from (Anderson *et al.* 1990).

2.5.2 Fatigue monitoring

Fatigue is a major cause of mechanical failure in industry. It is estimated that 90% of failures are caused by it. Fatigue basically means changes in a material when it is under a repeated load. The overall mechanism of fatigue can be divided into three phases. First, microscopic changes develop leading to crack initiation. This is followed by crack growth and finally by the fracture of the material. (Schijve 2001, Palit Sagar *et al.* 2005)

Many factors affect material fatigue. These factors are, for example, the initial microstructure, surface quality, the stress state of the specimen and the applied stress level (Schijve 2001). It has been shown in the literature that the magnetic properties of materials are related to fatigue-induced changes in the microstructure. Bi *et al.* (1997) found that coercivity and remanence change systematically with the number of stress cycles while Chen *et al.* (1994) found a logarithmic relation between coercivity and stress cycles. The changes in magnetic properties depend on the stage of the fatigue process. It has been reported that coercivity increases or decreases in the first stage depending on the material and remains stable in the second phase (Lo *et al.* 2000). Remanence, on

the other hand, is reported to increase in the first stage, remain stable in the second stage and decrease in the third stage (Lo *et al.* 2000).

Barkhausen noise measurement has also been proposed as a possible method for fatigue monitoring (Tomita *et al.* 1996, Vincent *et al.* 2005). It is typical that a nonlinear relationship is observed between Barkhausen noise and the number of stress cycles (Błachnio *et al.* 2002). This is quite natural because fatigue mechanisms in different stages affect Barkhausen noise in different ways. Furthermore, the factors affecting fatigue are multiple as listed above and along with microstructural changes the stress state of the material also changes during the fatigue process, which complicates the interpretation of the measurement signal (Lindgren & Lepistö 2003).

It has been reported that the different stages of fatigue can be distinguished using Barkhausen noise measurement (Moorthy *et al.* 1999a, Vincent *et al.* 2005). An increase in Barkhausen activity has been reported in the first stage of fatigue. This is followed by a decrease in the second stage and a further increase in the final stage. These observations have been justified by the changes in the microstructure caused by movements of dislocations. (Govindaraju *et al.* 1993, Lindgren & Lepistö 2003, Palit Sagar *et al.* 2005)

Even though similar results have been reported, it should be noted that the results are very case-dependent. For example, Lindgren & Lepistö (2000) found that the relationship is quite different for mild and high strength steel specimens. Palma *et al.* (2005) showed that the applied stress level affects the results. They observed no changes in Barkhausen noise if the applied load was below the fatigue limit and found Barkhausen noise inapplicable for fatigue monitoring where the applied load exceeded the yield strength (see Fig. 5) of the material.

2.5.3 Case depth evaluation

The fatigue life of materials depends on their surface condition. Consequently, materials are given surface treatments such as carburizing and induction hardening. In these processes, the chemical, mechanical and metallurgical properties of the components are changed to a certain depth below the surface. The knowledge about this case depth and the profiles of the aforementioned properties is advantageous for evaluating the quality life of components. (Vaidyanathan *et al.* 2000, Moorthy *et al.* 2004a)

The detection of the case depth with Barkhausen noise measurement is based on the hardness difference of the surface (case) and the interior of the material. It

is typical that two-peaked Barkhausen noise profiles are obtained with the first peak at lower applied field strength corresponding to the soft interior and the second peak corresponding to the case. When two peaks are obtained their ratio has been shown to correlate with the case depth (Bach *et al.* 1988, Vaidyanathan *et al.* 2000, Moorthy *et al.* 2004a).

The problem, however, is that two peaks are not always detected (Dubois & Fiset 1995, Moorthy *et al.* 2004a). Moorthy *et al.* (2004a) noticed that the second peak disappears if the case depth is too great. They showed that case depths up to 1 mm show the two-peaked profile and can be evaluated with the ratio of the two peak heights. Dubois & Fiset (1995) used the frequency spectrum instead and calculated the frequency content of a certain frequency bands. They found good correlations between the frequency content and case depth with many steel types.

Other features are also shown to correlate with the case depth. Wilson *et al.* (2009) showed that the peak height decreases with an increasing case depth. Santa-aho *et al.* (2010) observed the same relationship between the RMS value and the case depth. The same research later showed that the spectral density of the Barkhausen noise signal decreases exponentially with increasing case depths (Santa-aho *et al.* 2011). It was reported that case depths of up to 2 mm could be detected (Santa-aho *et al.* 2010, Santa-aho *et al.* 2011).

A quite different approach for the detection of the case depth is taken in (Santa-aho *et al.* 2012c). They studied voltage and frequency sweeps to evaluate the thickness of the surface layer. Usually these sweeps are only used for determining suitable voltage and frequency values for current measurements. In their study, it is observed that the ratio of two voltage sweeps with different frequencies provides a good estimation of the case depth.

Hardened surface layers increase the fatigue strength of the material. It is also possible that the surface undergoes a decarburizing process where the surface layer softens. The softening is detrimental in terms of the fatigue life of components. A couple of studies have shown that Barkhausen noise can also be used to detect the depth of this softened layer (Blaow *et al.* 2005, Stupakov *et al.* 2011). In (Blaow *et al.* 2005), it is reported that a single-peaked profile turns into a two-peaked one in decarburization. They observed that the height of the second peak at higher applied field strengths increased with increasing depths. Also, the position of the peak shifted to lower field strengths. In (Stupakov *et al.* 2011), it is shown that the RMS value first increases followed by a decrease with the increased depth of the decarburized layer. Therefore, they used Barkhausen noise coercivity and found that it decreases with increasing depths.

2.5.4 Grinding burn detection

The surface finishing of a component is accomplished in the grinding process. Grinding also removes a small amount of material so that the required dimensions are obtained. The processing step may, however, damage the surface. The damage may be temper (softening) or rehardening burns, tensile residual stress generation, or microcracking. The damage is caused by increased localized temperatures. (Malkin 1989)

Several studies have shown that Barkhausen noise can also be used in detecting grinding burns. The peak height and width have been shown to indicate grinding burns (Gupta *et al.* 1997, Moorthy & Shaw 2009). Moorthy *et al.* (2005) used low and high frequency measurements to detect grinding burns. High frequency measurement produced a single-peaked Barkhausen profile while low frequency measurement led to a two-peaked profile. They concluded that the peak height of the high frequency profile correlated well with the residual stress changes due to grinding damage. The second peak (higher applied field strength) of the low frequency profile was shown to reflect the changes in stresses and microstructure near the surface.

Some studies have used laser processing to produce artificial grinding burns in order to manufacture calibration blocks for industrial grinding burn detection (Sorsa *et al.* 2010a, Santa-aho *et al.* 2012b, Santa-aho *et al.* 2012d). These studies showed that laser processing could be applied in generating artificial grinding burns. However, the processing parameters must be carefully selected. The studies also showed that the changes in the material properties due to the processing correlated well with the changes in the Barkhausen noise features.

3 Variable selection and model identification

The amount of data in many fields of applications has increased dramatically in recent years. Such fields include chemometrics, quantitative structure-activity/property relationship (QSAR/QSPR) studies and sensor array applications. Obviously, huge amounts of data include irrelevant variables. Some variables may be noisy and some may be collinear with each other. As far as modelling is concerned, it is advantageous in many ways to select the significant variables from the whole set. If irrelevant, noisy or collinear variables are used in models it has been reported to lead to, for example,

- the deterioration of model performance (Alexandridis *et al.* 2005)
- increased time consumption in model training (Guyon & Elisseeff 2003)
- a more difficult interpretation of the developed model (Smit *et al.* 2008).

Also, the risk of overfitting and chance correlations increases (Baumann 2003, Dieterle *et al.* 2003). The variable selection task is a high dimensional optimization problem where the variable set is explored in order to find the most suitable subset. The usually exhaustive search which is guaranteed to find the optimal solution is not feasible because it is computationally too expensive. Consequently, many procedures have been proposed that are capable of selecting a subset of variables from a large variable set. They can be divided roughly into filters and wrappers (Kohavi & John 1997). In filter approaches, some ranking is formed and features are selected or eliminated according to that ranking. In wrapper approaches, the model itself is used in the selection. Filter approaches are computationally very efficient but the obtained model is seldom optimal (Guyon & Elisseeff 2003, Sayes *et al.* 2007).

The wrapper approaches need three components: a modelling technique, a search engine and an objective function (Baumann 2003). Typical modelling techniques are, for example, multivariate linear regression (MLR), principal component regression (PCR), partial least squares regression (PLSR) and artificial neural networks (ANN) and the objective function is typically based on the squared error of prediction. The search engines can be further divided into optimal, deterministic and stochastic. The optimal search engines are considered impractical but are still briefly described here. The deterministic methods are computationally more efficient but they may get trapped in local optima (Guyon & Elisseeff 2003, Sayes *et al.* 2007). Fig. 6 shows the classification of the search engines. The figure also shows the methods that are described later in this section.

This section presents the basic principles of feature selection. Firstly, some basic concepts of filter approaches are given although the main focus is on wrapper approaches. The only filter method that is described in more detail is the successive projections algorithm (SPA), which has been used in studies by the present author. Each aspect of the wrapper approaches is discussed separately starting with mathematical modelling techniques.

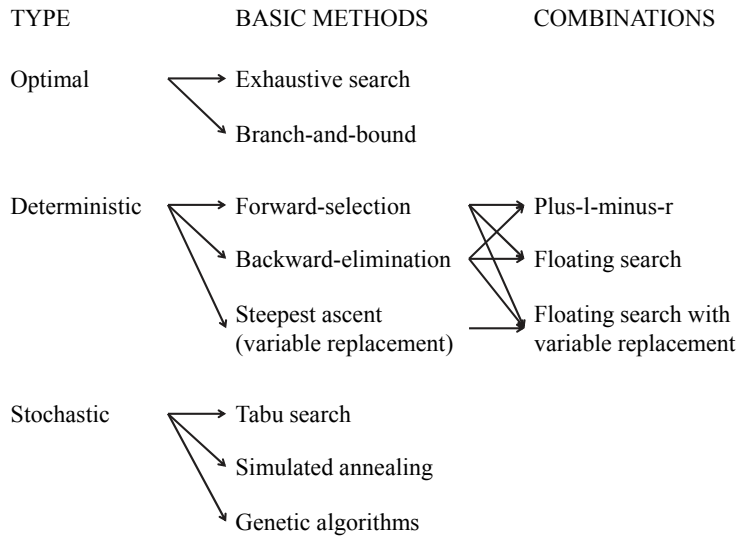


Fig. 6. The classification of search engines.

3.1 Filter methods

Fig. 7 shows a flowchart of a filter approach. The flowchart in the figure assumes that the data is split statically into training and external validation data sets. This issue is further discussed in section 3.6. In the filter methods, some index describing the relevance of the variables is calculated. The indices are then typically used to rank the variables or to remove the variables with low relevance. The basic idea in ranking approaches is that variables with a high rank are added to the model while low-rank variables can be eliminated from the set of candidate variables. (Guyon & Elisseeff 2003, Sayes *et al.* 2007)

A usual ranking criterion is correlation, which, however, is limited to linear correlations. Another common criterion is some measure of predictive power. It is

obtained by fitting a simple (typically linear regression) model between each one of the variables and the predicted output. The modelling error is then used as the criterion. Information theoretic criteria have also been used where the criterion is typically mutual information between the input and output variables. The benefit of this is that no assumptions about the underlying distributions or the nature of the relationships between the variables are made. (Guyon & Elisseeff 2003, Sayes *et al.* 2007)

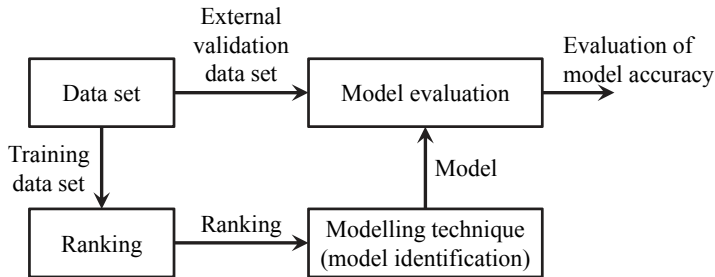


Fig. 7. Flowchart of a filter variable selection.

The successive projections algorithm (SPA) is introduced in (Araújo *et al.* 2001). The basic idea of the SPA algorithm is to form a ranking so that, when adding features to the model, each new feature has minimal collinearity with the features already added. This ranking is formed by calculating the orthogonal projections of the remaining features to the already selected features. The greatest Euclidean length of the projections then decides which feature is added next. The first feature may be selected manually (Araújo *et al.* 2001) or M (M is the number of features) different rankings may be formed starting from each of the features (Galvão *et al.* 2008). In the latter case, all the rankings are then tested and the best one is selected. Furthermore, the selected subset of features can be subjected to a backward-elimination procedure (Galvão *et al.* 2008). The details of the SPA and required equations are not presented here but can be found in (Araújo *et al.* 2001) and (Galvão *et al.* 2008).

The filter methods and the ranking criteria are not described here in more detail. Many applications can be found in the literature. For example, correlation-based feature selection is discussed in (Hall 1999) and information theoretic criteria used in (Peng *et al.* 2005). A good review of the filter methods applied in the field of bioinformatics is given in (Sayes *et al.* 2007).

3.2 Mathematical modelling techniques

As mentioned before, wrapper selection procedures need three components. A flowchart of the wrapper approach is shown in Fig. 8. As illustrated, one of the required components is a mathematical modelling technique. The most usual modelling techniques used are briefly introduced in this section.

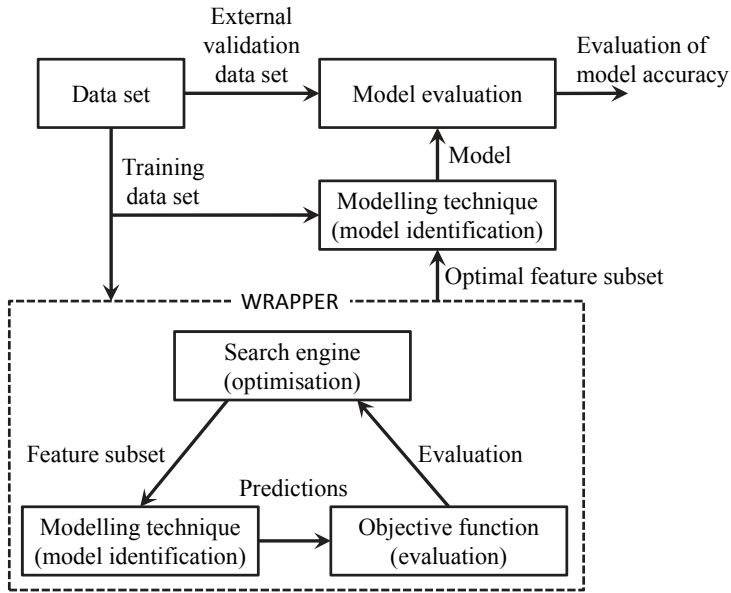


Fig. 8. Flowchart of a wrapper variable selection.

3.2.1 Multivariable linear regression

Linear regression models are used to map the linear interactions between input and output variables. Based on the collected data, the unknown model parameters are identified. A linear regression model is given by (Harrell 2001)

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_Mx_M = b_0 + \sum_{j=1}^M b_jx_j. \quad (2)$$

In (2), y is the output variable that is predicted, x_j are the input variables and b_j are the unknown model parameters. The parameter b_0 accounts for the bias. When there is only one input variable ($M = 1$), the model is considered to be a simple

regression model, but when $M > 1$ it is called a multiple regression model (MLR). The general multivariable regression model can be presented in matrix format as (Harrell 2001)

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}, \quad (3)$$

where \mathbf{X} is the matrix of input variables ($N \times M$) and \mathbf{b} is the matrix of regression coefficients ($M \times 1$). N and M are the number of data points and input variables, respectively. It should be noticed that the bias term is included in (3) as the vector x_0 , which is a vector of ones. It should also be noticed that only one output variable is written in (2) and (3) even though there may be more. Typically, a least squares solution of (3) gives the unknown model parameters. The least squares solution is based on the pseudoinverse of \mathbf{X} and is given by (Harrell 2001)

$$\hat{\mathbf{b}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}. \quad (4)$$

Many studies found in the literature use MLR models together with some wrapper selection procedure. Pires *et al.* (2008) studied simple forward-selection and backward-elimination procedures while a genetic algorithm was used in (Jouan-Rimbaud *et al.* 1996) and (Broadhurst *et al.* 1997). These studies are just examples and a lot more can be found in the literature.

3.2.2 Principal component regression

Principal component analysis (PCA) is a statistical technique usually used for data reduction. The data matrix is transformed into a new matrix so that the variables in the new matrix are uncorrelated and ordered on the basis of the portion of the variance they explain. The new variables are called principal components and are obtained from the covariance or correlation matrix of the original data matrix. Usually the first few variables explain the majority of the variance and thus they can be used to represent the whole data set without losing a significant amount of information. (Roffel & Betlem 2006)

The vectors containing the values of the new variables are called scores. They are obtained by multiplying the original variables by their weighting coefficients, which are called loadings. Thus a score vector is obtained by

$$\mathbf{t}_j = \mathbf{X}\mathbf{p}_j, \quad (5)$$

where \mathbf{t} and \mathbf{p} are the score and the loading vectors, respectively. The loading vectors are the eigenvectors of the covariance or correlation matrix of the original data matrix. The decomposition of the original data matrix into the matrices of scores and loadings can be expressed as (Roffel & Betlem 2006, Romagnoli & Palazoglu 2006)

$$\mathbf{X} = \mathbf{TP}^T . \quad (6)$$

In (5), \mathbf{P} is the loadings matrix ($M \times M$) including the eigenvectors of the correlation (or covariance) matrix of \mathbf{X} and \mathbf{T} is the score matrix ($N \times M$). The transformation is carried out so that each loading vector is a unit vector meaning that its Euclidean length is 1. The significance of the transformation is that the obtained component score vectors are uncorrelated. In addition, they are formed in such a way that the first score vector explains the maximum amount of the variance in the original data matrix. The second score vector explains the maximum amount of the remaining variance and so on. An illustrative figure of PCA is given in Fig. 9, where the components are shown for a two-dimensional case. The figure clearly shows how the first component captures the majority of the variance while the second component explains the remaining variance. The variance explained by each component is defined by the eigenvalue corresponding to the component's eigenvector (Romagnoli & Palazoglu 2006).

In addition to the uncorrelated component score vectors, another significant outcome of the principal component transformation is that typically the majority of the variance is explained by a number of components that is less than the original number of variables. Thus, the dimensionality of the modelling problem decreases. This dimensionality reduction is usually also the aim of the studies found in the literature where PCA is used (Barros & Rutledge 1998, Zhang 2007). Studies also exist where PCA is used to produce uncorrelated input variables for other modelling techniques such as ANNs (Zhang 2007).

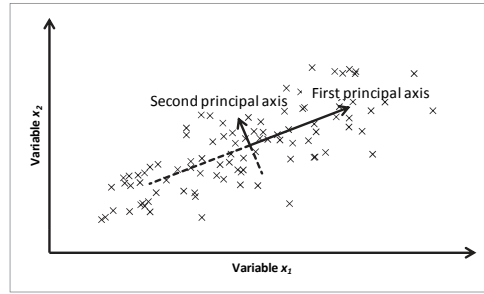


Fig. 9. A two-dimensional example of PCA.

The regression model that is identified with the component score vectors as input variables is called a principal component regression (PCR) model. It is given by

$$\hat{\mathbf{y}} = \mathbf{T}'\mathbf{b}, \quad (7)$$

where \mathbf{T}' is the matrix of component scores ($N \times m$) including only the selected number (m) of vectors ($m < M$). The model is identified in the least squares sense by (4). The number of component score vectors included in the model is usually defined by adding them one at a time to the model as long as the model behaviour improves. The model behaviour is typically evaluated by the prediction error of some validation data set (Depczynski *et al.* 2000). The score vectors are ranked according to their eigenvalues and added starting from the highest.

The traditionally used eigenvalue-based component ranking, however, has been reported as not leading to optimal results (Sun 1995, Depczynski *et al.* 2000). It has been shown that even though some variables explain only a small fraction of the total variance, they may contain significant information and thus cannot be neglected. Thus, a correlated PCR has been proposed where the correlation of component score vectors to the predicted output variable is used as the ranking criterion (Sun 1995, Depczynski *et al.* 2000). Furthermore, it has been suggested that no ranking at all be used but that the component score vector selection be subjected to a wrapper selection procedure. The proposed wrapper procedures use forward selection and backward elimination (Depczynski *et al.* 2000, Pires *et al.* 2008) and genetic algorithms (Barros & Rutledge 1998, Depczynski *et al.* 2000) as search engines. These search engines are explained later in Sections 3.4 and 3.5.

3.2.3 Partial least squares regression

Partial least squares regression (PLSR) is quite similar to the PCR described in the previous section. The original data matrix is also transformed into component score vectors and the regression model is built between these transformed vectors and the predicted variable as given in (7). However, the mechanism used in the transformation is different. While PCA only considers the internal variance of the data matrix, PLSR also considers the predicted output variable and forms the component score vectors so that the maximum amount of useful information is captured from the data matrix. The required number of vectors is usually lower than the number of original variables and thus PLSR is also used to decrease the dimensionality of the modelling task. (Garthwaite 1994, Höskuldsson 2001)

The calculation of PLS component score vectors is an iterative procedure where a desired number of vectors is formed. It is typical that the score vectors are calculated and added to the model until the model performance improves (Garthwaite 1994). The score vectors are obtained by identifying the corresponding loadings vector and then substituting them in (5). The first PLS component loadings vector is obtained by identifying a simple regression model between each of the input variables and the predicted output variable. Mean-centred or normalized variables are used and thus the regression coefficients are obtained from the covariance or correlation matrix, respectively. These regression coefficients form the first loadings vector so that the Euclidean length of the vector is 1. (Bastien *et al.* 2005)

The identification of the first loading vector and the calculation of the first score vector is followed by a set of simple regressions. In these regressions, both the input variable matrix and the output variable are predicted with the first score vector. The identification of the regression coefficients between these are obtained with the least squares solution given in (4) and the predictions are (Bastien *et al.* 2005)

$$\hat{\mathbf{X}} = \mathbf{T}_1 \mathbf{a}_1 \text{ and} \quad (8)$$

$$\hat{y} = \mathbf{T}_1 b_1. \quad (9)$$

Above, \mathbf{a}_1 is the vector of regression coefficients ($1 \times M$) between the data matrix \mathbf{X} and the first score vector \mathbf{T}_1 and b_1 is the regression coefficient between the output variable y and the first score vector. It should also be observed that b_1 is the first element of \mathbf{b} in (7). After the predictions given in (8) and (9), the prediction errors (residuals) are calculated by (Bastien *et al.* 2005)

$$\mathbf{E}_{x,1} = \hat{\mathbf{X}} - \mathbf{X} \text{ and} \quad (10)$$

$$\mathbf{e}_{y,1} = \hat{\mathbf{y}} - \mathbf{y} . \quad (11)$$

Now the residuals hold no correlation to the first score vector \mathbf{T}_1 and are used in identifying the second loadings vector. The procedure of identifying the second loadings vector is the same as the one described above but the original data matrix is replaced by $\mathbf{E}_{x,1}$ and the original output variable by $\mathbf{e}_{y,1}$. The same procedure is repeated until the desired number of score vectors is formed. (Bastien *et al.* 2005)

It has been reported in the literature that the prediction accuracy of a PLSR model is influenced by the variables used in generating the component score vectors (Baumann 2003, François *et al.* 2007). Consequently, the selection of the variables is essential before the identification of a PLSR model. It is also possible to set the loadings of the insignificant components to zero (Bastien *et al.* 2005). Many studies can be found in the literature where the variables are selected for PLSR models. The results are reported, for example, in (Leardi & Gonzáles 1998), (Andersen *et al.* 2006) and (McLeod *et al.* 2009). Andersen *et al.* (2006) and McLeod *et al.* (2009) use backward-elimination procedure while Leardi & Gonzáles (1998) use a genetic algorithm for the selection. Also, McLeod *et al.* (2009) use a genetic algorithm but only as a pre-selection step.

3.2.4 Artificial neural networks

An artificial neural network is a set of simple, parallel computational units, neurons (Fig. 10). The inputs of the neurons are weighted with the weighting coefficients (w_j) as shown in Fig. 10. The weighted inputs are typically summed and then passed through the activation function (f) to produce the output of the neuron (y). Fig. 11 shows a couple of often used activation functions. The sigmoid function is used in backpropagation neural networks (BPNN) while radial-basis functions are used in radial basis function neural networks (RBFNN). (Haykin 1999)

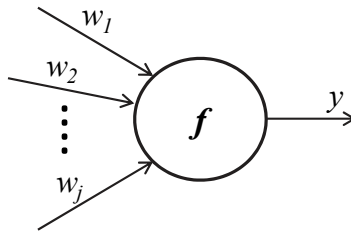


Fig. 10. A simple computational unit, neuron.

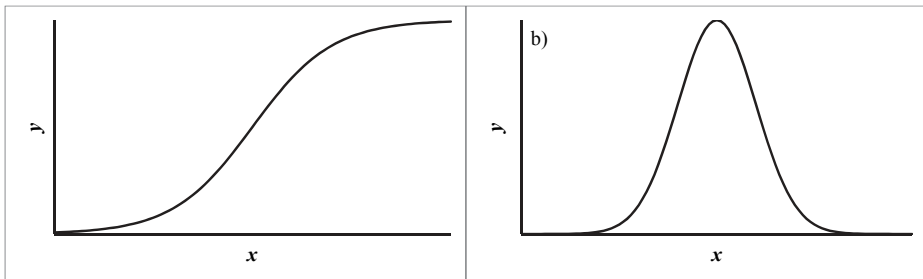


Fig. 11. Activation functions used in neural networks: a) sigmoidal and b) radial-basis function.

Neurons are arranged in layers. The networks include an input and an output layers and a fixed number of so-called hidden layers as shown in Fig. 12. No computations are carried out in the input layer as the input variables are only distributed to the neurons in the first hidden layer. Thus, the number of neurons in the input layer equals the number of input variables (N). The output layer obviously produces the outputs of the neural model. Consequently, the number of neurons in the output layer equals the number of output variables (P). (Haykin 1999)

The complexity of the network is governed by the number of hidden layers and the number of neurons in each hidden layer. Too many hidden layers and neurons result in overfitting while too simple a model structure results in poor prediction accuracy. Usually, one hidden layer is enough for the majority of the modelling tasks. (Babuska & Verbruggen 2003)

After the structure of the network is defined, it is trained. In training, the unknown parameters are defined. The training procedure used depends on the network type. In BPNNS, the parameters are found by first simulating the network (calculating the output) and then passing the error back to the input layer and

tuning the parameters starting from the output layer (Haykin 1999). With the RBFNNs, the training can be carried out through clustering or by adding one neuron at a time until a predefined number of neurons is reached. The applied clustering techniques are k-means and fuzzy c-means. (Wang & Xiang 2007)

Automated variable selection strategies are applied to ANN models. Forward selection was used, for example, in (Eklöv *et al.* 1999) and (Benoudjit *et al.* 2004). In the latter study, the forward-selection step was followed by backward elimination in order to find a variable subset as small as possible. Alexandridis *et al.* (2005) and Gualdron *et al.* (2006) used simulated annealing in variable selection while a genetic algorithm was used in (Gardner *et al.* 2005) and (Hemmateenejad 2005).

When variables are selected for ANN models, the problem is that the parameters defining the network structure must also be optimized and the network must be trained (Benoudjit *et al.* 2004, Wang & Xiang 2007). Thus, the overall procedure easily becomes computationally very expensive. Naturally, it is usual that the computational burden is decreased somehow. It is, for example, possible to use a linear model in a variable selection step instead of the nonlinear ANN model (Eklöv *et al.* 1999, Guyon & Elisseeff 2003). Also, the dimensionality of the problem can be decreased by a pre-selection step (Gualdron *et al.* 2006, Wang & Xiang 2007). The PCA described in Section 3.2.2 has also been applied to the input variable set followed by a stepwise addition of variables based on the eigenvalue (Zhang 2007) or correlation ranking (Hemmateenejad 2005).

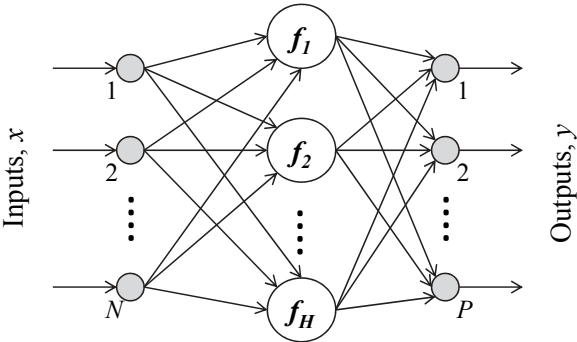


Fig. 12. Typical neural network structure with one hidden layer.

3.3 Optimal search engines

Two optimal search algorithms exist as shown in Fig. 6. These are the exhaustive search and the so-called branch-and-bound algorithm. The exhaustive search simply tests all the possible feature subsets in order to find the best one. The computational load increases exponentially with the increasing number of variables in the original subset (M) and becomes impractical with a rather small M . The branch-and-bound approach is computationally more feasible because it does not search the entire search space, but parts that are guaranteed not to lead to optimal results are rejected. The major drawback is that the objective function must be monotonic. This is not often the case especially with small sample sizes. Also, the complexity of the algorithm explodes with increasing M making it also impractical. (Jain & Zongker 1997, Nakariyakul & Casasent 2007)

Many improvements have been proposed for the branch-and-bound algorithm to enhance the computational speed. They are not presented here but can be found for example in (Nakariyakul & Casasent 2007).

3.4 Deterministic search engines

Deterministic wrapper methods are step-wise algorithms where one feature at a time is added, eliminated or replaced as shown in Fig. 6. These basic methods have been combined to obtain more efficient search engines. The deterministic methods are computationally more efficient than the stochastic methods and the results are reproducible.

The forward-selection procedure is very simple. It starts by identifying M models each having only one variable. The performance of these models is evaluated with the objective function and the model leading to the best performance determines the variable that is added into the model. The second step is to identify $M-1$ models so that each model contains the already selected variable and one of the remaining variables. The second variable is again added based on the behaviour of the models. This procedure is repeated as long as the model performance improves. The drawbacks of the procedure are that it easily gets trapped in a local optimum and only considers the effect of a single variable at a time. Thus, variables which are weak alone but strong together are rejected. Furthermore, once the variable is added its relevancy is not questioned later. (Kohavi & John 1997, Guyon & Elisseeff 2003)

In the backward-elimination procedure, the opposite is carried out. Firstly, a model with all the variables is identified followed by the identification of M models each with $M-1$ variables. The performance of these models is evaluated and the best model determines the eliminated feature. Then $M-1$ models each with $M-2$ variables are identified and evaluated to eliminate another variable. The same elimination procedure is continued for as long as the model performance improves. The backward elimination algorithm requires more computations than forward selection and is also prone to getting trapped in local optima (Kohavi & John 1997, Guyon & Elisseeff 2003)

A computationally intriguing variant of the above-mentioned procedures is their generalized version where multiple variables are added or eliminated simultaneously. However, such approaches have been reported to have what is called a “nesting problem” (Jain & Zongker 1997, Kudo & Sklansky 2000). This problem is avoided if the plus- l -minus- r method is used (Pudil *et al.* 1994, Eklöv *et al.* 1999). As expected, the procedure takes l steps forwards followed by r backward steps. The drawback is, however, that there is no method for determining the optimal l and r values.

A more advanced variant of the simple procedures is the sequential floating search method presented in (Pudil *et al.* 1994). The procedure contains repeated forward and backward steps and thus it is assumed that the method is able to escape local optima. Floating search has been reported to give better results than the other approaches mentioned above (Pudil *et al.* 1994, Jain & Zongker 1997). An adaptive variant of the floating search method is proposed in (Somol *et al.* 1999) where more than one variable can be eliminated in one step. The number of variables added or eliminated changes in the course of the search.

The steepest ascent algorithm is proposed in (Serpico & Bruzzone 2001). This algorithm starts with forward selection where the desired number of variables is selected. Then it is evaluated if any variable selected can be changed to improve the model performance. The drawback of this procedure is that it has to be repeated with different numbers of variables. Based on the steepest ascent algorithm, an improvement for the floating search algorithm is proposed in (Nakariyakul & Casasent 2009). This improvement adds a replacement step to the floating search procedure so that each selected variable is subjected both to elimination and replacement.

There are many reported applications of deterministic variables selection methods in the literature. The following section contains some examples. Forward selection has been used, for example, in (Xie & Kalivas 1997) and (Barros &

Rutledge 1998) to find the optimal variable subset for a PCR model while in (Paulsson *et al.* 2000) variables for a feedforward neural network model are sought. Backward elimination is seldom used alone but usually as a part of a more complex procedure. Such applications are reported in (Andersen *et al.* 2006), (Benoudjit *et al.* 2004a) and (Benoudjit *et al.* 2004b). It is also quite typical that simple forward-selection and backward-elimination procedures are carried out as a second selection step after some prior selection step. For example, McLeod *et al.* (2009) applied forward selection and Jouan-Rimbaud *et al.* (1996) backward elimination to a variable set obtained from a genetic algorithm selection.

3.5 Stochastic search engines

The deterministic search engines described in the previous section have some drawbacks. Firstly, they may be trapped in local optima and thus the result may not be optimal. Also, the variables are often evaluated independently and thus it is possible that variables which are weak alone but strong together with some other variable are not found. Thus, it may be desirable to use stochastic search methods which are more likely to find the global optimum. (Guyon & Elisseeff 2003)

Two stochastic search engines are mainly used in variable selection, simulated annealing and genetic algorithms. Another but quite seldom used algorithm is the tabu search (Baumann *et al.* 2002, Wang *et al.* 2009), although it is not described here in more detail.

3.5.1 Simulated annealing

Annealing is a process where a solid piece of material is first heated up and then cooled down slowly. With high temperatures the atoms within the material can arrange themselves quite freely. When the temperature is lowered, the movements of the atoms are hindered and finally the solid crystallizes into a permanent structure. Throughout cooling, the system energy is reduced and reaches the minimum when the structure crystallizes. The probability distribution of possible system energies follows the Boltzmann distribution given by (Pham & Karaboga 2000)

$$p(E) = \exp\left(-\frac{E}{kT}\right). \quad (12)$$

Above, $p(E)$ is the probability that the system is in a state of energy E , k is the Boltzmann's constant and T is the temperature. Fig. 13 shows probability distributions with different temperatures. The figure shows that at high temperatures all the system energies have probabilities close to one, but at lower temperatures the probabilities of higher energy states are reduced. (Pham & Karaboga 2000)

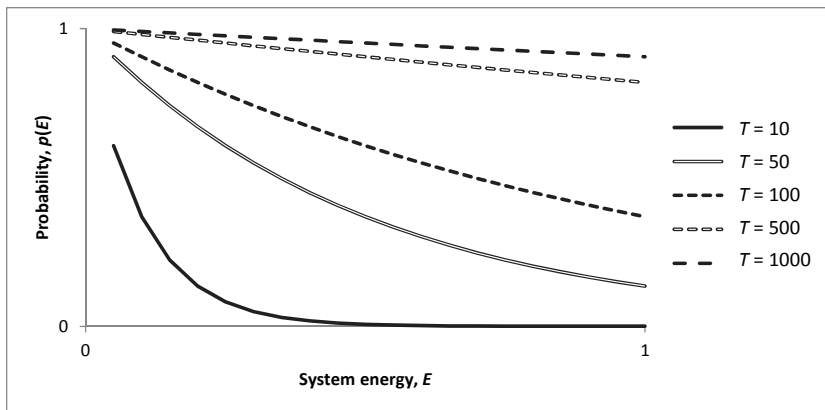


Fig. 13. Probability distributions at different temperatures ($k = 0.01$).

Procedure

This process of cooling and the subsequent crystallization of a solid is mimicked in simulated annealing. The state of the solid is represented by the possible solutions to the optimization problem and the system energy is replaced by the value of the objective function. The minimum energy state at the end of cooling corresponds to the optimal solution found by the algorithm. Fig. 14 shows a flowchart of a typical simulated annealing algorithm.

The algorithm is iterative and searches the neighbourhood of the current solution for better ones. With each iteration, a new solution is created by making a random modification to the current solution and the new solution is then evaluated with the objective function. The difference between the objective function values determines whether the new solution is accepted or not. If the difference is negative (i.e. the new solution is better), the new solution is accepted and the current solution is replaced by the new one. If the difference is positive (i.e. the new solution is not better), the new solution is accepted according to the

Metropolis criterion, which is based on Boltzmann's distribution given in (12). The new solution is accepted if (Pham & Karaboga 2000)

$$\delta \leq \exp\left(-\frac{J_{new} - J_{cur}}{T}\right), \quad (13)$$

where δ is a uniform random number between 0 and 1 and J_{new} and J_{cur} are the objective function values of the new and current solutions, respectively. Also, a general version of simulated annealing can be used. The new solution is then accepted if (Sutter & Kalivas 1993)

$$\delta \leq \exp\left(-\beta \frac{J_{new} - J_{cur}}{J_{bsf} - J_{cons}}\right). \quad (14)$$

Above, β is a constant, J_{bsf} is the objective function value of the best solution found so far and J_{cons} is a conservative guess of the minimum objective function value, respectively. A suitable value for β is reported to be such that the ratio of accepted and rejected detrimental moves is initially somewhere between 0.5 and 0.7 (Hörchner & Kalivas 1995).

Using simulated annealing

When using simulated annealing, a couple of selections must be made. These selections are (Pham & Karaboga 2000)

- representation of solutions
- definition of the objective function
- definition of the generation mechanism of new solutions
- definition of the cooling schedule.

The representation of the solutions may be binary, integer or real-valued. These possibilities are described in more detail in Section 3.5.2. The definition of the objective function is described in Section 3.6. The generation mechanism of new solutions makes a very small change in the current solution. It is quite typical that a small change is made only to one of the parameters to be optimized. Some mechanisms used with genetic algorithms (mutation) described in the following section may be used. (Pham & Karaboga 2000)

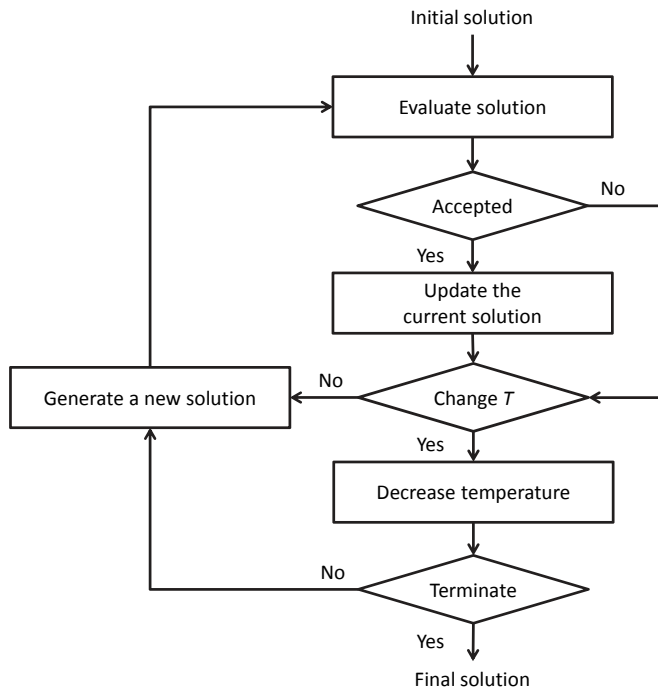


Fig. 14. A flowchart of a typical simulated annealing algorithm.

The cooling schedule needs four parameters to be defined: the initial temperature, the temperature update rule, the number of new solutions generated at each temperature and the stopping criterion for the search (Pham & Karaboga 2000). It is desirable for the cooling schedule to be such that initially almost all the new solutions are accepted and thus the search space is thoroughly covered. Later when the temperature is lowered, a lesser amount of detrimental moves are allowed and thus the search starts to converge to an optimum. Still, some detrimental moves are allowed and thus the algorithm is still able to escape local optima. When the temperature is low enough, practically no detrimental moves are allowed and the algorithm can converge to the global optimum. (Alexandridis *et al.* 2005) Typically, a stepwise cooling update rule is used. Such a rule is, for example, the geometric cooling rule given by (Pham & Karaboga 2000)

$$T_{i+1} = cT_i, \quad (15)$$

where c is a constant factor.

Simulated annealing has been used in many variable selection applications. For example, Sutter & Kalivas (1993) and Sutter *et al.* (1995) use simulated annealing to find a suitable subset of features for an MLR model applied in QSAR studies. A suitable feature subset for an MLR model is also sought in (Meiri & Zahavi 2006). They use the model in a marketing application. Swierenga *et al.* (1998) used simulated annealing to find the optimal parameters for a PLSR model while some applications where neural networks are used are presented in (Sutter *et al.* 1995), (Alexandridis *et al.* 2005), (Gualdrón *et al.* 2006) and (Llobet *et al.* 2007). Sutter *et al.* (1995) identified feedforward neural networks, Alexandridis *et al.* (2005) identified RBFNNs and Gualdrón *et al.* (2006) identified probabilistic neural networks (PNN).

When using simulated annealing in variable selection, it has been noticed that the obtained result is not stable but changes when the search procedure is repeated (Hörchner & Kalivas 1995). Thus, many studies suggest the use of some pre-selection procedure. Filter methods are the most usual methods used for this (Debus & Rayward-Smith 1997, Llobet *et al.* 2007). In (Hörchner & Kalivas 1995), simulated annealing itself is used in removing the irrelevant features before the actual selection, while Alexandridis *et al.* (2005) used genetic algorithms for this task.

3.5.2 Genetic algorithms

Genetic algorithms are a class of robust evolutionary optimization methods suitable for various types of optimization problems. The basic idea comes from biological evolution. During the development of genetic algorithms, new operators inspired by nature are added even though their significance is not clear. It has been assumed that if a certain operator has been favoured in nature, it must be advantageous. Crossing methods, for example, have been generated and selected on this basis. (Goldberg 1989)

A genetic algorithm is based on a population of chromosomes, each being a possible solution to the optimization problem. New populations are repeatedly generated to allow the population to evolve towards better solutions through genetic operators: reproduction and mutation. Reproduction further divides into selection and crossover and is basically responsible for convergence, while mutation allows the search to escape local minima. A flowchart of a typical genetic algorithm is shown in Fig. 15. (Goldberg 1989)

The algorithm starts from the initial population. It should be noted that the time needed to find a good solution depends directly on the initial population (Rahnamayan *et al.* 2007). When no a priori information about the solution is available, it is typical to take the initial population randomly from a uniform distribution. Another initialization method uses a quasi-random sequence of points (Maaranen *et al.* 2007) and opposite initialization has also been proposed (Rahnamayan *et al.* 2007). Population initialization requires the definition of the first tuning parameter: the population size (N_{pop}). This is simply the number of chromosomes used. When binary-coded algorithms are used the number of bits (N_{bits}) used per parameter must also be defined.

The algorithm continues with the evaluation of the initial population. The evaluation is carried out with the objective function. After the evaluation, a new population is created. First, an adequate number of parents are selected. For each crossover a random number is taken and compared to the predefined crossover probability (p_c). If the random number is smaller than the probability, the parents are crossed with the selected crossover method. Otherwise, the parents are moved directly to the new population. Thus, the crossover probability regulates the crossover rate. (Goldberg 1989)

The crossover is followed by mutation where random changes are made to the population. The mutation rate is regulated using the mutation probability (p_m). Mutation occurs if a random number is smaller than the mutation probability. The final step in creating the new population is elitism. New populations are created until the termination criterion is satisfied. Typically, the algorithm terminates when a predefined number of generations (N_{gen}) have been created. (Davis 1991)

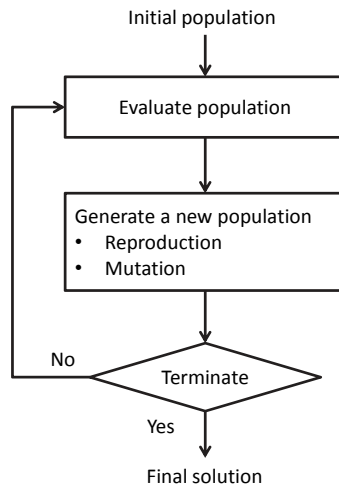


Fig. 15. A flowchart of a typical genetic algorithm.

Chromosome encoding and the objective function

Chromosomes form the population which develops throughout the genetic search. Each chromosome is a possible solution to the optimization problem. The solution is coded to the chromosomes. Binary, integer or real-valued coding can be used. Fig. 16 shows examples of different encodings. The early applications of genetic algorithms used binary coding. Binary coding, however, has some drawbacks when used for real-valued optimization problems. First, the accuracy of the solution depends on the number of bits used for each parameter. High accuracy needs more bits, leading to longer chromosomes, and may deteriorate the efficiency of reproduction. Another problem may be that a one-bit change may lead to great changes in the converted value. The conversion also needs some computations and thus the algorithm may become computationally very expensive when high accuracy is needed. (Michalewicz 1996) Consequently, it has been stated that, in general, real-valued genetic algorithms suit real engineering problems better (Chang 2007). Integer coded algorithms are quite seldom used. However, some applications can be found in the literature where they are used (Alexandridis *et al.* 2005, Gardner *et al.* 2005).

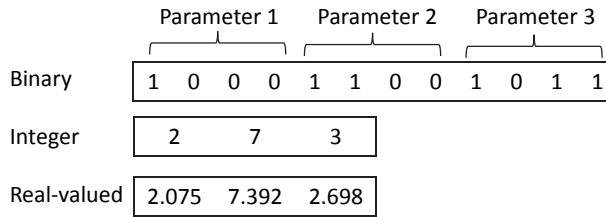


Fig. 16. Chromosomes with different encodings. Each chromosome includes three parameters to be optimized. In binary coding, four bits are used for each parameter.

In variable selection tasks, it is typical to use binary-coded genetic algorithms (Broadhurst *et al.* 1997, Barros & Rutledge 1998). In these applications, each variable in the variable set is represented by one bit. Thus, the length of each chromosome equals the number of candidate variables. If the bit corresponding to a certain variable is 0, the variable is not selected, while 1 indicates that the variable is selected. Integer coded algorithms have also been used for variable selection (Gardner *et al.* 2005). The chromosome then defines the indices of the selected variables. The drawback of this approach is that the number of selected variables must be predefined.

Many applications of genetic algorithms utilizing real-valued coding can be found in the literature. A typical application is, for example, a parameter identification problem where suitable parameter values are searched for a model. Results from such studies are reported, for example, in (Khalik *et al.* 2007), (Wang *et al.* 2008), (Sorsa *et al.* 2008b), (Ohenoja & Leiviskä 2010) and (Sorsa *et al.* 2010b).

The objective function forms the link between the genetic algorithm and the optimization problem. It is used to evaluate the fitness of the chromosomes to the optimization problem. A typical objective function is, for example, some criterion based on the squared error. The objective functions used in variable selection problems are further discussed in Section 3.6.

Selection

The generation of the new population begins with selection where chromosomes are selected for possible crossover. Selection is carried out so that better chromosomes have a higher probability to be selected. Thus, the properties of the

better chromosomes are more likely to be inherited by the chromosomes in the new population. (Davis 1991)

There are two common methods for selection: roulette wheel and tournament selection. In roulette wheel selection, the probability that a chromosome is selected is directly proportional to its fitness. It is still possible that the worst chromosome is selected, even though it is very unlikely. In roulette wheel selection, the cumulative sum of the chromosome fitness values is calculated and a uniform random number between 0 and the sum of the fitness values is taken. (Davis 1991) Table 1 shows an example with 10 chromosomes and their fitness values together with the cumulative sum. Fig. 17 shows the cumulative sum curve together with the random number. The random number is used in selection as shown in the figure. It is clear from Fig. 17 that the chromosomes with higher fitness values have a higher probability of being selected.

In tournament selection, a certain number of candidate chromosomes are selected randomly. From these chromosomes, the one that has the highest fitness value is selected. The tournament selection allows the weaker chromosomes to be selected only if they are in a tournament with even weaker chromosomes. The worst chromosome, however, can never be selected. (Fogel 2000)

The method that is used for selection has an influence on the efficiency of the algorithm. If mainly better chromosomes are selected, the algorithm converges fast. The convergence rate may be too fast, which increases the risk of premature convergence to a local optimum. When it is also allowed to select weaker chromosomes, the convergence rate is slower but the search space is better covered. (Michalewicz 1996)

Table 1. An example of population fitness values and the cumulative sum.

Chromosome	1	2	3	4	5	6	7	8	9	10
Fitness value	6.6	0.4	8.5	9.3	6.8	7.6	7.4	3.9	6.6	1.7
Cumulative sum	6.6	6.9	15.4	24.7	31.5	39.1	46.5	50.5	57.0	58.7

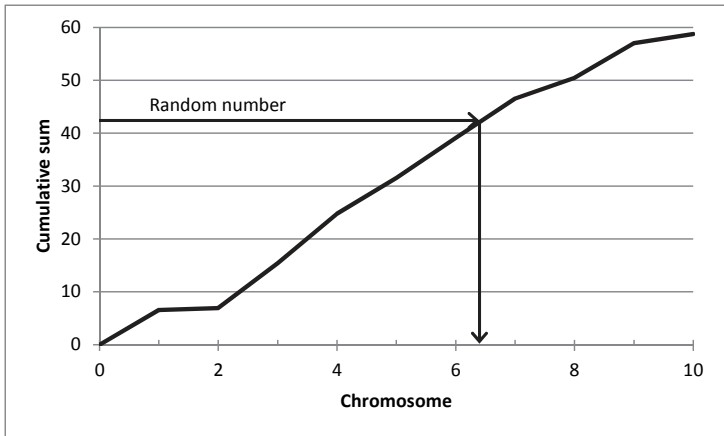


Fig. 17. Roulette wheel selection. In this case, chromosome 6 is selected.

Crossover

In crossover, the selected parent chromosomes are fused to produce offspring. Typically, two parents produce two offspring but this may change depending on the crossover method used. For example, heuristic crossover produces only one offspring from two parents. The main idea in crossover is that moderate chromosomes may include good values for some parameters. Thus, the combination of these good values leads to better chromosomes and thus the population evolves towards the optimum. Crossover rate is usually regulated with a predefined crossover probability. Thus, not all parents are necessarily crossed but are moved directly to the new population. (Goldberg 1989)

Many crossover methods exist. Some are applicable to binary-coded genetic algorithms while others only suit real-valued algorithms. With binary-coded algorithms, it is typical to use the n -point crossover where n points are selected randomly and the segments of parents determined by the random points are switched (Herrera *et al.* 1998). Fig. 18 illustrates crossover operations with one- and two-point crossover. Another method used with binary-coded algorithms is uniform crossover where each bit of the offspring is taken randomly from the parents (Davis 1991). Fig. 18 also shows uniform crossover.

With real-value coded genetic algorithms, a wide selection of crossover operators is available. These operators are, for example, arithmetic, linear and

heuristic crossover. These methods are not described here in more detail. Descriptions of them together with other methods can be found, for example, in (Deep & Thakur 2007) and (Sorsa *et al.* 2010b).

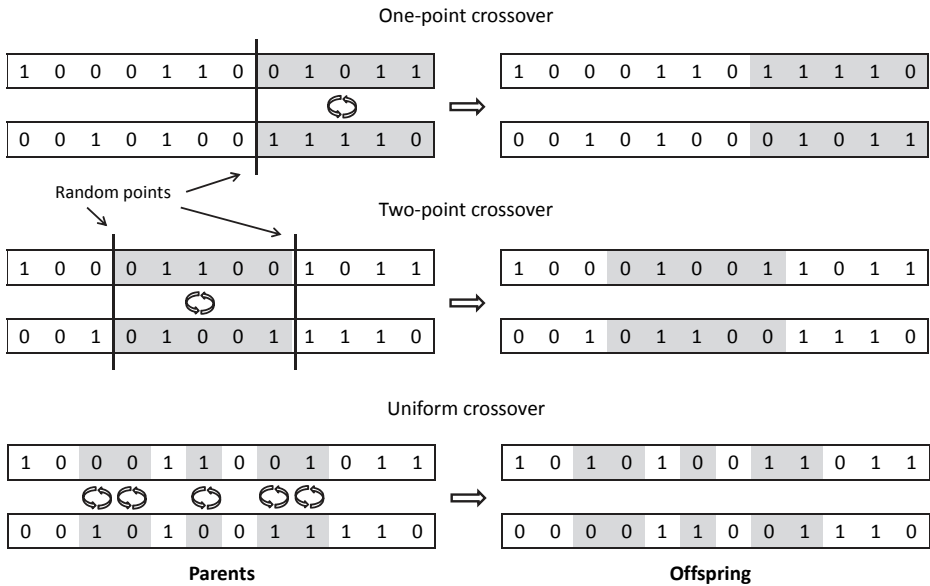


Fig. 18. Crossover operations used with binary-coded genetic algorithms.

Mutation and elitism

Mutation creates random changes in the population. It simply changes a random part of a chromosome. With binary-coded algorithms, all the digits are browsed and some of them are randomly flipped from 0 to 1 or from 1 to 0. The mutation rate is usually regulated with a mutation probability. Mutation is essential for the algorithm because it allows the search to escape local minima and also creates new genetic material for the population. (Davis 1991)

With real-value coded genetic algorithms, more options are available. Uniform mutation simply replaces the mutated parts of the chromosome with feasible new values, while boundary mutation replaces them with the lower or upper bounds of the feasible area. In non-uniform mutation, the feasible area changes with the number of generations created. During the early stages, non-

uniform mutation searches the space uniformly, but during the later stages it tends to search the space locally. (Michalewicz 1996)

It is possible that some good solutions disappear from the population through genetic operations. To avoid this, elitism can be applied. In elitism, a predefined number of the fittest chromosomes is moved directly to the new population. This may increase the dominance of the fittest chromosomes and thus decrease the diversity of the population. On the other hand, the convergence rate is increased. (Davis 1991)

Performance of the genetic algorithm

The performance of the genetic algorithm and thus the outcome of the optimization are closely dependent on the values of the tuning parameters (Katare *et al.* 2004, Eiben *et al.* 2007). The tuning parameters are crossover and mutation probabilities, rate of elitism, population size and the number of generations as mentioned above. The appropriate values of the parameters are highly problem-specific and often defined manually (Katare *et al.* 2004, Eiben *et al.* 2007). Finding the values may be a difficult task because of the complex interactions between the parameters (Deb & Agrawal 1999, Eiben *et al.* 2007). Furthermore, the optimization problem itself affects the interactions, making the tuning even more complicated (Deb & Agrawal 1999).

Crossover is the genetic operator responsible for convergence of the population. Generally, it can be said that the crossover probability should be high enough (> 0.6) for the search to be efficient (Eiben *et al.* 2007). Mutation is sometimes considered to be a secondary genetic operator (Grefenstette 1986). However, it is stated in (Nunez-Letamendia 2007) that mutation is the only way to produce genetic information not present in the initial population and thus is the mechanism that prevents the search from getting trapped in a local optimum. This is especially true with binary-coded algorithms where crossover only produces new combinations of the selected parents. Generally, it holds that higher mutation probabilities decrease the risk of premature convergence and guarantees that the population is diverse enough (Boyabatli & Sabuncuoglu 2004). On the other hand, if the mutation probability is too high, the risk of destroying already found good solutions is high (Deb & Agrawal 1999). In such cases, the performance of the algorithm approaches the performance of the random search (Grefenstette 1986).

The population size is also a very significant parameter in terms of the outcome of the genetic search (Grefenstette 1986). If the population size is too

small, the risk of finding only a local optimum increases (Nunez-Letamendia 2007). Also, the efficiency of crossover and thus the convergence rate may be inappropriately low with too small populations (Grefenstette 1986). Furthermore, too large populations may lead to low convergence rates because more objective function evaluations are needed per generation (Grefenstette 1986). This also increases the computational burden of the algorithm. Consequently, it is of major importance to use a population of appropriate size (Deb & Agrawal 1999).

The number of generations also affects the computational load of the search (Boyabatli & Sabuncuoglu 2004). If too few generations are generated it may be possible that chromosomes do not adapt to their environment (Deb & Agrawal 1999). On the other hand, too many generations increase the computational load drastically.

As mentioned above, the tuning of the parameters is typically carried out manually one parameter at a time. This may be time-consuming and does not necessarily lead to optimal parameter combinations. Another approach controls the parameters during optimization (Eiben *et al.* 2007). However, in this paper parameter control is not used and thus these methods are not discussed in more detail.

Genetic algorithms in variable selection problems

Many successful applications of genetic algorithms in variable selection can be found in the literature. Suitable feature subsets have been selected for MLR (Jouan-Rimbaud *et al.* 1996, Broadhurst *et al.* 1997), PCR (Barros & Rutledge 1998, Depczynski *et al.* 2000) and PLSR (Kimura *et al.* 1998, Leardi & González 1998, McLeod *et al.* 2009) models. Applications also exist where the input variables for ANN models are selected with genetic algorithms. Llobet *et al.* (2004) use genetic algorithms to find the optimal subset for a fuzzy artmap neural model suitable for classification problems. Gardner *et al.* (2005) select variables for a PNN model while (Alexandridis *et al.* 2005) and (Wang & Xiang 2007) and use RBFNNs. BPNNs are studied in (Hemmateenejad 2005).

Even though good results have been reported with genetic algorithms, some drawbacks have also been noticed. One is that genetic algorithms are computationally very expensive (Alexandridis *et al.* 2005). Another drawback is related to the huge search space that is typically associated with variable selection tasks. As mentioned earlier, the initial population has an effect on the outcome of the search (Rahnamayan *et al.* 2007) and thus it is assumed that only a small

fraction of the search space is covered with a single run of the algorithm (Leardi & González 1998). Some studies have also found that insignificant variables are selected among the significant ones when using genetic algorithms (Jouan-Rimbaud *et al.* 1996, Leardi & González 1998, Baumann 2003). Thus it has been suggested that genetic algorithms should not be used for very large-scale variable selection problems (Leardi & González 1998, Gauchi & Chagnon 2001).

To avoid the instability caused by the variations in initial populations and the selection of insignificant variables, cascaded selection strategies have been proposed. It has been reported that filter methods can be used for eliminating noisy and collinear variables before a genetic search. For example, a variance-based elimination is used in (Llobet *et al.* 2004, Wasim & Brereton 2004, Chu *et al.* 2007) to remove noisy variables. Collinear variables can be removed, for example, based on correlation (Llobet *et al.* 2004). The filter methods however, need a predefined threshold value and are use-dependent (Llobet *et al.* 2004, Chu *et al.* 2007). It is also possible to use genetic search as the pre-selection procedure and continue with some other method. For example, the results of the repeated runs of a genetic algorithm may be refined to a ranking according to which variables are added to the model (Leardi & González 1998, Dieterle *et al.* 2003). The results of a genetic search have also been used as a starting point for searches carried out with forward selection (McLeod *et al.* 2009), backward elimination (Jouan-Rimbaud *et al.* 1996) or simulated annealing (Alexandridis *et al.* 2005).

3.6 Objective functions

The objective function is critical in obtaining results with the desired properties. It may contain terms counting for predictive ability and complexity of the model. Basically, the objective function is a compromise between model accuracy and parsimony (Li *et al.* 2002, Guyon & Elisseeff 2003). If all the data is used for model training, the model is easily overfitted. An overfitted model performs very accurately (even perfectly) for the training data set but fails with other data sets (Hawkins 2004). Thus, it is desirable to split the data set into training and validation sets. Two different types of validation sets can be used: internal and external. An internal validation data set is separated from the training data set and is used when determining the optimal variable subset. An external validation data set is an independent data set that is used to guarantee the validity of the model for future use. (Harrell 2001) Fig. 19 illustrates the data splits.

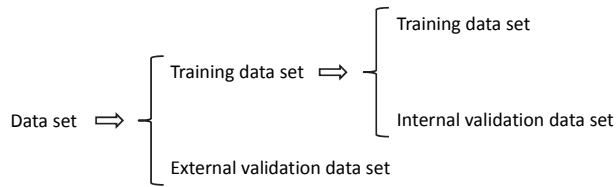


Fig. 19. Different data sets separated from the whole data set. Training and internal validation data sets are used for variable selection and model identification while the external validation data set is used for testing.

3.6.1 Resampling methods

When plenty of data is available, a static data split can be made. With small data sets, however, a static split may lead to an unacceptable loss of data for model training. In such cases, resampling methods can be used where all the data is used for training and validation. Typical resampling methods are cross-validation and bootstrapping. In bootstrapping, samples of size N (the number of data points) are drawn from the whole data set with replacement, meaning that the very same data point can be drawn multiple times. Then the model is fitted to the drawn sample and its performance is evaluated. An average of these evaluations is the estimate of the model performance. This procedure is referred to as simple bootstrap. An advanced version of bootstrap also estimates the optimism in model performance and subtracts it from the performance estimate. (Efron & Tibshirani 1993, Harrell 2001) Bootstrap methods are not that often used in variable selection problems even though some applications can be found (Amato & Vinzi 2003, Chu *et al.* 2004).

There are basically three resampling cross-validation methods: leave-one-out (LOO), leave-multiple-out (LMO) and k-fold. In each of these methods, some of the data is used for training the model and the remaining part is used for validation. In the LOO procedure, one data point is left out for validation while all the other data is used in training. This is repeated so that all the data points are left out for validation and the estimate of the model performance can be made. In k-fold cross-validation, k data sets are randomly drawn without replacement. One of these sets is left out for validation while the others are used for training. The LMO procedure uses data so that multiple data points are left out for validation at a time. Also in this case, the samples used for training and validation are drawn without replacement. The k-fold and LMO procedures use random data splits and

thus it is advantageous to make multiple splits to estimate model performance. This makes both of these methods computationally more expensive. These algorithms also possess some tuneable parameters, which makes their use more difficult. (Baumann 2003)

Sometimes the data set is so small that no data at all can be separated for external validation. In such cases, double cross-validation (also called model cross-validation) can be used (Wessels *et al.* 2005, Anderssen *et al.* 2006, Foca *et al.* 2009). In double cross-validation, external validation is also carried out with resampling methods, leading to a great increase in the computational burden.

3.6.2 Criteria used

The definition of an appropriate objective function may be difficult. The objective function may include terms accounting for predictive power and model complexity, as mentioned earlier. These terms are to be used so that the result of the variable selection is an accurate model with as few variables as possible (Guyon & Elisseeff 2003, Hawkins 2004).

The simplest approach to reaching these goals is to add a penalty constant that penalizes the number of variables in the model (Kohavi & John 1997, Barros & Rutledge 1998, Dieterle *et al.* 2003). Barros & Rutledge (1998) multiplied the sum of the squared error of predictions (SSEP) by the number of features. They also used a term (Durbin-Watson criterion), which penalizes the solution if the modelling residuals are not random. A very small penalty constant is used in (Kohavi & John 1998) to break ties in favour of a smaller subset of variables. A tuneable penalty constant is proposed in (Zhang 1992) where the number of variables is simply multiplied by an appropriate constant. Some criteria readily hold the terms accounting for model parsimony. Such criteria are, for example, the Bayesian information criterion (BIC), Akaike information criterion (AIC) and Mallows's C_p criterion (Li *et al.* 2002, Wisnowski *et al.* 2003).

More parsimonious models are also obtained if only a near optimal solution is searched (Wisnowski *et al.* 2003). This has been studied, for example, by Li *et al.* (2002) who compared Wold's R criterion to its adjusted versions and also to AIC. Wold's R criterion is obtained as a ratio of the sum of the squared prediction error (SSEP) values of two models. Furthermore, the SSEP values are obtained through some cross-validation procedure. The ratio is then compared to unity to determine which model is better. In the adjusted version of Wold's criterion, the ratio is compared to thresholds smaller than unity to favour more parsimonious models.

Li *et al.* (2002) found that when model parsimony is the major concern, the adjusted Wold's R criterion is to be used and when model accuracy is the main goal, they proposed the use of Wold's original criterion or the AIC.

Another approach for finding more parsimonious models is to restrict model complexity in the selection procedure. This can be implemented by simply defining the maximum number of variables that can be added to the model (Swierenga *et al.* 1998, Baumann 2003). A similar limitation is also used in (Gardner *et al.* 2005), who used an integer-coded genetic algorithm for selecting variables. In their approach, chromosomes include information about the indices of the variables included in the model. Thus the chromosome length must be defined, which also defines the number of variables used in the model.

3.7 Functions used in the studies

Some functions used in the studies by the author are presented in this section. The functions are used in the feature generation step described in Section 5.2. The first set of functions is fitted into the BN profiles presented in Fig. 4. These functions are normal distribution, triangular function, trapezoidal function and the two-sided composite of two normal distributions. They are given respectively by

$$f(H) = \exp\left(-\frac{(H - H_1)^2}{2s^2}\right), \quad (16)$$

$$f(H) = \begin{cases} 0, & H < H_1 \\ k_1H + a_1, & H_1 \leq H < H_2 \\ k_2H + a_2, & H_2 \leq H \leq H_3 \\ 0, & H > H_3 \end{cases}, \quad (17)$$

$$f(H) = \begin{cases} 0, & H < H_1 \\ k_1H + a_1, & H_1 \leq H < H_2 \\ 0, & H_2 \leq H < H_3 \\ k_2H + a_2, & H_3 \leq H \leq H_4 \\ 0, & H > H_4 \end{cases} \quad \text{and} \quad (18)$$

$$f(H) = \begin{cases} \exp\left(-\frac{(H-H_1)^2}{2s_1^2}\right), & H < H_1 \\ 1, & H_1 \leq H < H_2 \\ \exp\left(-\frac{(H-H_2)^2}{2s_2^2}\right), & H \geq H_2 \end{cases} . \quad (19)$$

Above, H is the applied magnetic field, H_i refers to a certain strength of the external magnetic field, s denotes the spread of the normal distribution and k is the slope of the triangular or trapezoidal function.

Another set of functions used is certain factors. They are applied directly to the BN signal. These factors are crest, clearance, impulse and shape. They are given by (Lei *et al.* 2008, Yadav & Wadhvani 2011)

$$\text{Crest factor} = \frac{\text{peak}}{\text{RMS}}, \quad (20)$$

$$\text{Clearance factor} = \frac{\text{peak}}{\frac{1}{N} \left(\sum_{i=1}^N \sqrt{|x_i|} \right)^2}, \quad (21)$$

$$\text{Impulse factor} = \frac{\text{peak}}{\frac{1}{N} \sum_{i=1}^N |x_i|} \text{ and} \quad (22)$$

$$\text{Shape factor} = \frac{\text{RMS}}{\frac{1}{N} \sum_{i=1}^N |x_i|}, \quad (23)$$

where N is the number of data points, RMS is as given in (1) and

$$\text{peak} = 0.5(\max(x_i) - \min(x_i)). \quad (24)$$

4 Experimental set up

This section provides information about the basic set up of the research. The studied materials and the equipment used are described. In addition, a description is given of the data sets used.

4.1 Studied materials

The studied material was carburizing case-hardened steel 18CrNiMo7-6 (EN 10084). Its typical chemical composition is given in Table 2. This material is typically used in wind turbine gears and other industrial transmission applications. The studied samples underwent different procedures so that the final residual stress and hardness varied. For example, different tempering temperatures and times or different laser processing parameters were used. The procedures are described in more detail in Section 4.3.

Table 2. The nominal chemical composition of case-hardened steel 18CrNiMo7-6. (Sorsa *et al.* 2008a).

Element	C	Cr	Si	Mo	Mn	Ni	P	V	S	Cu	Al	Sn	N
%	0.17	1.65	0.19	0.32	0.56	1.52	<0.01	<0.01	<0.004	0.02	0.028	0.002	0.012

4.2 Measurements

The Barkhausen noise measurements were carried out with a Rollscan 300 instrument. The magnetising frequency used was 45 or 125 Hz while the magnetizing voltage was 4.3 Vpp (voltage peak to peak). The measured Barkhausen noise signals were stored with a Microscan software program. The residual stress measurements were carried out using the X-ray diffraction method. The measurement device used was an XStress 300 with CrK α radiation and the chi method. The voltage used was 30 kV, the current 6.7 mA and the collimator diameter 3 mm. Both of the above-mentioned measurement devices were manufactured by Stresstech Oy (Finland). The hardness measurements were performed with a Matsuzawa NMT-X7 (Matsuzawa Co. LTD, Japan) hardness tester. The applied load for these Vickers hardness measurements was 1 kg.

4.3 Data sets

Three data sets were used in the studies. The first and the third data sets were obtained from the samples with varying tempering temperature and times. The temperatures varied between 180 and 350 °C and the times between 90 and 240 min. The samples in the second data set were processed by laser to vary the surface hardness and residual stress state. Different laser processing parameters were used.

Regardless of the data set, the measurement signals underwent certain pre-processing steps. First, the signals may have included incomplete magnetizing loops which were removed. The second step in pre-processing was Boxcar averaging where a certain number of data points were replaced by their average. In the data sets used, the window size of 10 was used. This step was observed not to impair the quality of the data (in fact quite the opposite). It efficiently reduces the size of the data which is desired for future processing.

The further processing of the signals included moving window-based filtering. RMS- and entropy-based filters were applied. The Barkhausen noise profiles were obtained by means of this filtering step. Table 3 shows the dimensions of the data sets.

Table 3. The dimensions of the data sets used.

	Data set 1	Data set 2	Data set 3
Number of features	72	72	55
Number of data points	60	24	115

5 Results

This section presents the research contributions of the author. The aim of the whole study was to obtain a suitable methodology for building prediction models of material properties based on Barkhausen noise (BN) measurements. The overall modelling approach is presented in Section 5.1. Then the steps of model development are presented in Sections 5.2–5.4. The results are summarised in Section 5.6.

5.1 Modelling approach

The development of a prediction model includes four steps: feature generation, feature selection, model identification and model validation. Fig. 20 illustrates the overall prediction scheme. In feature generation, the measured BN signal is analysed and different computational techniques are used to extract information from the signal. The feature generation step produces a large set of features which is then subjected to the feature selection procedure where a suitable feature subset is selected. Different selection procedures are studied and their applicability and efficiency is analysed. The amount of data was quite limited and thus resampling cross-validation methods were used in this step, producing information about model validity.

Feature selection is closely related to model identification as shown in Fig. 20. Model identification generally includes the model structure and parameter identification steps. In these studies, however, the model structure was predefined and thus only the model identification step was carried out. After the model identification step, independent test data was used to validate the model. Also, the performance criteria from the feature selection step were used in validation.

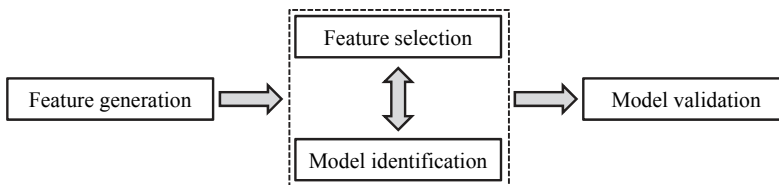


Fig. 20. The applied prediction scheme.

5.2 Feature generation

The aim of feature generation is to extract information from the measured BN signal. Features can be extracted basically from the time or frequency domain signals. In the studies included in this thesis, the time domain features are used mainly. Features are obtained from the measuring device, directly from the measured signal or from the BN profile. The features obtained directly from the measuring device are shown in Table 4.

Table 4. Features from the measuring device.

Feature	Explanation
RMS	Associated with the overall Barkhausen activity; calculated from (1)
Peak height	The height of the BN profile
Peak position	The position of the BN profile
FWHM	The full width at half maximum; Can be associated with peak width.
Coercivity *	Magnetic field strength needed to obtain zero magnetization of the material. See Section 2.1.
Remanence *	The remaining magnetization of the material after the external field is removed. See Section 2.1.
Permeability *	The degree of magnetization of the material when placed in an external magnetic field.
Loop area *	Area of the hysteresis loop. See Section 2.1.
Spectrum area	Amplitude spectrum area

* These are not the standard quantities but only computational features from the BN signal. However, they can be associated with the standard quantities (Sorsa *et al.* 2012b).

5.2.1 Features calculated directly from the BN signal

The features calculated directly from the BN signal include (Sorsa & Leiviskä 2008, Sorsa & Leiviskä 2009a, Sorsa & Leiviskä 2012a, Sorsa & Leiviskä 2012b)

- statistical features: average, standard deviation, skewness, kurtosis, quantiles and inter quartile range (IQR)
- BN energy and BN entropy
- crest, clearance, impulse and shape factors.

Entropy is a measure of the uncertainty of a random variable. Entropy is calculated by utilising the probability mass function (pmf), $p(x)$, of the variable. The sum of the elements in the pmf equals 1. The pmf is obtained from the histogram of the signal, which is affected by the bin size used. Fig. 21 shows

typical histograms of the BN signal and illustrates the effect of the bin size. After the pmf is obtained, entropy is given by (Cover & Thomas 2005)

$$H(X) = -\sum p(x) \log p(x). \quad (25)$$

In (25), the base of the logarithm is typically 2 but other bases can also be used. In the studies presented here, a 10-based logarithm is used. It should be noticed that a problem exists when some component of $p(x)$ equals zero. Thus, it is defined that $0 \log 0 = 0$ (Cover & Thomas 2005). Consequently, a zero component has no effect on the entropy. The use of BN entropy as a feature was first suggested in (Sorsa & Leiviskä 2009a).

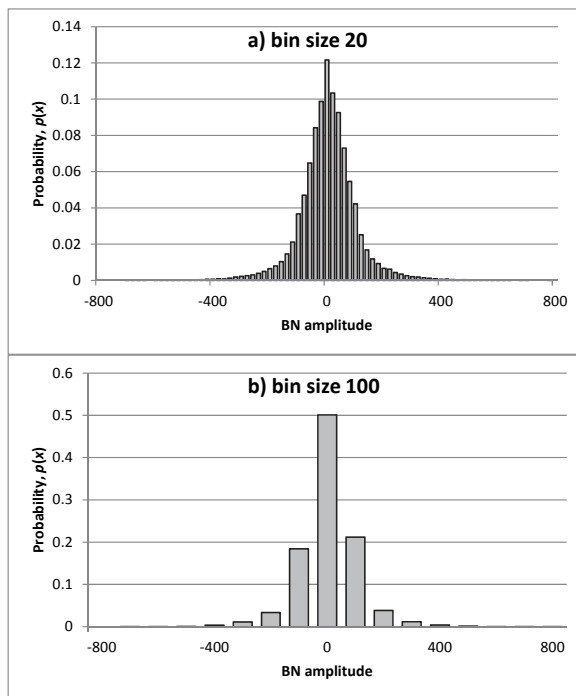


Fig. 21. Typical histograms of the BN signal with bin size a) 20 and b) 100.

5.2.2 Features of the BN profile

The BN profile is obtained by applying some moving window-based filtering to the BN signal and then considering it as a function of the applied magnetic field. It is typical to use moving RMS filtering but entropy (Sorsa & Leiviskä 2010, Sorsa & Leiviskä 2011c), power spectral density (Piotrowski *et al.* 2010) and pulse count (Augustyniak *et al.* 2010) have also been used. In the studies presented in this thesis, moving RMS and moving entropy with two bin sizes are used. Typical features obtained from the BN profile are peak height, position and width as mentioned in Section 2.3. Even though these features are obtained readily from the measuring device (Table 4), they are calculated for all the different BN profiles because different computational procedures may lead to different information content.

Peak height (h) and position (p) are obtained by fitting a second order polynomial to the profile top. Usually 25% of the highest values are used as the profile top. The polynomial then defines the height and position. These are used when calculating the peak width (w) as shown in Fig. 22. Widths can also be calculated corresponding to heights other than $0.5h$ such as $0.25h$ and $0.75h$. (Sorsa & Leiviskä 2009a, Sorsa & Leiviskä 2010).

It has been stated in the literature that more robust features are obtained if some function is fitted to the BN profile and then the function parameters are used as the features (Stewart *et al.* 2004). Thus, in the studies presented in this thesis, many different functions are used. The fitted functions are normal distribution, triangular function, trapezoidal function and the two-sided composite of two normal distributions (Sorsa *et al.* 2008a, Sorsa *et al.* 2012a). The functions are given in Section 3.7 and shown in Fig. 23. The figure also shows the features obtained from the functions. The functions also need two parameters accounting for scaling issues. Thus all the fitted functions are of the form (Sorsa *et al.* 2008a, Sorsa *et al.* 2012a)

$$\hat{M} = af(H) + c, \quad (26)$$

where M is the filtered BN, H is the applied magnetic field and a and c are the scaling parameters. The functions are fitted using the nonlinear least squares method. The features obtained from the fitted functions are given in Table 5.

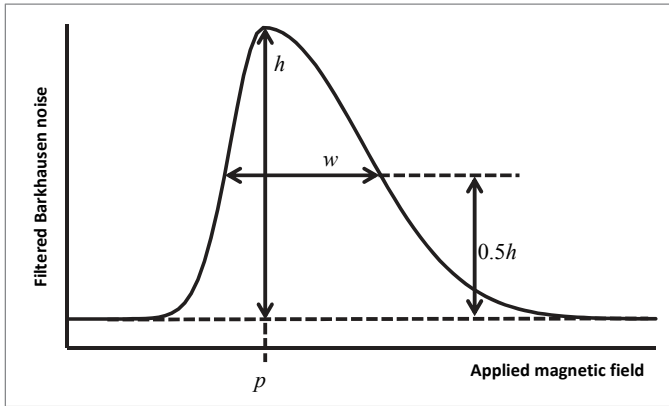


Fig. 22. Calculation of peak width (h is peak height, p is peak position and w is peak width).

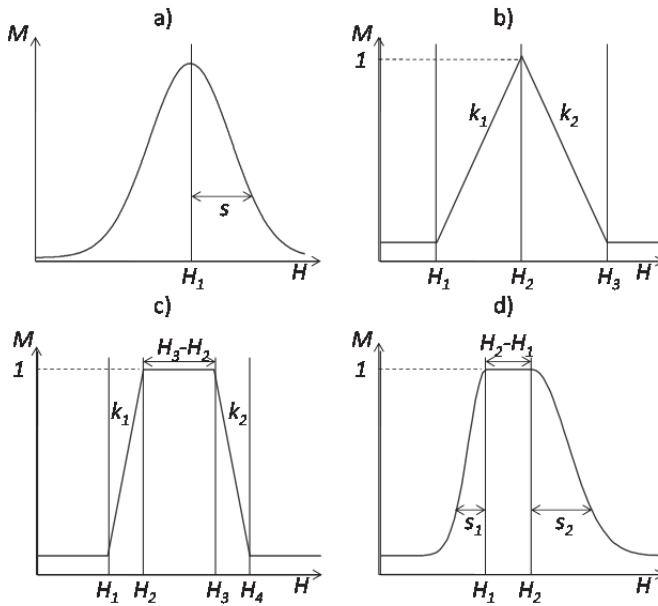


Fig. 23. The fitted functions: a) normal distribution, b) triangular function, c) trapezoidal function and d) two-sided composite function.

Table 5. The features obtained from the fitted functions. (Sorsa *et al.* 2008a, Sorsa *et al.* 2012a)

Normal distribution	Triangular function	Trapezoidal function	Two-sided composite function
H_1, s	$H_1, H_2, H_3,$ $H_3 - H_2$ k_1, k_2	$H_1, H_2, H_3, H_4,$ $H_3 - H_2$ k_1, k_2	H_1, H_2 s_1, s_2 $H_2 - H_1$

5.3 Feature selection

The feature generation procedure described in the previous section produces a large set of features. It is typical of Barkhausen noise studies that the number of measurements is low and thus selection must be carried out carefully. The complete feature set includes 241 features but is never used as such in these studies. Instead, reduced but still relatively large sets have been used. The number of features varies depending on the study but the selected features are indicated with the indices of the complete set to compare the results better. In other words, the feature set among which the features are selected may contain 72 features but feature x_{179} may still have been selected. In this section, the results from the studied selection schemes are presented. The selection procedures were implemented as described in Section 3. The problem-specific modifications (cross-validation, objective function, tuning parameters) are given in this section together with the results.

5.3.1 Manual selection

Manual selection obviously does not use any automatic selection procedures but is merely based on the literature and the experience of the modeller. Also, some indices can be used to guide the selection. When such an index is used the selection basically becomes a filter (see Section 3.1).

Results from a study where the features were manually selected are presented in (Sorsa *et al.* 2008a) and (Sorsa *et al.* 2012a). The results in (Sorsa *et al.* 2012a) are presented here in more detail. In the study, residual stress is predicted for data sets 1 and 2. Data set 1 is used in feature selection and both of the data sets in model identification and validation. A simple MLR model given in (2) is identified. The selection is carried out in three steps. The first step groups the features based on their cross-correlations. In the second step, one feature from each group is selected and the irrelevant features are removed in the final step.

The first step where the features are grouped produces eight groups and 14 individual features that hold no significant correlation to other features. One feature from each group is then selected, based on the correlation between the features and residual stress. If no feature in a group has significant correlation, the entire group is eliminated. The same procedure is also carried out for the individual features. The selection procedure produces a subset of ten features. These features undergo an elimination procedure where the influence of each feature on prediction accuracy is evaluated. If no significant deterioration in model behaviour is observed, the feature is eliminated. (Sorsa *et al.* 2012a)

The features selected in the prediction model are the FWHM value (x_4), coercivity (x_5), Pearson mode skewness (x_{14}), crest factor (x_{28}), the fitting parameter H_3 of a trapezoidal function (x_{62}) and the fitting parameters s_2 (x_{68}) and H_2 (x_{69}) of a two-sided composite function. Table 6 shows the regression coefficients of these features and Fig. 24 the model accuracy. The figure shows that the model is good for both data sets. The average error (RMSEP) for data set 1 is 57.68 MPa and 139.37 MPa for data set 2. The correlation coefficients between the measured and predicted stresses are 0.85 and 0.91 for data sets 1 and 2, respectively. (Sorsa *et al.* 2012a)

Table 6. The regression coefficients of the model obtained through the manual feature selection procedure (Sorsa *et al.* 2012a).

Parameter	b_4	b_5	b_{14}	b_{28}	b_{62}	b_{68}	b_{69}
Data set 1	-72.5	36.1	-10.8	50.9	19.1	-15.2	32.0
Data set 2	99.7	55.0	-9.6	392.8	-17.4	2.3	-52.0

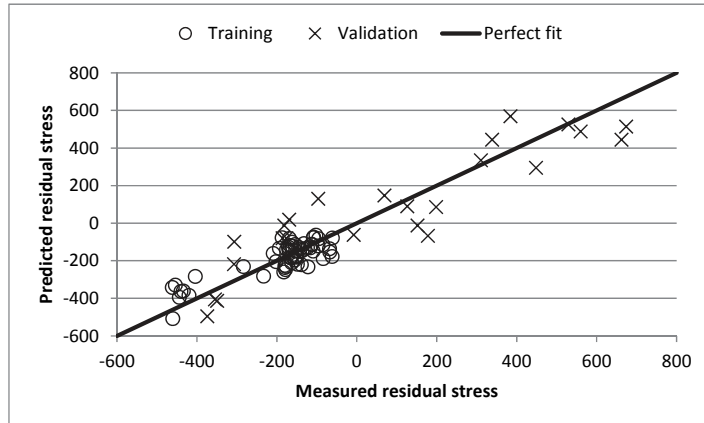


Fig. 24. The prediction accuracy of the models given in Table 6. Drawn based on the data in (Sorsa *et al.* 2012a)

5.3.2 Forward selection and backward elimination

Based on the literature, backward elimination is not expected to perform well when applied alone (Kohavi & John 1997). Therefore it was not used in these studies except in (Sorsa & Leiviskä 2011a) where its performance proved to be insufficient. The aim of this section is to present the results from the studies where forward selection was used. However, it should be noticed that a procedure similar to backward elimination is used in many studies as the final elimination step of the features. Forward selection has been used by the present author in many studies (Sorsa & Leiviskä 2010, Sorsa & Leiviskä 2011a, Sorsa & Leiviskä 2011c, Sorsa & Leiviskä 2012b). The results from (Sorsa *et al.* 2012b) are given below.

In the study, an MLR model was used for predicting residual stress and hardness. The data usage is the same as in the study presented in the previous section where manual selection was used (Sorsa *et al.* 2012a). Thus, data set 1 is used for feature selection while both of the data sets are used for model identification and validation. Altogether 72 features are included in the data set. The objective function used is

$$MSEP_{CV} = \frac{1}{N_{CV}} \sum_{i=1}^{N_{CV}} (\sigma_i - \hat{\sigma}_i)^2, \quad (27)$$

where N_{CV} refers to the number of data points used for internal validation and σ is the residual stress. The LMO cross-validation procedure is used in internal validation with half of the data points left out for validation and $2N$ (N is the number of data points) different data splits. (Sorsa *et al.* 2012b)

Fig. 25a and Fig 25b show the objective function value as a function of the number of features added when a feature subset is searched for predicting residual stress and hardness, respectively. As seen from the figures, the model improves each time a new feature is added but the internal validation procedure reveals the suitable number of features. The minimum is reached with six and four features for residual stress and hardness predictions, respectively. However, an almost equal but more parsimonious model should be preferred (Wisnovski *et al.* 2003, Alexandridis *et al.* 2005). Thus five features are selected for the residual stress predictions, i.e. the FWHM value (x_4), coercivity (x_5), crest factor (x_{28}), peak position obtained from the moving entropy profile, (x_{45}) and the fitting parameter H_4 of the trapezoidal function when fitted to the moving entropy profile (x_{191}). For hardness predictions, the model includes three features, i.e. peak position from the measuring device (x_4), remanence (x_6) and the fitting parameter H_1 of the trapezoidal function when fitted to the moving entropy profile (x_{188}). The values of the regression coefficients for the models are given in Table 7 and Table 8.

The accuracy of the prediction models is presented in Fig. 26 and the performance indices in Table 9. The accuracy is good for residual stress predictions and also for data set 1 of the hardness predictions, although a decrease in model performance is observed with data set 2. This indicates that the hardness model may contain some excess features. (Sorsa *et al.* 2012b)

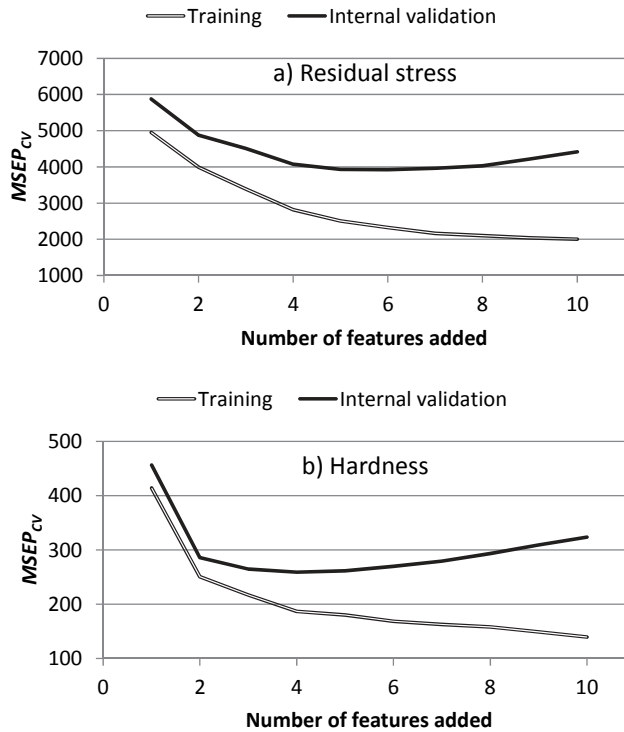


Fig. 25. The objective function value as a function of the number of features added when predicting a) residual stress and b) hardness. Drawn based on the data in (Sorsa *et al.* 2012b)

Table 7. The regression coefficients of the model for the prediction of residual stress obtained through the forward-selection procedure (Sorsa *et al.* 2012b).

Parameter	b_0	b_4	b_5	b_{28}	b_{45}	b_{191}
Data set 1	-184.7	-49.7	13.6	83.2	24.8	52.7
Data set 2	88.6	-26.0	190.6	263.2	-147.7	-81.9

Table 8. The regression coefficients of the model for the prediction of hardness obtained through the forward-selection procedure (Sorsa *et al.* 2012b).

Parameter	b_0	b_3	b_6	b_{188}
Data set 1	654.6	-13.6	-52.4	-14.0
Data set 2	563.9	-32.4	-90.6	-9.8

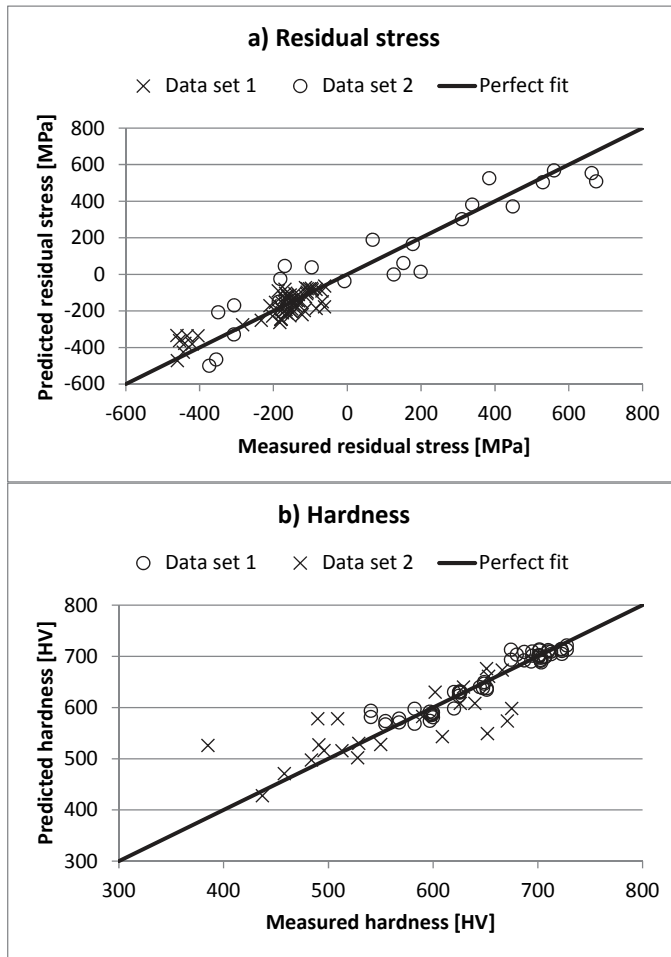


Fig. 26. The prediction accuracy of the prediction models using forward selection. Drawn based on the data in (Sorsa *et al.* 2012b).

Table 9. The performance indices of the models obtained through forward selection.

	Data set 1		Data set 2	
	RMSEP	R	RMSEP	R
Residual stress	53.11 MPa	0.87	111.82 MPa	0.94
Hardness	14.96 HV	0.96	53.63 HV	0.76

5.3.3 Simulated annealing

Only one study has been carried out using simulated annealing (Sorsa & Leiviskä 2011a). In this study, hardness is predicted with an MLR model. Data sets 1 and 2 are normalised and then fused. The objective function is as given in (27) where the LMO cross-validation procedure is used.

The simulated annealing algorithm deployed uses a geometrical cooling schedule as given in (15) with $c = 0.75$. The initial temperature is set to 10 so that almost all the moves are accepted in the early stages of the search. The minimum temperature is set to 0.001. Thus, the final temperature is $T = T_{40}$ and almost all detrimental moves are then prohibited. During each temperature step, 200 new solutions are generated before the temperature is lowered. Also, if 20 consecutive successful trials are generated, the temperature is lowered. The solution is represented by a string of M (M is the number of features) binary digits. If the digit is 1, the feature is selected, but if it is 0, the feature is not selected. The new solution is generated by changing one bit of the current solution. The algorithm stops if the minimum temperature is reached or if 1000 consecutive rejections are obtained. (Sorsa & Leiviskä 2011a)

Because the result of simulated annealing may vary, the search is carried out 100 times. On average, 8.5 features are selected while the best solution found contains eight features. The selected features are the FWHM value (x_4), signal-to-noise ratio (x_{15}), the fitting parameters H_1 (x_{60}) and H_4 (x_{63}) of the trapezoidal function fitted to moving RMS profile, the fitting parameter H_1 of the normal distribution function fitted to the moving entropy profile (x_{179}) and the fitting parameters a (x_{187}), H_1 (x_{188}) and H_2 (x_{189}) of the trapezoidal function fitted to the moving entropy profile.

The prediction accuracy of the model is reasonable. The average error (RMSEP) for the training and external validation data sets is 43.13 and 42.99 HV, respectively. The correlation coefficients are 0.77 and 0.86, respectively. However, the predictions are slightly biased. (Sorsa & Leiviskä 2011a)

5.3.4 Genetic algorithms

Many studies have been carried out where a genetic algorithm is used for feature selection. The results of these studies are reported in (Sorsa & Leiviskä 2009b), (Sorsa *et al.* 2010b) and (Sorsa & Leiviskä 2011a). Furthermore, two more studies have been carried out but the results have not yet been published. The genetic

algorithm used in all these studies utilises binary coding, where the length of the chromosome equals the number of features. If a bit corresponding to a certain feature is 1, the feature is selected while 0 means that the feature is not selected. The tuning parameters of the algorithm were typically determined through systematic testing. Either one-point or uniform crossover was used.

In the study presented in (Sorsa & Leiviskä 2009b) and (Sorsa *et al.* 2010b), data set 3 is used and residual stress is predicted. The data set contains 115 data points and 51 features. The objective function utilises a penalty constant so that more parsimonious models are favoured. The objective function is

$$J = SSEP_{CV} + \lambda m, \quad (28)$$

where λ is a tuneable penalty constant, m is the selected number of features and $SSEP_{CV}$ is given by

$$SSEP_{CV} = \sum_{i=1}^{N_{cr}} (\sigma_i - \hat{\sigma}_i)^2. \quad (29)$$

10-fold cross-validation is used in this study with 11 repetitions. The median of these repetitions is then taken as the $SSEP_{CV}$ value. The genetic search is repeated 50 times where the tuning parameters are: $N_{pop} = 200$, $N_{gen} = 20$, $p_c = 0.9$ and $p_m = 0.02$. The initial population is created so that it contains only 10% of ones and 90% of zeros. That is to favour more parsimonious solutions because the optimal solution is expected to contain quite a small number of features. (Sorsa & Leiviskä 2009b, Sorsa *et al.* 2010b)

The results obtained indicate that a suitable subset contains four features. However, good solutions are also found where more features are selected. Fig. 27 shows the occurrence of the number of features in the subsets found in the 50 repetitions of the algorithm. Even though the results strongly indicate that four features should be used, all the suitable feature subsets are further subjected to the elimination procedure described earlier (Section 5.3.1). Thus, the models are identified without one of the features at a time and unless the removal of the feature is detrimental in terms of model accuracy, it is eliminated. This procedure shows that only the subset with four features contains only significant features and thus four is indeed the most suitable number of features in this case. The selected features are the RMS value (x_1), coercivity (x_5), the loop area (x_8) and the fitting parameter H_l of the trapezoidal function fitted to the moving RMS profile (x_{60}). The performance of the model found is reasonable because the average error

(RMSEP) is 20.32 MPa and the correlation coefficient between the measured and predicted stresses is 0.86. (Sorsa & Leiviskä 2009b, Sorsa *et al.* 2010b)

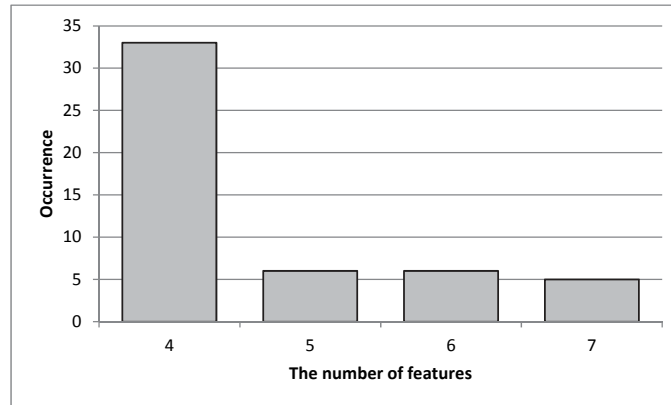


Fig. 27. The occurrence of a different number of features in 50 repetitions of the genetic algorithm.

5.3.5 Two-step selection strategies

One drawback of genetic algorithms in feature selection problems is that the subset found may contain irrelevant features as mentioned in Section 3.5.2. Thus there are some studies in the literature where cascaded selection strategies with genetic algorithms are used. In some studies, the genetic search is preceded by a pre-selection step (Llobet *et al.* 2007, Wasim & Brereton 2004, Chu *et al.* 2007) and in other studies genetic algorithms perform the pre-selection (Jouan-Rimbaud *et al.* 1996, Leardi & González 1998, Dieterle *et al.* 2003, Alexandridis *et al.* 2005, McLeod *et al.* 2009). Such algorithms were also studied by the author of this thesis and the results from these studies are reported in this section.

Successive projections algorithm in pre-selection

In the first two-step approach, the prediction model (RBFNN) for residual stress is identified so that the dimension of the original feature set is reduced with a filter method and then the actual selection is carried out with a genetic algorithm. The filter method used is the successive projections algorithm (SPA). Data sets 1 and 2 are fused in this study and thus 111 features and 84 data points are available.

The objective function is given in (27) and the LMO procedure is used in the computation of the $MSEP_{CV}$ value. An external validation data set is separated and contains 12 data points.

In this study, SPA is used to eliminate features so that the resulting subset contains as few useful features as possible, but almost no information is lost. This is achieved by selecting the first feature so that it has the highest correlation to the predicted material property (residual stress in this study) and then proceeding according to SPA. Thus starting from the first feature, the ranking is generated using SPA. According to the ranking, features are added one at a time until 99% of the variance of the original data set is explained. Fig. 28 shows the variance explained as a function of the number of added features. Based on the figure, 67 features are eliminated, leaving only 44 for the genetic search.

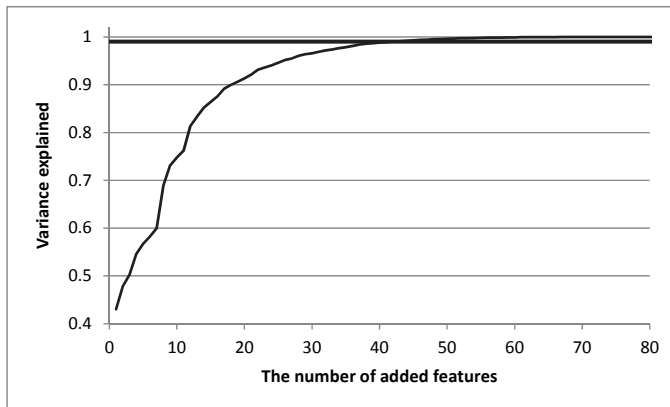


Fig. 28. The variance explained as a function of the number of features added.

To evaluate the efficiency of SPA in pre-selection, features are also selected directly from the whole data set. The genetic algorithms are implemented as described in Section 5.3.4 with the tuning parameters given in Table 10. The algorithm is repeated 100 times. The aim is to identify a nonlinear neural network model although the genetic algorithm uses a linear MLR model. A similar approach has been used earlier with good results, for example, in (Eklöv *et al.* 1999) and (Paulsson *et al.* 2000).

Table 10. The tuning parameters of the genetic algorithms used in the first two-step selection procedure.

	N_{pop}	N_{gen}	P_c	P_m
GA	100	30	0.9	0.01
SPA+GA	100	30	0.8	0.025

When the selection is carried out directly from the whole data set, 4.78 features on average are selected. The occurrence of features in the identified subsets is shown in Fig. 29. As seen from the figure, some features are selected more than others but quite a lot of variation can be noticed. Thus the RBFNN models are identified with both the best solution and the most frequent solution found. When the selection is carried out from the reduced data set, the selection is almost unanimous, as shown in Fig. 30.

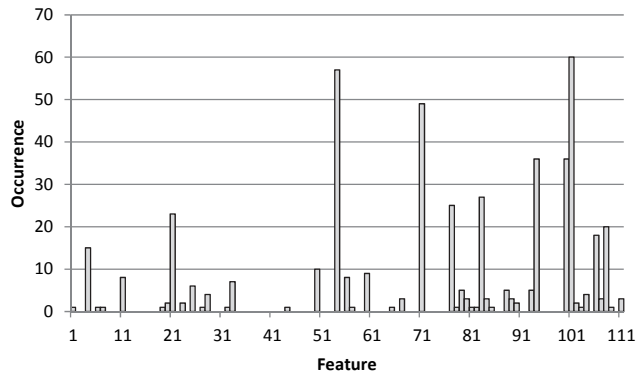


Fig. 29. The occurrence of features in the 100 repetitions of the genetic algorithm starting from the whole data set.

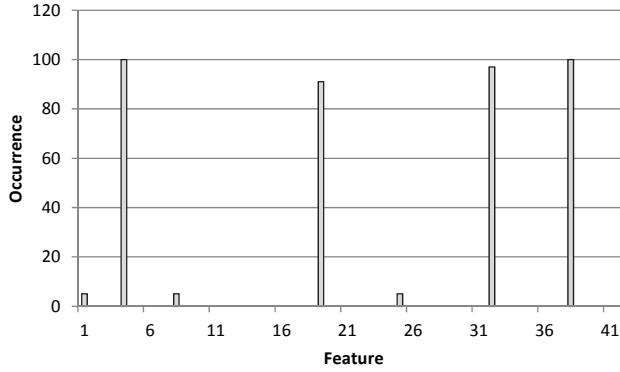


Fig. 30. The occurrence of features in the 100 repetitions of the genetic search after the application of SPA.

The RBFNN models are identified for the three subsets of features. These subsets are the best and the most frequent subsets found by the genetic algorithm when the selection starts from the whole data set and the subset found after applying SPA. The identification of the RBFNN model proceeds as follows. The activation function used in the RBFNN models is (Ramuhalli *et al.* 2002)

$$f_{ih} = \exp\left(-\frac{\|x_i - c_h\|^2}{2\sigma_h^2}\right), \quad (30)$$

where c_h is the centre of the basis function, σ_h is the spread of the basis function and $\|x_i - c_k\|$ is the Euclidean distance between the input vector and the centre of the basis function. The training of the network is an iterative process where one neuron at a time is added until the predefined number of neurons is reached (Wang and Xiang 2007). New neurons are created based on the input vector that leads to the greatest improvement in network performance. The network output can be written in a matrix format as

$$\mathbf{Y} = \mathbf{F}\mathbf{w}, \quad (31)$$

where \mathbf{w} is the weight matrix and \mathbf{F} includes the outputs of the hidden layer neurons. The weight matrix can be solved in the least squares sense as given in (4). The structural parameters (σ_h and the number of neurons) are determined in this study through systematic testing.

The performance indices of the identified RBFNN models are given in Table 11. It can be observed quite clearly that the very best solution found by the genetic algorithm when starting from the whole data set is overfitted. It contains nine features and fails in the predictions with the external validation data set. The most frequently found solution, however, includes five features and performs well with the external validation data set. Also, the solution found after the application of SPA is good. In fact, almost no deterioration in model performance with the external validation data set is noticed and thus it is safe to conclude that SPA can be used effectively in pre-selection. Also, the MLR models are identified with the selected feature subsets. These results are presented later when different modelling techniques are compared in Section 5.4.

The most frequently found solution when starting from the whole data set includes the following features: the width of the top of the trapezoidal function ($H_3 - H_2$) fitted to the moving RMS profile (x_{109}), the fitting parameters a (x_{193}), s_1 (x_{194}) and H_1 (x_{195}) of the two-sided composite function fitted to the moving entropy profile and the slope k_2 of the triangular function fitted to the moving entropy profile (x_{229}). The selected features when SPA is applied first are the RMS value (x_1), the width of the top of the trapezoidal function ($H_3 - H_2$) fitted to the moving RMS profile (x_{109}), the slope k_1 of the triangular function fitted to the moving entropy profile (x_{164}) and the fitting parameter H_1 of the two-sided composite function fitted to the moving entropy profile (x_{195}). Only two of the selected features are the same. The significance of the selected features is discussed later in Section 6.5.

Table 11. The performance indices of the RBFNN models.

Index	MSEP		Correlation coefficient	
	Training	External validation	Training	External validation
GA (best)	0.059	0.727	0.970	0.525
GA (most frequent)	0.343	0.246	0.808	0.897
SPA + GA	0.411	0.247	0.764	0.872

Forward selection and genetic algorithms in pre-selection

In this study, residual stress and hardness are predicted with an MLR model. The results of this study have not yet been published. The data set used contains 101 features of 84 measurements (data sets 1 and 2). The selection procedures used are forward selection and genetic algorithms. The selection is divided into pre-

selection and actual selection steps. Both of the procedures are tested in both steps. The objective function used is given in (29). The idea is that in the pre-selection step, an overoptimistic objective function is used. Thus, the LOO cross-validation procedure is used in pre-selection while the actual selection step uses the LMO cross-validation procedure. The tuning parameters of the genetic algorithm in the pre-selection step are: $N_{pop} = 50$, $N_{gen} = 30$, $p_c = 0.9$ and $p_m = 0.005$. The tuning parameters in the actual selection are: $N_{pop} = 50$, $N_{gen} = 30$, $p_c = 0.9$ and $p_m = 0.02$. The initial population in pre-selection is created so that it contains 10% of zeros while the initial population in the actual selection contains 50% of zeros. The genetic search is always repeated 100 times.

The results of the pre-selection show that when a genetic algorithm is applied to the whole feature set, the results vary similarly to those shown in Fig. 29. Thus, a threshold value has to be determined to identify the features for the actual selection step. In this study, the threshold is set to 25 and 12 features continue to the actual selection step when hardness is predicted. For residual stress predictions, the preselected subset contains 19 features. Deterministic forward selection always produces the same result when LOO cross-validation is used. In this case, the feature subsets continuing to the actual selection step include 11 features for hardness predictions and 15 for residual stress predictions. The preselected subsets contain some mutual features. For hardness predictions, there are six mutual features while 11 mutual features are observed for residual stress predictions.

The actual selections are carried out starting from the feature sets from the various pre-selections. The actual selection uses the LMO cross-validation procedure and thus also the results of forward selection vary. It is repeated 100 times. Table 12 shows the statistical values of the performance of the algorithms when predicting residual stress. It can be seen that using a genetic algorithm in pre-selection produces better results on average. As expected, more variability is observed when a genetic algorithm is used in the actual selection, as can be seen from the standard deviation of the result.

When the results are carefully analysed, the selected feature subset for hardness prediction includes the following features: coercivity (x_5), peak height from the moving RMS profile (x_{32}), peak position from the moving entropy profile (x_{45}) and the fitting parameters a (x_{65}) and s_1 (x_{66}) of the two-sided composite function when fitted to the moving RMS profile. This subset of features is obtained with all the approaches except FS / FS. The selected features for residual stress prediction are: Pearson mode skewness (x_{14}), crest factor (x_{28}),

the fitting parameter H_1 of the normal distribution function (x_{51}), the fitting parameters H_1 (x_{55}) and b (x_{58}) of triangular function, the fitting parameter H_4 of the trapezoidal function (x_{63}), the slope k_2 of the trapezoidal function (x_{103}) and the width of the bottom of the triangular function ($H_3 - H_1$) (x_{108}).

The prediction of hardness is not as complex a problem as the prediction of residual stresses. This can be seen from the results since the best feature subset for hardness predictions is found with all the selection schemes except FS / FS. On the other hand, using only genetic algorithms in both selection steps provided the best feature subset when predicting residual stresses. This indicates that forward selection is adequate for simple selection tasks but that genetic algorithms should be used, at least in the pre-selection step, for more complex problems.

Table 12. The statistics of objective function values from 100 repetitions of the selection procedures for residual stress predictions.

Pre-selection / selection	Mean	Standard deviation	Minimum	Maximum
FS / FS	26.19	0.59	25.09	27.38
FS / GA	27.37	0.91	23.66	29.32
GA / FS	25.44	0.40	24.34	26.62
GA / GA	24.19	0.58	23.06	26.05

5.4 Modelling techniques

In the studies by the author, many modelling techniques have been used. The results provided in this section are from the studies where different techniques have been compared. These studies have been carried out by the present author but the results have not yet been published. The majority of the studies mentioned earlier use MLR models. Linear models are used mainly due to the complexity of the Barkhausen phenomenon. The MLR model also operates on original features and thus the models obtained and the selected features can be more easily analysed. However, the study with unpublished results compares the use of partial least squares regression (PLSR), principal component regression (PCR) and radial basis function neural network (RBFNN) models. In this study, data set 3 is used containing BN measurements carried out in different directions. Altogether 115 measurements are available with 55 features. Another study also with yet unpublished results compares the use of MLR models and neural networks. In this study, data sets 1 and 2 are merged. The results with the RBFNN model are

presented in Section 5.3.5 but in this section a comparison with the MLR model is made.

Comparison of PCR, PLSR and RBFNN models

In this study, PCR, PLSR and RBFNN models are identified. The objective function used is given in (29). 10-fold cross-validation with 100 repetitions is used in the study and the median values are reported. In the development of the PCR model, the components are formed as described in Section 3.2.2. The components are added to the model based on their eigenvalues (PCR). However, eigenvalue-based addition may not lead to optimal solution (Sun 1995, Depczynski *et al.* 2000) and thus correlation-based addition (CPCR) is also tested. Table 13 shows the performance of the PCR and CPCR models with increasing complexity. It is seen that a suitable number of components for PCR and CPCR models is three and four, respectively. The CPCR model performs a little bit better but the difference is not very significant.

The PLS components are formed as described in Section 3.2.3. It is stated in the literature that it may be advantageous to set very small coefficients of components to zero (Bastien *et al.* 2005). The effect of this is also studied here as a small threshold value is defined for setting some coefficients to zero. The results with and without two different threshold values are given in 0. The table shows that two components are to be used and that the best performance is obtained with the threshold value of 0.2.

The RBFNN model uses the feature subset selected with a genetic algorithm in (Sorsa & Leiviskä 2009b) and (Sorsa *et al.* 2010b) (see Section 5.3.5). The structural parameters (σ_h and the number of neurons) are identified using the central composite experimental design (CCD). The CCD is carried out as described in (Diamond 1981). The outcome of the CCD is a response surface that can be used for identifying the optimal parameter values. Fig. 31 shows the response surface and the optimum. The RBFNN model is then identified as described in Section 5.3.5. The $SSEP_{CV}$ value of the identified neural network model is 37.6.

When comparing the results, it can be seen that the linear modelling techniques perform almost equally. A slightly better solution is found with the PLSR technique. The major difference is in model complexity because the PCR model needs at least three components while the PLSR model needs only two. However, the number of components is fairly irrelevant because both of these

models can be expressed as the usual MLR model given in (2) and (3). The nonlinear RBFNN model performs better than the linear models. A decrease of about 10–15% is observed in the $SSEP_{CV}$ value with the neural network model.

Table 13. The $SSEP_{CV}$ values of the PCR models with increasing complexity.

Number of components	1	2	3	4	5	6	7	8
PCR	67.4	50.0	44.2	45.2	45.9	46.9	47.7	47.1
CPCR	67.3	50.0	46.4	43.9	45.2	47.4	48.7	49.5

Table 14. The $SSEP_{CV}$ values of the PLSR models with increasing complexity.

Number of components	1	2	3	4	5
No threshold	58.3	45.3	54.6	88.3	95.3
Threshold = 0.1	54.3	44.1	57.1	80.5	87.7
Threshold = 0.2	49.1	43.1	58.3	76.7	80.5

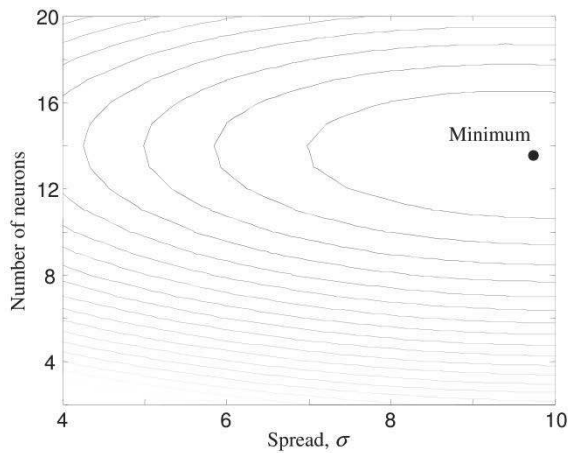


Fig. 31. The contour plot of the response surface and the optimum.

Comparison of MLR and RBFNN models

The second study by the present author compares the performance of the MLR model to the nonlinear neural network model. Features are selected for the models

using SPA-based feature elimination followed by a genetic search (Section 5.3.5). Table 15 shows the results with both models. It is clear that the nonlinear neural network model outperforms the linear modelling technique. In particular, the performance with the external validation data set is much better with the RBFNN model.

Table 15. The performance indices of the MLR and RBFNN models with the feature subsets obtained with the SPA + GA approach.

Index Data set	<i>MSEP</i>		Correlation coefficient	
	Training	External validation	Training	External validation
RBFNN	0.411	0.247	0.764	0.872
MLR	0.522	0.499	0.687	0.718

5.5 Interpretation and validity of the results

The interpretation of the results presented above may appear somewhat confusing. The published papers had different aims and thus different aspects were emphasized. In addition, the properties of the data set used influence the prediction accuracy. The results in Section 5.3 are converted into actual units (MPa for residual stress and HV for hardness) to make their comparison easier even though normalised variables are used in some of the original publications. However, the data usage varies in these studies and direct comparison is still slightly questionable. The studies presented in Section 5.4 focus on a comparison of methods. Normalised variables are used in these studies and the results are presented as in the original publications.

The presented performance indices are valid because appropriate cross-validation procedures are used in model identification. Usually, the cross-validation method used is leave-multiple-out, which is reported to be realistic (Baumann 2003). Also, the studies where an external validation data set is used show that the results are not too optimistic.

When a limited number of data points is available, the risk of overfitting the model increases. The risk is even higher if the number of model parameters is great. In the studies presented in this section, any possible overfitting is avoided in two ways: realistic cross-validation procedures are used in model identification, and the models are kept simple enough that the number of parameters is low compared to the number of data points.

5.6 Summary of the results

This section has presented the results obtained in many studies carried out by the author. The results have been presented considering the different steps of the overall modelling scheme presented in Fig. 20. The outcome of the presented scheme is a model that can be used in predicting changes in material properties. Residual stress and hardness changes are predicted in the studies presented here.

The first step is feature generation where different mathematical procedures are used to extract knowledge from the BN signals measured. Many different kinds of features are calculated. Some are obtained directly from the measuring device; some are calculated directly from the BN signal, while others are calculated from different BN profiles. The outcome of this step is a large set of features. This set includes noisy, collinear and irrelevant features and thus the selection of features must be carried out carefully.

Feature selection is a complex problem, consisting of three components: the search engine, the mathematical modelling technique and the objective function. In the literature, a variety of methods have been proposed for use in feature selection. These methods are divided roughly into filters and wrappers. In the results presented here, wrapper approaches are used because they are expected to give better results. The methods used are forward selection, backward elimination, simulated annealing and genetic algorithms. As expected, based on the literature, backward elimination alone does not perform well. Forward selection, on the other hand, is efficient and the results obtained are reasonable. However, even better results are obtained with the stochastic methods, simulated annealing and genetic algorithms. These methods, however, are more expensive computationally. For example, the forward-selection scheme described in Section 5.3.2 needs 483 objective function evaluations to reach the solution, while a single run of the genetic algorithm described in Section 5.3.4 needs 4000 evaluations. The computational cost is, however, case-dependent. The result of the genetic search also varies as shown in this section. This leads to the need for careful examination of the results after repeated runs of the search.

Feature selection is a large-scale problem and thus it may be advantageous to perform it in two steps. In the first step, the dimension of the problem is decreased, while the second step performs the actual selection. The results of a couple of such approaches are presented here. In these studies, the SPA procedure, forward selection and genetic algorithms have been used in the first step and mainly a genetic algorithm in the second step. With these approaches, the

variability in results when using a genetic algorithm has been decreased. Consequently, these approaches seem to be usable when reliable feature selection is desired.

In the studies, different modelling techniques have been tested. In general, MLR models are used because of the complexity of the Barkhausen phenomenon. However, the results indicate that nonlinear neural network models perform better and thus are recommended for use. However, the feature selection step is challenging when applied to a neural network model.

6 Discussion and future work

The previous section provided a lot of results. In this section, these results are discussed and the methods used are compared. The analysis of the selected features is carried out only for the predictions of residual stress because the results for hardness predictions are limited. In addition, future prospects are presented.

6.1 Feature generation

The feature generation procedure itself does not require a lot of discussion. The mathematical procedures used produce a large set of features and the results show a lot of variability, indicating that the prediction task is very case-dependent. Thus, it is quite difficult to exclude any feature from the set. In fact, it seems that prediction accuracy can be further improved by introducing new features.

The procedure of fitting functions to BN profiles seems to be beneficial because in the majority of the studies, some fitting parameter is selected. The parameters most typically selected can be referred to as peak position (such as H_l of normal distribution) or peak width (such as s of normal distribution). However, the fitting parameters selected also refer to BN properties which are not typically included in BN studies, such as the initiation or termination of Barkhausen activity (for example H_l or H_d of the trapezoidal function) or the slope of the descending or ascending side of BN profiles.

Section 5.2 presented the features that are used in the studies presented in this thesis. As mentioned there, these features are mainly obtained through the time domain analysis of the signal. More time domain features can be obtained by analysing the histograms of the BN signal (such as shown in Fig. 21). The data points in each segment can be used as features. The difficulty is, however, that a fixed or relative bin size can be used. It seems that relative bin size (for example related to standard deviation) can produce more suitable features. That is because the level of background noise (and the whole signal) varies and thus a fixed bin size must be defined for each measurement individually. It has been noticed that the background noise usually varies between $\pm 2s$ (s is the standard deviation of the signal) and thus the higher multiples of s could possibly be used as the bin edges when producing the features.

Another approach for obtaining more time domain features is to take a derivative of the signal (Juuso & Lahdelma 2010) and repeat the feature

generation procedures described in Section 5.2 for this signal. It is also possible to use the second or even higher derivatives. This approach has not been studied yet but is worth studying in the future.

Features can also be obtained through a frequency domain analysis. The frequency domain signal is obtained through the Fourier transformation. The frequency domain signal can then be presented as a power spectrum of the signal from which the power spectral density (PSD) can be obtained. Frequency domain analysis is used, for example, in (Davut & Gür 2007) and (Santa-aho *et al.* 2011). It can also be used to obtain a BN profile (Piotrowski *et al.* 2010). The spectrum presents the amplitude of the signal as a function of frequency. Thus, it is also possible to obtain the spectral density of some well-defined frequency interval (Yamaura *et al.* 2001). Such a feature could be very informative but unfortunately the identification of a suitable interval can be very challenging and case-dependent. However, this aspect needs more research.

Sophisticated data analysis techniques such as wavelets may be used in feature generation. Approaches already exist where a wavelet transformation is used to obtain features used as input variables for a prediction model (Chu *et al.* 2007, Piotrowski *et al.* 2010). The use of wavelets needs to be studied thoroughly.

6.2 Feature selection procedures

In the studies presented in Section 5.3, many different feature selection methods have been used. These methods include manual selection, forward selection, simulated annealing and genetic algorithms and two-step algorithms. The outcome of these varies.

Manual selection

Manual selection is based on knowledge about the modelled phenomenon. The prediction models identified in the studies presented include 3–7 features. The model where only three features are used has a slightly more complex model structure, as the interaction terms of the features are also included. The identified MLR models have reasonable prediction accuracy. However, the solutions found are not optimal because other methods give even better results. For example, the same data sets are used in (Sorsa *et al.* 2012a) and (Sorsa *et al.* 2012b). The first one uses manual selection and the second one forward selection. The correlation

coefficient between the measured and predicted residual stresses is 0.91 and 0.94, respectively, for data set 2.

Forward selection

Forward selection is very efficient in computational terms, as stated in the literature (Guyon & Elisseeff 2003). It is also usually robust against overfitting (Guyon & Elisseeff 2003). The drawback of the method is that when a feature is selected, its relevance is not questioned again. Thus, it tends to find features that are strong alone but fails in the selection of features that are weak alone but strong together with some other feature (Kohavi & John 1997). These observations are also seen in the results presented here. In (Sorsa *et al.* 2011a), forward selection found a subset with five features while simulated annealing and genetic algorithms found a subset with eight features. The latter solution is evaluated to be better in this case but it is obvious that the solution found by forward selection is suboptimal but not overfitted. Also, when forward selection is used in pre-selection (Section 5.3.5), it has been observed that some significant features are not found (Unpublished). Despite the drawbacks, the studies show that forward selection can successfully be applied in BN applications. This holds true especially with simpler prediction tasks such as the prediction of hardness as presented in Section 5.3.5. The performance of forward selection can be further improved if the suggestions in (Nakariyakul & Casasent 2009) are implemented. However, these improvements increase the computational load, which may be unacceptable in some cases. Thus, the following conclusions about forward selection in BN applications can be made:

- It outperforms manual selection and finds reasonable feature subsets.
- The solution found may be suboptimal but still usable.
- Its computational efficiency makes it interesting with regard to on-line applications.
- Its performance can be improved at the expense of the computational load.

Simulated annealing

The results with simulated annealing are quite limited but show that it gives results comparable to genetic algorithms. However, it has been noticed that the feature subsets obtained are bigger on average than those found by genetic

algorithms (Sorsa *et al.* 2011a). Thus, there may be a greater chance of overfitting. Compared to genetic algorithms, the use of simulated annealing is advantageous because it is guaranteed to converge to an optimal solution (Alexandridis *et al.* 2005). In spite of this, studies found in the literature seem to prefer genetic algorithms over simulated annealing.

Genetic algorithms

The results presented show that genetic algorithms give the best results in general. However, the drawbacks presented in the literature are valid. Fig. 27 and Fig. 29 clearly show that the feature subset found has great variability. Fig. 27 together with the results in (Sorsa & Leiviskä 2009b) and (Sorsa *et al.* 2010b) show that the selected subset may also contain some irrelevant features. The computational load of the genetic algorithm is significant and becomes even more so if the algorithm needs to be repeated due to the variability in the results. If the computational cost is significant and it is desired that a single run of a genetic algorithm produce a good and reliable solution (such as in on-line applications), the variability of the results must be decreased. Thus, it has been shown that it is advantageous to use two-step algorithms when applying genetic algorithms. The studied pre-selection approaches are SPA-based feature elimination, forward selection and genetic algorithms. All the pre-selection schemes decreased the variability of the results. However, it has been observed that forward selection missed some significant features and thus led to a suboptimal result. In contrast, the other two approaches performed well. Based on these observations, the following conclusions about the use of a genetic algorithm can be drawn:

- It gives better results than forward selection.
- The computational load may be too high for some applications.
- It is advantageous to use some pre-selection procedure before the actual selection.

Two-step procedures

Two-step procedures are advantageous in a couple of ways. Firstly, it is possible to use some computationally efficient method to reduce the dimensionality of the search problem (Llobet *et al.* 2007). Secondly, the selection problem is usually a compromise between model accuracy and parsimony and thus these aspects can

be highlighted in the different steps (Alexandridis *et al.* 2005). The two-step procedures presented here give good results in general. A genetic search following pre-selection gives results that are almost unanimous. This result also holds true for simulated annealing. The pre-selection procedures used are quite different. The SPA procedure is computationally very efficient because it is a filter approach. It leads, however, to a small decrease in the model performance as shown in Table 11. Forward selection used in pre-selection does not perform that well. Indeed, some significant features are missed and thus the actual selection is not able to find the optimal solution. When a genetic algorithm is used in pre-selection, the best results are obtained but the computational load is increased significantly. The results, however, indicate that it is adequate to use a genetic algorithm only in pre-selection and the actual selection can utilise computationally more efficient methods such as forward selection. The following conclusions can be drawn about two-step procedures:

- Their use is recommended if a single run of a genetic algorithm (or simulated annealing) would give a reliable solution.
- SPA is computationally very efficient but may lead to a small decrease in model performance.
- Forward selection is inadequate in pre-selection because some useful feature combinations may be lost.
- Genetic algorithms in pre-selection give good results but the computational load increases.

6.3 Modelling techniques and model accuracy

The modelling techniques used are multivariable linear regression (MLR), principal component regression (PCR), partial least squares regression (PLSR) and artificial neural networks (ANN). The majority of the studies use linear methods but the use of neural networks is considered in the later studies.

The MLR models are seen to capture the major interactions between BN and the material properties studied. The Barkhausen phenomenon is complex and there are major interactions between the features and the material properties under study. The data sets are typically quite small and thus there is a high probability of chance correlations. To avoid overfitting, the use of simple MLR models is suggested. Also, the selection of suitable features should use some cross-

validation procedure to avoid excess features being added due to chance correlations.

The unpublished results presented indicate that it is not recommended to use PCR and PLSR models. The results in (Sorsa & Leiviskä 2009b) and (Sorsa *et al.* 2010b) even indicate that the MLR model also outperforms radial basis function neural networks. At this point it should be recognised that these conclusions are not necessarily valid, for the following reasons. Firstly, the comparative studies mentioned above were carried out in the early stages of the research and thus the understanding of the Barkhausen phenomenon and the data sets was limited. Secondly, the data set used (data set 3) is very challenging when it comes to predictions because it includes measurements in different directions. A challenging data set may favour the simpler modelling approaches. Also, the more recent (but yet unpublished) study where a simpler data set was used showed that the use of RBFNN is to be preferred (Section 5.4). However, there is no evidence that PCR and PLSR models perform better than MLR models. Both of these model structures operate with components, which makes the interpretation of the model harder. Also, a suitable number of components needs to be determined. However, it is worth the effort to re-evaluate the usability of these model structures in the future. The effect of feature selection on the outcome of these modelling techniques should be investigated. In addition, the components used when building a PCR model can be selected with the methods described in this thesis and in (Barros & Rutledge 1998).

As mentioned before, the earlier studies indicate that neural networks do not improve the model accuracy while the more recent studies clearly indicate the opposite. Indeed, the use of neural networks must be researched more in the future. Only the RBFNN models have been used so far, so other neural network model structures may lead to even better results. It is also definitely worth integrating the neural network model structure identification procedure into the feature selection procedure. RBFNN model development can be carried out by utilizing the fuzzy c-means algorithm, as reported in (Alexandridis *et al.* 2005).

6.4 Objective functions

The objective functions used are based on the squared error. Sum, mean and root mean squared error of prediction (SSEP, MSEP and RMSEP, respectively) criteria have all been used. If these criteria are applied directly to the training data set, they produce overoptimistic results as stated in the literature (Baumann 2003,

Guyon & Elisseeff 2003) and in Section 3.6. Therefore, resampling cross-validation procedures are used.

The resampling cross-validation methods used are leave-one-out (LOO), k -fold and leave-multiple-out (LMO). The literature states that the LOO procedure still produces overoptimistic results (Baumann 2003). K -fold is also reported to do so (Baumann 2003) but this depends on the value of k . Indeed, the 2-fold procedure equals the LMO procedure where half of the data points are left out for validation. It is, however, typical that five- or 10-fold procedures are used which may provide overoptimistic results. The LMO procedure is reported to give realistic results. This is also seen from the results in (Sorsa *et al.* 2012b) where the feature subset selected (residual stress prediction) includes only five features and is directly applicable for prediction purposes. However, the results depend on the number of data points left out for validation. The suggestion is that 40–60% of data points are used for training and the rest used for validation (Baumann 2003). The drawback of the LMO procedure (and also k -fold) is that the random data split may have a serious effect on the estimated values. Thus, multiple data splits must be done, which increases the computational load of the methods (Baumann 2003). For example, one suggestion is that at least $2N$ (N is the number of data points) data splits are used in the LMO procedure (Baumann 2003).

The results obtained prove these observations. For example, the pre-selection in the latter study in Section 5.3.5 used LOO cross-validation and resulted in a subset of 19 features whereas the LMO procedure reduced the subset to eight features. It is observed in (Sorsa & Leiviskä 2010) and (Sorsa & Leiviskä 2011c) that five-fold cross-validation also produced overoptimistic results because it indicated that 10–12 features were to be used while the subsequent elimination step reduced the number of features to five. Even though these methods may provide overoptimistic results they also have some advantages. The first advantage is that the chance of underfitting is negligible (Wisnovski *et al.* 2003) and thus they can be safely used for data reduction purposes as in the yet unpublished study described in Section 5.3.5. The LOO procedure is also computationally efficient because the data split is always the same and thus no repetitions are needed. To balance the overoptimistic performance criteria, it is also possible to add a term to the objective function that penalizes for excess features (Kohavi & John 1997, Barros & Rutledge 1998). A penalty constant with good results is also used in (Sorsa & Leiviskä 2009b) and (Sorsa & Leiviskä 2010b).

6.5 The significance of the selected features

The feature selection results show great variability, as presented in Section 5.3. Nevertheless, some features remain the same regardless of the study. Furthermore, the selected features may vary but still they correspond to certain BN properties. Based on the literature, the overall Barkhausen activity (Lindgren & Lepistö 2001, Moorthy *et al.* 2004b, Stewart *et al.* 2004), the shape and position of the BN profile (Moorthy *et al.* 2003, Stewart *et al.* 2004, Wilson *et al.* 2009) and some specific points of the profile (Wilson *et al.* 2009), for example are influenced by changes in residual stress. Table 16 shows the selected features for residual stress predictions categorized as above.

Some observations can be made from Table 16. Firstly, in almost every selected feature subset there is a feature corresponding to overall Barkhausen activity. This is to be expected based on the literature because in the majority of studies a feature corresponding to overall Barkhausen activity is used. Based on Table 16, the most usual such feature in the studies presented here is the crest factor. The second observation is that peak width too is represented by some feature in all of the subsets. Usually the feature is FWHM but also the width of the profile top ($H_3 - H_2$) and the fitting parameters referring to the spread of the fitted functions are sometimes selected. Also, peak position, coercivity or H_I from the normal distribution (f_1 in Table 16) is present in almost all the feature subsets. The two prior features are reported to be closely related (Stewart *et al.* 2004, Davut & Gür 2007) and obviously the fitting parameter mentioned is closely related to peak position. Thus, it is a little controversial that two of these features are sometimes selected at the same time. Based on the selections, it seems that the other possibly important points in the BN profile are the field strength where the Barkhausen activity initiates and terminates. Table 16 shows that even though skewness has been used (Stewart *et al.* 2004), it is quite seldom selected.

Table 16 shows that if a fitting parameter is selected, it is quite often obtained from the trapezoidal (f_3) or the two-sided composite (f_4) function. This is quite interesting and may be due to the shape of these functions, which may be closest to the actual BN profile. The normal distribution function assumes a symmetrical profile and the triangular function assumes a sharp-peaked profile. The trapezoidal and two-sided composite functions, however, do not make any assumptions about the shape of the profile and thus can probably produce the most useful features.

To evaluate whether the selected features behave as reported in the literature, Table 17 shows the regression coefficients of the features most usually used. The magnitudes of the coefficients cannot be compared in this case because the data sets have undergone different pre-processing procedures. Instead, the directions of the interactions can be evaluated. Table 17 shows that the regression coefficient of the features corresponding to overall Barkhausen activity is always positive. This is in accordance with the literature as it is commonly accepted that Barkhausen activity decreases with increasing compressive stress and increases with increasing tensile stress (for example Lindgren & Lepistö 2001, Moorthy *et al.* 2004, Stewart *et al.* 2004, Blaow *et al.* 2004).

The regression coefficients of the features corresponding to peak width are usually negative. This is in accordance with the literature since it has been reported that increased compression shifts the profile to higher field strengths and thus increases the value of the peak position (Stewart *et al.* 2004). It is quite surprising that the coefficient is positive for data set 2 in (Sorsa *et al.* 2012a). The reason for this is not clear but it is possible that the other selected features have this kind of influence on the FWHM value. An interesting detail is seen in the coefficients of s_1 and s_2 from the two-sided composite function. The former refers to the width of the ascending side of the profile and the latter to the width of the descending side. The sign of the latter is negative while the sign of the former is positive. Thus, it is possible that the changes in peak width are usually stronger in the descending side of the profile although sometimes the change may also be strong in the ascending side. In such cases, the width changes may cancel each other out or the influence of peak width may be reversed. However, this is only hypothetical and more research is needed to verify this observation.

As shown in Table 17, the regression coefficient of the features related to peak position is usually positive, which is not in accordance with the literature. It has been reported that coercivity (and thus also peak position) increases with increasing compressive stress (Stewart *et al.* 2004, O'Sullivan *et al.* 2004). This may be due to other, unmeasured material properties which affect the peak position because changes in coercivity are usually associated with changes in hardness (O'Sullivan *et al.* 2004, Davut & Gür 2007). A drastic change is observed in the regression coefficients for data sets 1 and 2 in (Sorsa *et al.* 2012b). This change is similar to that observed with the FWHM value above. This indicates that there are some serious changes in microstructure between these two series of samples. These changes seem to modify the interactions between the

features. Consequently, depending on the selected subset of features, drastic differences may be observed.

Table 16. The features selected in the studies presented in Section 5.3 for prediction of residual stress.

Overall Barkhausen activity	Profile shape	Certain points of the profile	Other	Reference
BN entropy		coercivity, peak position		Sorsa <i>et al.</i> 2010a
crest factor	FWHM, s_2 from f_4 *	coercivity, H_3 from f_3 , H_2 from f_4	Pearson mode skewness	Sorsa <i>et al.</i> 2012a
RMS	s_1 and s_2 from f_4	peak position, H_1 from f_3	skewness	Sorsa <i>et al.</i> 2008a
	FWHM	coercivity, remanence, permeability, peak position		Sorsa & Leiviskä 2010, Sorsa & Leiviskä 2011c
crest factor, peak height	k_2 from f_2	H_1 from f_3 , H_1 from f_1		Sorsa <i>et al.</i> 2011a (FS)
crest factor	FWHM	coercivity, peak position, H_4 from f_3		Sorsa <i>et al.</i> 2012b
signal-to-noise ratio **, a from f_3 ***	FWHM	H_1 , H_2 and H_4 from f_3 , H_1 from f_1		Sorsa & Leiviskä 2011a (SA and GA)
RMS		coercivity, H_1 from f_3	loop area	Sorsa & Leiviskä 2009b, Sorsa & Leiviskä 2010b
a from f_4 ***	$H_3 - H_2$ from f_3 , s_1 from f_4 , k_2 from f_2	H_1 from f_4		Unpublished Section 5.3.4
RMS	$H_3 - H_2$ from h_3 , k_1 from f_2	H_1 from f_4		Unpublished Section 5.3.5
crest factor, b from f_2 ***	k_2 from f_3 , $H_3 - H_1$ from f_2	H_1 from f_1 , H_1 from f_2 , H_4 from f_3	Pearson mode skewness	Unpublished Section 5.3.5

* The fitted functions are denoted here by: f_1 - normal distribution, f_2 - triangular function, f_3 -trapezoidal function and f_4 - two-sided composite function.

** Signal-to-noise ratio is closely related to the reciprocal of RMS.

*** The fitting parameters can be associated with the overall Barkhausen activity.

Table 17. The regression coefficients of certain features.

Overall Barkhausen activity	Peak width	Peak position	Reference
crest factor: 1.46		peak position: -0.84	Sorsa <i>et al.</i> 2010a
crest factor: 50.9	FWHM: -72.5	coercivity: 36.1	Sorsa <i>et al.</i> 2012a (Data set 1)
crest factor: 392.8	FWHM: 99.7	coercivity: 55	Sorsa <i>et al.</i> 2012a (Data set 2)
RMS: 0.7	s_1 : 0.05 / s_2 : -0.32 FWHM: -0.881	peak position: 0.09 peak position: 0.276	Sorsa <i>et al.</i> 2008a Sorsa & Leiviskä 2010, Sorsa & Leiviskä 2011c
crest factor: 83.2	FWHM: -49.7	peak position: 24.8	Sorsa <i>et al.</i> 2012b (Data set 1)
crest factor: 263.2	FWHM: -26.0	peak position: -147.7	Sorsa <i>et al.</i> 2012b (Data set 1)

6.6 Meeting the objectives of the research

The goal of this work was to find a suitable methodology for predicting material properties based on Barkhausen noise measurements as described in Section 1.2. This objective basically means the development of efficient methodology for feature generation and identification of suitable feature selection procedures. Feature generation is carried out mainly from the time domain signal. However, the results show that the feature set obtained includes information that can be used to predict material properties with reasonable accuracy. Thus it can be concluded that the applied feature generation is adequate even though some future work is also needed. The feature selection task is widely studied in the literature. The methods used in these studies do not cover the whole field but nevertheless essential knowledge considering the methods and their usability is gained. The methods can be used in the future as presented or as a reference when studying some other potential methods. The combination of the basic selection methods presented also provides potential approaches for improving the efficiency and reliability of the selection task. The potential is shown in this thesis by a couple of successful studies. Consequently, it can be concluded that the objectives set for the research have been met.

7 Conclusions

Materials can be tested destructively or non-destructively. Non-destructive testing is obviously preferable if it can be applied because it does not compromise the future use of the material. For ferromagnetic materials, an intriguing non-destructive testing method is Barkhausen noise measurement, where the tested sample is placed in an external varying magnetic field. The varying magnetic field causes the magnetic domains within the material to change their sizes and orientations. These changes are hindered by pinning sites. The phenomenon of domain walls getting trapped in these pinning sites and then breaking out of them causes abrupt changes in the magnetization of the measured sample. These abrupt changes lead to a noise-like signal, which is recorded.

Analysis of the Barkhausen noise signal usually utilises some computational features. Typical features are, for example, the RMS value and peak height, width and position. The three latter are obtained from the Barkhausen noise profile. The calculated features are compared to changes in the studied material properties. It is, however, typical that only qualitative comparisons are carried out. Better utilisation of the method needs quantitative results.

The goal of the research presented in this thesis is to find a suitable methodology for quantitative prediction of material properties based on the Barkhausen noise signal. It is a very challenging task due to the complex interactions between material properties and Barkhausen noise. This thesis considers a prediction scheme that is divided into the following steps: feature generation, feature selection together with model identification and model validation. The first step uses different mathematical procedures to extract information from the Barkhausen noise signal. The outcome of feature generation is a large set of features, from which the most suitable ones are to be selected for the prediction model.

Feature selection is a complex task that requires three elements: a mathematical modelling technique, an objective function and a search engine. In the studies presented in this thesis, the modelling techniques used are multivariable linear regression (MLR), partial least squares regression (PLSR), principal component regression (PCR) and artificial neural networks (ANN). The search engines used are manual selection, forward selection, simulated annealing and genetic algorithms. The objective function is based on the squared error criteria. These criteria are obtained through cross-validation procedures because the number of data points is limited and thus a static split would decrease the

amount of data for training and validation purposes significantly. The procedures used are leave-one-out, k-fold and leave-multiple-out cross-validations.

The results of the studies presented show that the proposed modelling scheme can be used in predicting material properties based on Barkhausen noise measurement. A closer look at the results show that the feature generation procedures used are applicable but that more features can improve the prediction accuracy of the models. More features can be obtained through a more thorough time domain analysis of the signal, through frequency domain analysis or by using more sophisticated tools such as wavelets. The comparison of the selection procedures reveals that even though forward selection gives rather good solutions, the identified solution is only suboptimal. Better results are obtained with simulated annealing and genetic algorithms. However, the improvement in the results is obtained at the expense of the computational load, which increases drastically with these methods. The problem with simulated annealing and genetic algorithms is that the obtained result varies. Consequently, two-step algorithms are proposed in this thesis. Based on the results, the pre-selection step can successfully utilise the successive projections algorithm or genetic algorithms. The use of a genetic algorithm in pre-selection obviously increases the computational load even more.

The majority of the studies presented used MLR models. The early results indicated that the MLR model outperforms all the other modelling techniques. However, the later studies have shown that the use of ANN models improves the prediction accuracy. Therefore, it is worth re-evaluating the applicability of PLSR and PCR modelling techniques in the future. Furthermore, the utilisation of ANN models should be studied further.

The use of different objective functions revealed that LOO and k-fold procedures easily lead to overoptimistic results. However, these methods are still usable because they can be used effectively in pre-selection. Also, these cross-validation procedures can be complemented with a penalty term that favours the more parsimonious solutions and thus compensates the over-optimism. The LMO procedure is shown to lead to more realistic solutions but at the expense of the computational load.

Overall, the results obtained are very promising. The models identified mainly include reasonable terms and the prediction accuracy is fairly good considering the challenge. However, some conflicts with the literature have been revealed as the application of Barkhausen noise measurement is highly case-dependent. In addition, changes in unmeasured material properties may lead to

the unexpected behaviour of some features. Finally, it can be concluded that the objectives set for this research have been well met.

References

- Alexandridis A, Patrinos P, Sarimveis H & Tsekouras G (2005) A two-stage evolutionary algorithm for variable selection in the development of RBF neural network models. *Chemometrics and Intelligent Laboratory Systems* 75: 149–162.
- Amato S & Vinzi VE (2003) Bootstrap-based \hat{Q}_{kh}^2 for the selection of components and variables in PLS regression. *Chemometrics and Intelligent Laboratory Systems* 68: 5–16.
- Anderson JC, Leaver KD, Rawlings RD & Alexander JM (1990) *Materials Science*, fourth edition. London, Chapman & Hall.
- Anderssen E, Dyrstad K, Westad F & Martens H (2006) Reducing over-optimism in variable selection by cross-model validation. *Chemometrics and Intelligent Laboratory Systems* 84: 69–74.
- Anglada-Rivera J, Padovese LR & Capó-Sánchez J (2001) Magnetic Barkhausen Noise and hysteresis loop in commercial carbon steel: influence of applied tensile stress and grain size. *Journal of Magnetism and Magnetic Materials* 231: 299–306.
- Araújo MCU, Saldanha TCB, Galvão RKH, Yoneyama T, Chame HC, Visani V (2001) The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometrics and Intelligent Laboratory Systems* 57: 65–73.
- Augustyniak B, Piotrowski L, Chmielewski M, Kosmas K & Hristoforou E (2010) Barkhausen noise properties measured by different methods for deformed Armcoc samples. *IEEE Transactions on Magnetics* 46: 544–547.
- Babuska R & Verbruggen H (2003) Neuro-fuzzy methods for nonlinear system identification. *Annual Reviews in Control* 27: 73–85.
- Bach G, Goebbels K & Theiner WA (1988) Characterization of Hardening Depth by Barkhausen Noise Measurements. *Materials Evaluation* 46: 1576–1580.
- Barros AS & Rutledge DN (1998) Genetic algorithm applied to the selection of principal components. *Chemometrics and Intelligent Laboratory Systems* 40: 65–81.
- Bastien P, Vinzi VE & Tenenhaus M (2005) PLS generalized linear regression. *Computational Statistics and Data Analysis* 48: 17–46.
- Baumann K, Albert H & von Korff M (2002) A systematic evaluation of the benefits and hazards of variable selection in latent variable regression. Part I. Search algorithm, theory and simulations. *Journal of Chemometrics* 16: 339–350.
- Baumann K (2003) Cross-validation as the objective function for variable-selection techniques. *TrAC Trends in Analytical Chemistry* 22: 395–406.
- Bastien P, Vinzi VE & Tenenhaus M (2005) PLS generalized linear regression. *Computational Statistics and Data Analysis* 48: 17–46.
- Benoudjit N, Cools E, Meurens M & Verleysen M (2004a) Chemometric calibration of infrared spectrometers: selection and validation of variables by non-linear models. *Chemometrics and Intelligent Laboratory Systems* 70: 47–53.

- Benoudjit N, François D, Meurens M & Verleysen M (2004b) Spectrophotometric variable selection by mutual information. *Chemometrics and Intelligent Laboratory Systems* 74: 243–251.
- Bi Y, Govindaraju MR & Jiles DC (1997) The dependence of magnetic properties on fatigue in A533B nuclear pressure vessel steels. *IEEE Transactions on Magnetics* 33: 3928–3930.
- Birkett AJ, Corner WD, Tanner BK & Thompson SM (1989) Influence of plastic deformation on Barkhausen power spectra in steels. *Journal of Physics D: Applied Physics* 22: 1240–1242.
- Błachnio JB, Dutkiewicz J & Salamon A (2002) The effect of cyclic deformation in a 13% Cr ferritic steel on structure and Barkhausen noise level. *Materials Science and Engineering A* 323: 83–90.
- Blaow M, Evans JT & Shaw B (2004) Effect of deformation in bending on magnetic Barkhausen noise in low alloy steel. *Materials Science and Engineering A* 386: 74–80.
- Blaow M, Evans JT & Shaw BA (2005) Surface decarburisation of steel detected by magnetic Barkhausen emission. *Journal of Materials Science* 40: 5517–5520.
- Blaow M, Evans JT & Shaw BA (2006) Effect of hardness and composition gradients on Barkhausen emission in case hardened steel, *Journal of Magnetism and Magnetic Materials* 303: 153–159.
- Blaow M, Evans JT & Shaw BA (2007) The effect of microstructure and applied stress on magnetic Barkhausen emission in induction hardened steel. *Journal of Materials Science* 42: 4364–4371.
- Boyabatli O & Sabuncuoglu I (2004) Parameter selection in genetic algorithms. *Journal of Systems, Cybernetics and Informatics* 2: 78–83.
- Broadhurst D, Goodacre R, Jones A, Rowland JJ & Kell DB (1997) Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry. *Analytica Chimica Acta* 348: 71–86.
- Capó-Sánchez J, Pérez-Benitez JA, Padovese LR & Serna-Giraldo C (2004) Dependence of the magnetic Barkhausen emission with carbon content in commercial steels. *Journal of Materials Science* 39: 1367–1370.
- Chang W-D (2007) Nonlinear system identification and control using a real-coded genetic algorithm. *Applied Mathematical Modelling* 31: 541–550.
- Chen ZJ, Jiles DC & Kameda J (1994) Estimation of fatigue exposure from magnetic coercivity. *Journal of Applied Physics* 75: 6975–6977.
- Chu Y-H, Lee Y-H & Han C (2004) Improved quality estimation and knowledge extraction in a batch process by bootstrapping-based generalized variable selection. *Industrial & Engineering Chemistry Research* 43: 2680–2690.
- Chu Y-H, Kim D, Han C & Yoon E-S (2007) Two-stage variable selection using the wavelet transform of batch trajectories for data interpretation and construction of parsimonious quality-estimation models. *Industrial & Engineering Chemistry Research* 46: 7188–7197.
- Cover TM & Thomas JA (2005) *Elements of Information Theory*. Hoboken, N.J, Wiley.

- Cullity BD (1972) Introduction to magnetic materials. Reading MA, Addison-Wesley Publishing Company.
- D'Amato C, Verdu C, Kleber X, Regheere G & Vincent A (2003) Characterization of Austempered Ductile Iron Through Barkhausen Noise Measurements. *Journal of Nondestructive Evaluation* 22: 127–139.
- Davis L (1991) Handbook of genetic algorithms. New York, Van Nostrand Reinhold.
- Davut K & Gür CH (2007) Monitoring the Microstructural Changes During Tempering of Quenched SAE 5140 steel by Magnetic Barkhausen Noise. *Journal of Nondestructive Evaluation* 26: 107–113.
- Deb K & Agrawal S (1999) Understanding interactions among genetic algorithms parameters. In: Banzhaf W., Reeves C. (Eds.) *Foundations of Genetic Algorithms*, vol. 5. San Francisco, Morgan Kaufman Publishers, Inc.
- Debusse JCW & Rayward-Smith VJ (1997) Feature subset selection within a simulated annealing data mining algorithm. *Journal of Intelligent Information Systems* 9: 57–81.
- Deep K & Thakur M (2007) A new crossover operator for real coded genetic algorithms. *Applied Mathematics and Computation* 188: 895–911.
- Depeczynski U, Frost VJ & Molt K (2000) Genetic algorithms applied to the selection of factors in principal component regression. *Analytica Chimica Acta* 420: 217–227.
- Desvaux S, Duqunoy M, Gualandri J & Ourak M (2004) The evaluation of surface residual stress in aeronautic bearings using the Barkhausen noise effect. *NDT&E International* 37: 9–17.
- Dhar A, Clapham L & Atherton DL (2001) Influence of uniaxial plastic deformation on magnetic Barkhausen noise in steel. *NDT & E International* 34: 507–514.
- Diamond WJ (1981) *Practical Experiment Designs for Engineers and Scientists*. New York, Van Nostrand Reinhold.
- Dieterle F, Kieser B & Gauglitz G (2003) Genetic algorithms and neural networks for the quantitative analysis of ternary mixtures using surface plasmon resonance. *Chemometrics and Intelligent Laboratory Systems* 65: 67–81.
- Dubois & Fiset (1995) Evaluation of case depth on steels by Barkhausen noise measurements. *Materials Science and Technology* 11: 264–267.
- Durin G & Zapperi S (2006) The Barkhausen effect. In: Bertotti G & Mayergoz I (eds) *The science of hysteresis*. Amsterdam, Elsevier: 181–267.
- Efron B & Tibshirani RJ (1993) *An introduction to the bootstrap*. New York, Chapman & Hall, Inc.
- Eiben AE, Michalewicz Z, Schoenauer M & Smith JE (2007) Parameter Control in Evolutionary Algorithms. *Studies in Computational Intelligence (SCI)* 54: 19–46.
- Eklöv T, Mårtensson P & Lundström I (1999) Selection of variables for interpreting multivariate gas sensor data. *Analytica Chimica Acta* 381: 221–232.
- Foca G, Cocchi M, Li Vigni M, Caramanico R, Corbellini M & Ulrici A (2009) Different feature selection strategies in the wavelet domain applied to NIR-based quality classification models of bread wheat flours. *Chemometrics and Intelligent Laboratory Systems* 99: 91–100.

- Fogel DB (2000) *Evolutionary computation: toward a new philosophy of machine intelligence*. Piscataway, IEEE Press.
- Galvão RKH, Araújo MCU, Fragoso WD, Silva EC, José GE, Soares SFC & Paiva HM (2008) A variable elimination method to improve the parsimony of MLR models using the successive projections algorithm. *Chemometrics and Intelligent Laboratory Systems* 92: 83–91.
- Gardner JW, Boilot P & Hines EL (2005) Enhancing electronic nose performance by sensor selection using a new integer-based genetic algorithm approach. *Sensors and Actuators B: Chemical* 106: 114–121.
- Garthwaite PH (1994) An interpretation of partial least squares. *Journal of the American Statistical Association* 89: 122–127.
- Gatelier-Rothea C, Chicois J, Fourgeres R & Fleischmann PP (1998) Characterization of pure iron and (130 PPPPM.) carbon-iron binary alloy by Barkhausen noise measurements: Study of the influence of stress and microstructure. *Acta Materialia* 46: 4873–4882.
- Gauchi J-P & Chagnon P (2001) Comparison of selection methods of explanatory variables in PLS regression with application to manufacturing process data. *Chemometrics and Intelligent Laboratory Systems* 58: 171–193.
- Gauthier J, Krause TW & Atherton DL (1998) Measurement of residual stress in steel using the magnetic Barkhausen technique. *NDT&E International* 31: 23–31.
- Goldberg DE (1989) *Genetic algorithms in search, optimization, and machine learning*. Reading, Addison-Wesley.
- Govindaraju MR, Strom A, Jiles DC, Biner SB & Chen Z-J (1993) Evaluation of fatigue damage in steel structural components by magnetoelastic Barkhausen signal analysis. *Journal of Applied Physics* 73: 6165–6167.
- Grefenstette JJ (1986) Optimization of control parameters for genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics* 16: 122–128.
- Gualdrón O, Llobet E, Brezmes J, Vilanova X & Correig X (2006) Coupling fast variable selection methods to neural network-based classifiers: Application to multisensor systems. *Sensors and Actuators B: Chemical* 114: 522–529.
- Gupta H, Zhang M & Parakka AP (1997) Barkhausen effect in ground steels. *Acta Materialia* 45: 1917–1921.
- Guyon I & Elisseeff A (2003) An introduction to variable and feature selection. *Journal of Machine Learning Research* 3: 1157–1182.
- Hall M (1999) *Correlation-based feature selection for machine learning*. PhD Thesis, Department of Computer Science, Waikato University, New Zealand.
- Harrell Jr FE (2001) *Regression modelling Strategies with applications to linear models, logistic regression, and survival analysis*. New York, Springer.
- Haykin S. (1999) *Neural Networks: A Comprehensive Foundation*. Upper Saddle River (NJ), Prentice Hall.
- Hawkins DM (2004) The problem of overfitting. *Journal of Chemical Information and Computer Sciences* 44: 1–12.

- Hemmateenejad B (2005) Correlation ranking procedure for factor selection in PC-ANN modeling and application to ADMETox evaluation. *Chemometrics and Intelligent Laboratory Systems* 75:231–245.
- Herrera F, Lozano M & Verdegay JL (1998) Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis. *Artificial Intelligence Review* 12: 265–319.
- Hwang DG & Kim HC (1988) The influence of plastic deformation on Barkhausen effects and magnetic properties in mild steel. *Journal of Physics D: Applied Physics* 21: 1807–1813.
- Hörchner U & Kalivas JH (1995) Further investigation on a comparative study of simulated annealing and genetic algorithm for wavelength selection. *Analytica Chimica Acta* 311: 1–13.
- Höskuldsson A (2001) Variable and subset selection in PLS regression. *Chemometrics and Intelligent Laboratory Systems* 55: 23–38.
- Iordache VE, Hug E & Buiron N (2003) Magnetic behaviour versus tensile deformation mechanisms in a non-oriented Fe–(3 wt.%) Si steel. *Materials Science and Engineering A* 359: 62–74.
- Jagadish C, Clapham L & Atherton DL (1990) Influence of uniaxial elastic stress on power spectrum and pulse height distribution of surface Barkhausen noise in pipeline steel. *IEEE Transactions on Magnetics* 26: 1160–1163.
- Jain A & Zongker D (1997) Feature selection: evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19: 153–158.
- Jiles DC (2000) Dynamics of domain magnetization and Barkhausen effect. *Czechoslovak Journal of Physics* 50: 893–924.
- Jouan-Rimbaud D, Massart DL & de Noord OE (1996) Random correlation in variable selection for multivariate calibration with a genetic algorithm. *Chemometrics and Intelligent Laboratory Systems* 35: 213–220.
- Juuso EK & Lahdelma S (2010) Cavitation detection in power control of Kaplan water turbines. *Proceedings of the 51st Conference on Simulation and Modelling, SIMS 2011, October 14–15, Oulu, Finland, 7p.*
- Jiles DC (2000) Dynamics of domain magnetization and Barkhausen effect. *Czechoslovak Journal of Physics* 50: 893–924.
- Kaplan M, Gür CH & Erdogan M (2007) Characterization of Dual-Phase Steels Using Magnetic Barkhausen Noise Technique. *Journal of Nondestructive Evaluation* 26: 79–87.
- Katara S, Bhan A, Caruthers JM, Delgass WN & Venkatasubramanian V (2004) A hybrid genetic algorithm for efficient parameter estimation of large kinetic models. *Computers and Chemical Engineering* 28: 2569–2581.
- Khalik MA, Sherif M, Saraya S & Areed F (2007) Parameter identification problem: Real-coded GA approach. *Applied Mathematics and Computation* 187: 1495–1501.

- Kim D-W & Kwon D (2003) Quantification of the Barkhausen noise method for the evaluation of time-dependent degradation. *Journal of Magnetism and Magnetic Materials* 257: 175–183.
- Kim HC, Lee KH, Kim CG & Hwang DG (1992) Barkhausen noise in ferromagnetic metallic glass Fe₄₀Ni₃₈Mo₄B₁₈. *Journal of Applied Physics* 72: 3626–3633.
- Kimura T, Hasegawa K & Funatsu K (1998) GA Strategy for Variable Selection in QSAR Studies: GA-Based Region Selection for CoMFA Modeling. *Journal of Chemical Information and Computer Sciences* 38: 276–282.
- Kinser ER, Lo CCH, Barsic AJ & Jiles DC (2005) Modeling microstructural effects on Barkhausen emission in surface-modified magnetic materials. *IEEE Transactions on Magnetics* 41: 3292–3294.
- Kleber X & Vincent A (2004) On the role of residual internal stresses and dislocations on Barkhausen noise in plastically deformed steel. *NDT & E International* 37: 439–445.
- Kleber X, Hug-Amalric A & Merlin J (2008) Evaluation of the Proportion of Phases and Mechanical Strength of Two-Phase Steels Using Barkhausen Noise Measurements: Application to Commercial Dual-Phase Steel. *Metallurgical and Materials Transactions A* 39: 1308–1318.
- Kohavi R & John GH (1997) Wrappers for feature subset selection. *Artificial Intelligence* 97: 273–324.
- Koo KM, Yau MY, Ng DHL & Lo CCH (2003) Characterization of pearlite grains in plain carbon steel by Barkhausen emission. *Materials Science and Engineering A* 351: 310–315.
- Kudo M & Sklansky J (2000) Comparison of algorithms that select features for pattern classifiers. *Pattern Recognition* 33: 25–41.
- Learidi R & González AL (1998) Genetic algorithms applied to feature selection in PLS regression: how and when to use them. *Chemometrics and Intelligent Laboratory Systems* 41: 195–207.
- Lei Y, He Z, Zi Y & Hu Q (2008) Fault diagnosis of rotating machinery based on a new hybrid clustering algorithm. *The International Journal of Advanced Manufacturing Technology* 35: 968–977.
- Li B, Morris J & Martin EB (2002) Model selection for partial least squares regression. *Chemometrics and Intelligent Laboratory Systems* 64: 79–89.
- Lindgren M (2003) Barkhausen noise: application to residual stress measurements and fatigue damage estimation. Doctoral Thesis. Tampere University of Technology, Department of Materials Science.
- Lindgren M & Lepistö T (2000) Application of a novel type Barkhausen noise sensor to continuous fatigue monitoring. *NDT&E International* 33: 423–428.
- Lindgren M & Lepistö T (2001) Effect of prestraining on Barkhausen noise vs. stress relation. *NDT&E International* 34: 337–344.
- Lindgren M & Lepistö T (2002) Application of Barkhausen noise to biaxial residual stress measurements in welded steel tubes. *Materials Science and Technology* 18: 1369–1376.

- Lindgren M & Lepistö T (2003) Effect of cyclic deformation on Barkhausen noise in a mild steel. *NDT&E International* 36: 401–409.
- Llobet E, Brezmes J, Gualdrón O, Vilanova X & Correig X (2004) Building parsimonious fuzzy ARTMAP models by variable selection with a cascaded genetic algorithm: application to multisensor systems for gas analysis. *Sensors and Actuators B* 99: 267–272.
- Llobet E, Gualdrón O, Vinaixa M, El-Barbri N, Brezmes J, Vilanova X, Bouchikhi B, Gómez R, Carrasco JA & Correig X (2007) Efficient feature selection for mass spectrometry based electronic nose applications. *Chemometrics and Intelligent Laboratory Systems* 85: 253–261.
- Lo CCH, Tang F, Biner SB & Jiles DC (2000) Effects of fatigue-induced changes in microstructure and stress on domain structure and magnetic properties of Fe–C alloys. *Journal of Applied Physics* 87: 6520–6522.
- Maaranen H, Miettinen K and Mäkelä MM (2004) Quasi-random initial population for genetic algorithms. *Computers and Mathematics with Applications* 47: 1885–1895.
- Malkin S (1989) *Grinding technology: Theory and applications of machining with abrasives*. Chichester Horwood, American Society of Civil Engineers.
- McGonnagle WJ (1963) *Nondestructive testing*. New York, McGraw-Hill Book Company, Inc.
- McLeod G, Clelland K, Tapp H, Kemsley EK, Wilson RH, Poulter G, Coombs D & Hewitt CJ (2009) A comparison of variate pre-selection methods for use in partial least squares regression: A case study on NIR spectroscopy applied to monitoring beer fermentation. *Journal of Food Engineering* 90: 300–307.
- McMaster RC (1963) *Nondestructive testing handbook, Volume*. New York, The Ronald Press Company.
- Meiri R & Zahavi J (2006) Using simulated annealing to optimize the feature selection problem in marketing applications. *European Journal of Operational Research* 171: 842–858.
- Mészáros I & Szabó PJ (2005) Complex magnetic and microstructural investigation of duplex stainless steel. *NDT & E International* 38: 517–521.
- Michalewicz Z (1996) *Genetic algorithms + data structures = evolution programs*, 3rd edition. Berlin, Springer.
- Mierczak L, Jiles DC & Fantoni G (2011) A New Method for Evaluation of Mechanical Stress Using the Reciprocal Amplitude of Magnetic Barkhausen Noise. *IEEE Transactions on Magnetics* 47: 459–465.
- Moorthy V, Choudhary BK, Vaidyanathan S, Jayakumar T, Rao KBS & Raj B (1999a) An assessment of low cycle fatigue damage using magnetic Barkhausen emission in 9Cr–1Mo ferritic steel. *International Journal of Fatigue* 21: 263–269.
- Moorthy V & Shaw BA (2009) Magnetic Barkhausen emission measurements for evaluation of depth of grinding damage. In: *Proceedings of 7th International Conference on Barkhausen noise and micromagnetic testing*, July 15–16, 2009, Aachen, Germany, 21–32.

- Moorthy V, Shaw BA & Brimble K (2004a) Testing of Case Depth in Case carburized Gear steels using Magnetic Barkhausen Emission technique. *Material Evaluation* 62: 523–527.
- Moorthy V, Shaw BA & Day S (2004b) Evaluation of applied and residual stresses in case-carburised En36 steel subjected to bending using the magnetic Barkhausen emission technique. *Acta Materialia* 52: 1927–1936.
- Moorthy V, Shaw BA & Evans JT (2003) Evaluation of tempering induced changes in the hardness profile of case-carburised EN36 steel using magnetic Barkhausen noise analysis. *NDT&E International* 36: 43–49.
- Moorthy V, Shaw BA, Mountford P & Hopkins P (2005) Magnetic Barkhausen emission technique for evaluation of residual stress alteration by grinding in case-carburised En36 steel. *Acta Materialia* 53: 4997–5006.
- Moorthy V, Vaidyanathan S, Jayakumar T & Raj B (1997a) Microstructural characterization of quenched and tempered 0.2% carbon steel using magnetic Barkhausen noise analysis. *Journal of Magnetism and Magnetic Materials* 171: 179–189.
- Moorthy V, Vaidyanathan S, Jayakumar T & Raj B (1998) Microstructural characterization of quenched and tempered 0.2% carbon steel using magnetic Barkhausen noise analysis. *Journal of Magnetism and Magnetic Materials* 171: 179–189.
- Moorthy V, Vaidyanathan S, Jayakumar T, Raj B & Kashyap BP (1999b) Effect of tensile deformation on micromagnetic parameters in 0.2% carbon steel and 2.25Cr–1Mo steel. *Acta Materialia* 47: 1869–1878.
- Moorthy V, Vaidyanathan S, Laha K, Jayakumar T, Rao KBS & Raj B (1997b) Evaluation of microstructures in 2.25Cr-1Mo and 9Cr-1Mo steel weldments using magnetic Barkhausen noise. *Materials Science and Engineering A* 231: 98–104.
- Moorthy V, Vaidyanathan S, Raj B, Jayakumar T & Kashyap BP (2000) Insight into the microstructural characterization of ferritic steels using micromagnetic parameters. *Metallurgical and Materials Transactions A* 31: 1053–1065.
- Nakariyakul S & Casasent DP (2007) Adaptive branch and bound algorithm for selecting optimal features. *Pattern Recognition Letters* 28: 1415–1427.
- Nakariyakul S & Casasent DP (2009) An improvement on floating search algorithms for feature subset selection. *Pattern Recognition* 42: 1932–1940.
- Ng DHL, Cheng K, Cho KS, Ren ZY, Ma XY & Chan SLI (2001) Nondestructive Evaluation of Carbon Contents and Microstructures in Plain Carbon Steel Bars by Barkhausen Emission. *IEEE Transactions on Magnetics* 37: 2734–2736.
- Nunez-Letamendia L (2007) Fitting the control parameters of a genetic algorithm: An application to technical trading system design. *European Journal of Operational Research* 179: 847–868.
- Ohenoja M & Leiviskä K (2010) Validation of genetic algorithm results in a fuel cell model. *International Journal of Hydrogen Energy* 35: 12618–12625.
- O'Sullivan D, Cotterell M, Tanner DA & Mészáros I (2004) Characterisation of ferritic stainless steel by Barkhausen techniques. *NDT & E International* 37: 489–496.

- Palit Sagar S, Parida N, Das S, Dobmann G & Bhattacharya DK (2005) Magnetic Barkhausen emission to evaluate fatigue damage in a low carbon structural steel. *International Journal of Fatigue* 27: 317–322.
- Palma ES, Mansur TR, Ferreira Silva Jr S & Alvarenga Jr A (2005) Fatigue damage assessment in AISI 8620 steel using Barkhausen noise. *International Journal of Fatigue* 27: 659–665.
- Paulsson N, Larsson E & Winquist F (2000) Extraction and selection of parameters for evaluation of breath alcohol measurement with an electronic nose. *Sensors and Actuators A: Physical* 84: 187–197.
- Peng H, Long F & Ding C (2005) Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27: 1226–1238.
- Pérez-Benitez JA, Capó-Sánchez J, Anglada-Rivera J & Padovese LR (2005) A model for the influence of microstructural defects on magnetic Barkhausen noise in plain steels. *Journal of Magnetism and Magnetic Materials* 288: 433–442.
- Pham DT & Karaboga D (2000) *Intelligent Optimisation Techniques: Genetic Algorithms, Tabu Search, Simulated Annealing and Neural Networks*. London, Springer.
- Piotrowski L, Augustyniak B, Chmielewski M, Hristoforou EV & Kosmas K (2010) Evaluation of Barkhausen noise and magnetoacoustic emission signals properties for plastically deformed Armco iron. *IEEE Transactions on Magnetics* 46: 239–242.
- Pires JCM, Martins FG, Sousa SIV, Alvim-Ferraz MCM & Pereira MC (2008) Selection and validation of parameters in multiple linear and principal component regressions. *Environmental Modelling & Software* 23: 50–55.
- Pudil P, Novovičová J & Kittler J (1994) Floating search methods in feature selection. *Pattern Recognition Letters* 15: 1119–1125.
- Rahnamayan S, Tizhoosh HR & Salama MMA (2007) A novel population initialization method for accelerating evolutionary algorithms. *Computers and Mathematics with Applications* 53: 1605–1614.
- Ramuhalli P, Udpa L & Udpa SS (2002) Electromagnetic NDE signal inversion by function-approximation neural networks. *IEEE Transactions on Magnetics* 38: 3633–3642.
- Ranjan R, Jiles D & Rastogi P (1987) Magnetic properties of decarburized steels: An investigation of the effects of grain size and carbon content. *IEEE Transactions on Magnetics* 23: 1869–1876.
- Roffel B & Betlem B (2006) *Process dynamics and control*. Chichester, England, John Wiley & Sons Ltd.
- Romagnoli JA & Palazoglu A (2006) *Introduction to process control*. Boca Raton (FL), Taylor & Francis.
- Santa-aho S (2008) Characterization of gear grinding burns by multiparameter noise measurements. M.Sc. Thesis. Tampere University of Technology, Department of Materials Science.

- Santa-aho S, Vippola M, Lepistö T & Lindgren M (2009) Characterisation of case-hardened gear steel by multiparameter Barkhausen noise measurements. *Insight - Non-Destructive Testing and Condition Monitoring* 51: 212–216.
- Santa-aho S, Hakalahti J, Vippola M, Lepistö T, Lindgren M, Hakanen M (2010) The use of Barkhausen noise to detect the case-depth of induction hardened samples. In: *Proceedings of 8th International Conference on Barkhausen noise and micromagnetic testing*, February 11–12, Kalpakkam, India, 12p.
- Santa-aho S, Vippola M, Lindgren M, Sorsa A, Leiviskä K & Lepistö T (2011) Use of low frequency Barkhausen noise for inspection of induction hardened rod samples. In: *Proceedings of 9th International conference on Barkhausen noise and micromagnetic testing*, June 28–29, Hejnice, Czech Republic, 12p.
- Santa-aho S, Vippola M, Saarinen T, Isakov M, Sorsa A, Lindgren M, Leiviskä K & Lepistö T (2012a) Barkhausen Noise characterization during elastic bending and tensile-compression loading of case-hardened and tempered samples. *Journal of Materials Science* 47: 6520–6428.
- Santa-aho S, Vippola M, Sorsa A, Latokartano J, Lindgren M, Leiviskä K & Lepistö T (2012b) Development of Barkhausen noise calibration blocks for reliable grinding burn detection. *Journal of Materials Processing Technology* 212: 408–416.
- Santa-aho S, Vippola M, Sorsa A, Leiviskä K, Lindgren M & Lepistö T (2012c) Utilization of Barkhausen noise magnetizing sweeps for case-depth detection from hardened steel. *NDT & E International* XX: XXX-XXX.
- Santa-aho S, Vippola M, Sorsa A, Lindgren M, Latokartano J, Leiviskä K & Lepistö T (2012d) Optimized laser processing of calibration blocks for grinding burn detection with Barkhausen noise. *Journal of Materials Processing Technology* 212: 2282–2293.
- Saquet O, Chicois J & Vincent A (1999) Barkhausen noise from plain carbon steels: analysis of the influence of microstructure. *Materials Science and Engineering A* 269: 73–82.
- Sayes Y, Inza I & Larrañaga P (2007) A review of feature selection techniques in bioinformatics. *Bioinformatics* 23: 2507–2517.
- Schijve J (2001) *Fatigue of structures and materials* (electronic book). Dordrecht, Kluwer Academic. <http://site.ebrary.com/lib/oulu/docDetail.action?docID=10067462>
- Serpico SB & Bruzzone L (2001) A new search algorithm for feature selection in hyperspectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing* 39: 1360–1367.
- Sipahi LB (1994) Effects of creep damage, shot peening, and case hardening on magnetic Barkhausen noise analysis. *IEEE Transactions on Magnetics* 30: 4830–4832.
- Smit S, Hoefsloot HCJ & Smilde AK (2008) Statistical data processing in clinical proteomics. *Journal of Chromatography B* 866: 77–88.
- Somol P, Pudil P, Novovičová J & Pačlíc P (1999) Adaptive floating search methods in feature selection. *Pattern Recognition Letters* 20: 1157–1163.
- Sorsa A & Leiviskä K (2009a) An entropy-based approach for the analysis of the Barkhausen noise signal. In: *Proceedings of 7th International Conference on Barkhausen noise and Micromagnetic Testing*, July 15–16, Aachen, Germany, 12p.

- Sorsa A & Leiviskä K (2009b) Feature selection from Barkhausen noise data using genetic algorithms with cross-validation. In: Kolehmainen M, Toivanen P & Beliczynski B (eds.) ICANNGA 2009, Lecture Notes in Computer Science 5495: 213–222.
- Sorsa A & Leiviskä K (2010) Simultaneous prediction of residual stress and hardness from Barkhausen noise signal. In: Proceedings of 8th International Conference on Barkhausen noise and Micromagnetic Testing, February 11–12, Kalpakkam, India, 12p.
- Sorsa A & Leiviskä K (2011a) Comparison of feature selection methods applied to Barkhausen noise data set. In: Bittanti S, Cenedese A & Zampieri S (eds.) IFAC World Congress, volume 18, 14699–14704.
- Sorsa A & Leiviskä K (2011b) Predicting residual stress based on the Barkhausen noise measurements: a neural network approach. In: Proceedings of 9th International Conference on Barkhausen noise and Micromagnetic Testing, June 28–29, Hejnice, Czech Republic, 12p.
- Sorsa A & Leiviskä K (2011c) Simultaneous prediction of residual stress and hardness from Barkhausen noise signal. *NDT World Review* 4: 78–83. (In Russian)
- Sorsa A, Leiviskä K & Santa-aho S (2008a) Prediction of residual stress from Barkhausen noise signal. In: *NDT 2008*, 15–18 September, Cheshire, UK, 10p.
- Sorsa A, Leiviskä K, Santa-aho S & Lepistö T (2012a) A data-based modelling scheme for estimating residual stress from Barkhausen noise measurements. *Insight - Non-Destructive Testing and Condition Monitoring* 54: 278–283.
- Sorsa A, Leiviskä K, Santa-aho S & Lepistö T (2012b) Quantitative prediction of residual stress and hardness in case-hardened steel based on the Barkhausen noise measurement. *NDT&E International* 46: 100–106.
- Sorsa A, Leiviskä K, Santa-aho S, Vippola M & Lepistö T (2010a) A study on laser processed grinding burn simulation and analysis based on Barkhausen noise measurement. *Insight - Non-Destructive Testing and Condition Monitoring* 52: 293–297.
- Sorsa A, Peltokangas R & Leiviskä K (2008b) Real-coded genetic algorithms and nonlinear parameter identification. *IS'08*, 4th International IEEE Conference Intelligent Systems, volume 2, 10–42 - 10–47.
- Sorsa A, Peltokangas R & Leiviskä K (2010b) Case studies for genetic algorithms in system identification tasks. In: Kacprzyk J, Hadjiski M & Sgurev V (eds.) *Intelligent Systems: From Theory to Practice*. Berlin-Heidelberg, Springer.
- Stefanita C-G, Clapham L & Atherton DL (2000a) Plastic versus elastic deformation effects on magnetic Barkhausen noise in steel. *Acta Materialia* 48: 3545–3551.
- Stefanita C-G, Clapham L & Atherton DL (2000b) Subtle changes in magnetic Barkhausen noise before the macroscopic elastic limit. *Journal of Materials Science* 35: 2675–2681.
- Stewart DM, Stevens KJ & Kaiser AB (2004) Magnetic Barkhausen noise analysis of stress in steel. *Current Applied Physics* 4: 308–311.

- Stupakov O, Perevertov O, Tomáš I & Skrbek B (2011) Evaluation of surface decarburization depth by magnetic Barkhausen noise. *Journal of Magnetism and Magnetic Materials* 323: 1692–1697.
- Sun J (1995) A correlation principal component regression analysis of NIR data. *Journal of Chemometrics* 9: 21–29.
- Sutter JM & Kalivas JH (1993) Comparison of forward selection, backward elimination, and generalized simulated annealing for variable selection. *Microchemical Journal* 47: 60–66.
- Sutter JM, Dixon SL & Jurs PC (1995) Automated descriptor selection for quantitative structure- activity relationships using generalized simulated annealing. *Journal of Chemical Information and Computer Sciences* 35: 77–84.
- Swierenga H, de Groot PJ, de Weijer AP, Derksen MWJ & Buydens LMC (1998) Improvement of PLS model transferability by robust wavelength selection. *Chemometrics and Intelligent Laboratory Systems* 41: 237–248.
- Tomita Y, Hashimoto K & Osawa N (1996) Nondestructive estimation of fatigue damage for steel by Barkhausen noise analysis. *NDT&E International* 29: 275–280.
- Vincent A, Pasco L, Morin M, Kleber X & Delnondedieu M (2005) Magnetic Barkhausen noise from strain-induced martensite during low cycle fatigue of 304L austenitic stainless steel. *Acta Materialia* 53: 4579–4591.
- Wang G, Feng E & Xiu Z (2008) Modeling and parameter identification of microbial bioconversion in fed-batch cultures. *Journal of Process Control* 18: 458–464.
- Wang Y, Li L, Ni J & Huang S (2009) Feature selection using tabu search with long-term memories and probabilistic neural networks. *Pattern Recognition Letters* 30: 661–670.
- Wang Y & Xiang B (2007) Radial basis function network calibration model for near-infrared spectra in wavelet domain using a genetic algorithm. *Analytica Chimica Acta* 602: 55–65.
- Wasim M & Brereton RG (2004) Determination of the number of significant components in liquid chromatography nuclear magnetic resonance spectroscopy. *Chemometrics and Intelligent Laboratory Systems* 72: 133–151.
- Wessels LFA, Reinders MJT, Hart AAM, Veenman CJ, Dai H, He YD & van't Veer LJ (2005) A protocol for building and evaluating predictors of disease state based on microarray data. *Bioinformatics* 21: 3755–3762.
- Wilson JW, Tian GY, Moorthy V & Shaw BA (2009) Magneto-Acoustic Emission and Magnetic Barkhausen Emission for Case Depth Measurement in En36 Gear Steel. *IEEE Transactions on Magnetics* 45: 177–183.
- Wisnowski JW, Simpson JR, Montgomery DC & Runger GC (2003) Resampling methods for variable selection in robust regression. *Computational Statistics & Data Analysis* 43: 341–355.
- Withers PJ & Bhadeshia HKDH (2001a) Residual stress, Part 1 – Measurement techniques. *Materials Science and Technology* 17: 355–365.
- Withers PJ & Bhadeshia HKDH (2001b) Residual stress, Part 2 – Nature and origins. *Materials Science and Technology* 17: 366–375.

- Wilson JW, Tian GY, Moorthy V & Shaw BA (2009) Magneto-Acoustic Emission and Magnetic Barkhausen Emission for Case Depth Measurement in En36 Gear Steel. *IEEE Transactions on Magnetics* 45: 177–183.
- Xie Y-L & Kalivas JH (1997) Evaluation of principal component selection methods to form a global prediction model by principal component regression. *Analytica Chimica Acta* 348: 19–27.
- Yadav M & Wadhvani S (2011) Vibration analysis of bearing for fault detection using time domain features and neural network. *International Journal of Applied Research in Mechanical Engineering* 1: 69–74.
- Yamaura S, Furuya Y & Watanabe T (2001) The effect of grain boundary microstructure on Barkhausen noise in ferromagnetic materials. *Acta Materialia* 49: 3019–3027.
- Zhang P (1992) On the distributional properties of model selection criteria. *Journal of the American Statistical Association* 87: 732–737.
- Zhang YX (2007) Artificial neural networks based on principal component analysis input selection for clinical pattern recognition analysis. *Talanta* 73: 68–75.

426. Kukka, Hannu (2012) Case studies in human information behaviour in smart urban spaces
427. Koivukangas, Tapani (2012) Methods for determination of the accuracy of surgical guidance devices : A study in the region of neurosurgical interest
428. Landaburu-Aguirre, Junkal (2012) Micellar-enhanced ultrafiltration for the removal of heavy metals from phosphorous-rich wastewaters : From end-of-pipe to clean technology
429. Myllymäki, Sami (2012) Capacitive antenna sensor for user proximity recognition
430. Jansson, Jussi-Pekka (2012) A stabilized multi-channel CMOS time-to-digital converter based on a low frequency reference
431. Soini, Jaakko (2012) Effects of environmental variations in *Escherichia coli* fermentations
432. Wang, Meng (2012) Polymer integrated Young interferometers for label-free biosensing applications
433. Halunen, Kimmo (2012) Hash function security : Cryptanalysis of the Very Smooth Hash and multicollisions in generalised iterated hash functions
434. Destino, Giuseppe (2012) Positioning in Wireless Networks : Non-cooperative and cooperative algorithms
435. Kreku, Jari (2012) Early-phase performance evaluation of computer systems using workload models and SystemC
436. Ossiannilsson, Ebba (2012) Benchmarking e-learning in higher education : Lessons learned from international projects
437. Pouttu, Ari (2012) Performance Analysis of mMCSK-mMFSK modulation variants with comparative discussion
438. Kupiainen, Laura (2012) Dilute acid catalysed hydrolysis of cellulose – extension to formic acid
439. Kurtti, Sami (2012) Integrated receiver channel and timing discrimination circuits for a pulsed time-of-flight laser rangefinder
440. Distanont, Anyanitha (2012) Knowledge transfer in requirements engineering in collaborative product development
441. Keskinarkaus, Anja (2012) Digital watermarking techniques for printed images

S E R I E S E D I T O R S

A
SCIENTIAE RERUM NATURALIUM

Senior Assistant Jorma Arhippainen

B
HUMANIORA

University Lecturer Santeri Palviainen

C
TECHNICA

Professor Hannu Heusala

D
MEDICA

Professor Olli Vuolteenaho

E
SCIENTIAE RERUM SOCIALIUM

University Lecturer Hannu Heikkinen

F
SCRIPTA ACADEMICA

Director Sinikka Eskelinen

G
OECONOMICA

Professor Jari Juga

EDITOR IN CHIEF

Professor Olli Vuolteenaho

PUBLICATIONS EDITOR

Publications Editor Kirsti Nurkkala

ISBN 978-952-62-0067-5 (Paperback)

ISBN 978-952-62-0068-2 (PDF)

ISSN 0355-3213 (Print)

ISSN 1796-2226 (Online)

