ISSN 2303-4521

Periodicals of

Engineering and Natural Sciences

PEN – VOL. 1 NO. 2
(2013)

Published by International University of Sarajevo

Periodicals of Engineering and Natural Sciences (PEN)

Periodicals of Engineering and Natural Sciences (ISSN: 2303-4521) is an international open access single-blind review journal published online.

Publication frequency: Semiyearly (1. January - June; 2. July - December).

Publication Fees: No fee required (no article submission charges or processing charges).

Digital Object Identifier DOI: 10.21533/pen

Editorial Office

Mailing Address

Hrasnicka cesta 15
71000 Sarajevo
Bosnia

Principal Contact

Benjamin Durakovic
Managing Editor
International University of Sarajevo
Hrasnicka cesta 15
71000 Sarajevo
Bosnia
Phone: +387 33 957 229
Email: pen@ius.edu.ba

Available online at:

http://pen.ius.edu.ba

Editorial Team

Editor in Chief

Prof. Dr. Fehim Fındık, International University of Sarajevo, Bosnia

Managing Editor

Benjamin Durakovic, International University of Sarajevo, Bosnia

Editorial Board

1. Prof. Dr. Ozer Cinar, IUS, BIO/Environmental Eng., BiH
2. Prof. Dr. Orhan Torkul, IE, Sakarya University, TR
3. Prof. Dr. Mehmet Sabih Aksoy, CS, King Saud University, SA
4. Prof. Dr. İsmail Hakkı Cedimoğlu, IE, Sakarya University, TR
5. Prof. Dr. Mehmet Can, Applied Math, IUS Emeritus Professor, BiH
6. Prof. Dr. İbrahim Özsert, ME, Sakarya University, TR
7. Prof. Dr. Adem Demir, Mater. Sci, Sakarya University, TR
8. Prof. Dr. Hüseyin Uzun, Mater. Sci Sakarya University, TR
9. Prof. Dr. Seong Jin Park, ME, Pohang University of Science and Technology, KR
10. Prof. Dr. Muzaffar A. Shaikh, IE, Florida Institute of Technology, US
11. Prof. Dr. John Dee, ARCH, IUS FENS, AUS
12. Assoc. Prof. Dr. Mustafa Akay, ME, Ulster University, NIR
13. Assoc. Prof. Dr. Ümit Alver, EE, IUS FENS, BiH
14. Assoc. Prof. Dr. Cemil Öz, CS, Sakarya University, TR
15. Assoc. Prof. Dr. Harun Reşit Yazgan, IE, Sakarya University, TR
16. Assoc. Prof. Dr. Raşit Köker, EE, Sakarya University, TR 17.
17. Asst. Prof. Dr. Abdurrazag Ali Aburas, CS, IUS FENS, BiH
18. Asst. Prof. Dr. Muhamed Hadziabdic, ME, IUS, BiH
19. Asst. Prof. Dr. Muris Torlak, ME, University of Sarajevo, BiH
20. Asst. Prof. Dr. Mümtaz İpek, IE, Sakarya University, TR
21. Asst. Prof. Dr. Semra Boran, IE, Sakarya University, TR
22. Asst. Prof. Dr. Jasminka Telalovic, CS, IUS FENS, BiH
23. Asst. Prof. Dr. Edin Jahic, ACH, IUS FENS, BiH
24. Asst. Prof. Dr. Mirza Suljagic, BIO, IUS FENS, BiH
25. Asst. Prof. Dr. Izudin Dzafic, EE, IUS FENS, BiH
26. Asst. Prof. Dr. Kevser Cirik, BIO/Environmental Eng., TR
27. Dr. Youssef Hammi, ME, Mississippi State University, US
28. Dr. Nezih Mred, ME, National Research Council, CA

DOI: 10.21533/pen

Publication Frequency

Expected frequency of publication is twice per year:
1. January - June;
2. July - December.

Journal items can be published collectively, as part of an issue with its own Table of Contents. Alternatively, individual items can be published as soon as they are ready, by adding them to the "current" volume's Table of Contents.

Open Access Policy

This journal provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge.

Journal Ethics and Malpractice Statement

PEN Journal is committed to ensure and uphold standards of ethical behavior at all stages of the publication process, whereby such values also rely on editors, reviewers, contributors and authors who are expected to behave ethically. The standards are based on Committee on Publication Ethics' (COPE) code of conduct, and provide guidelines for best practices in order to meet these requirements. The following ethical guidelines are only intended to give a summary of our key expectations of editors, peer-reviewers, and authors but if you have any questions or concerns please also feel free to contact the Editor of the Journal.

1. Ethical Expectations

Editors' responsibilities

- To carry out their duties in a fair, objective and consistent manner, without discrimination on grounds of gender, sexual orientation, religious or political beliefs, ethnic or geographical origin of the authors.
- To take care that each article undergoes the proper peer-review process.
- To promote consistent ethical policies of PEN journal, and to ensure the confidentiality of the reviewing process.
- To uphold integrity in their work as editors of the journal, doing away with any personal interest, so that articles are considered and accepted solely on their academic merit and without commercial influence.
- To work with authors, reviewers, and the Editorial Board members to ensure the implementation of journals' ethics and publishing policies.
- To adopt and follow reasonable procedures in the event of complaints of an ethical or conflict nature, in accordance with the policies and procedures. To handle process of complaints and giving authors an opportunity to respond any complaints. All complaints should be investigated and the documentation associated with any such complaints should be retained.

Reviewers' responsibilities

- To act objectively, fairly and in a timely manner in reviewing submitted manuscript with the aim of improving its quality such as pointing out relevant published work, which is not cited, etc.

DOI: 10.21533/pen

- To alert PEN in case of any competing interest that could affect the impartiality of their reviewing, or any potential conflict of interest that includes any relationship between reviewer and author, or any content that is substantially similar to that under review.
- To conduct themselves fairly and impartially.
- To keep the confidentiality of the review process and to not retain or copy the manuscript.

Authors' responsibilities

- To ensure that their work submitted to the journal is original and authored by them and has not been previously published nor under consideration or accepted for publication elsewhere.
- To ensure that original ideas, data, findings and materials taken from other sources (including their own published writing) are properly documented and cited. Any content reproduced from other sources author should have permission.
- To ensure that their data is their won, true and not manipulated. Thus, authors are responsible to maintain accurate records and to provide access to their data associated with their manuscript.
- To ensure their work does not violate any rights of others, including privacy rights and intellectual property rights. In addition, authors should ensure that any studies involving human or animal subject are in accordance with local lows and requirements.
- To declare any real or apparent conflicting or competing interest at any stage during the publication process that could be considered or viewed as exerting an undue influence on his/her duties.
- To alert PEN in case if a significant error in their publication is identified and correct any errors prior or subsequent to publication of their work.
- To adhere to all research ethics guidelines of their discipline, and to ensure that authorship and/or co-authorship of the paper was accurately represented.

PEN responsibilities

- The journal and the publisher shall ensure that good practice is maintained to the standards outlined above.
- To deal with research misconduct allegations appropriately when it is occurred. The journal and the publisher shall undertake reasonable actions to identify and prevent the publication of papers where research misconduct has occurred.
- To maintain the editorial independence of journal editors.
- To support journal editors in ethical and academic matters.
- To ensure critical and objective assessment of all articles by reviewers and referees. To keep an accurate and transparent record, including publishing corrections and retractions when necessary.
- The following commonly recognized procedure for dealing with unethical behavior is adopted.

2. Procedures for Dealing With Unethical Behavior

Identification of unethical behavior

- Misconduct and unethical behavior may be identified and brought to the attention of the editor and publisher at any time, by anyone.
- Misconduct and unethical behavior may include, but need not be limited to, examples as outlined above.
- For an investigation sufficient information and evidence should be provided. All allegations should be taken seriously and treated in the same way, until a successful decision or conclusion is reached.

Investigation

- An initial decision should be taken by the editor, who should consult with or seek advice from the publisher, if appropriate.

DOI: 10.21533/pen

- Evidence should be gathered, while avoiding spreading any allegations beyond those who need to know.

Minor breaches

- Minor misconduct might be dealt with without the need to consult more widely. In any event, the author should be given the opportunity to respond to any allegations.

Serious breaches

- Serious misconduct might require that the employers of the accused be notified. The editor, in consultation with the publisher should make the decision whether or not to involve the employers, either by examining the available evidence themselves or by further consultation with a limited number of experts.

Outcomes (in increasing order of severity; may be applied separately or in conjunction)

- Informing or educating the author or reviewer where there appears to be a misunderstanding or misapplication of acceptable standards.
- A more strongly worded letter to the author or reviewer covering the misconduct and as a warning to future behavior.
- Publication of a formal notice detailing the misconduct.
- Publication of an editorial detailing the misconduct.
- A formal letter to the head of the author's or reviewer's department or funding agency.
- Formal retraction or withdrawal of a publication from the journal, in conjunction with informing the head of the author or reviewer's department, Abstracting & Indexing services and the readership of the publication.
- Imposition of a formal embargo on contributions from an individual for a defined period.
- Reporting the case and outcome to a professional organization or higher authority for further investigation and action.

Originality and Plagiarism Policy

Authors by submitting their manuscript to PEN declare that their work is original and authored by them, their work submitted to PEN has not been previously published, original ideas, data, findings and materials taken from other sources (including their own published writing) are properly documented and cited, their work does not violate any rights of others, including privacy rights and intellectual property rights, their data is their won, true and not manipulated. Plagiarism in whole or in part without proper citation is not tolerated by the Journal. Manuscripts submitted to the journal will be checked for originality using anti-plagiarism software.

DOI: 10.21533/pen

ISSN 2303-4521
PERIODICALS OF ENGINEERING AND NATURAL SCIENCES
Vol. 1 No. 2 (2013)
Available online at: http://pen.ius.edu.ba

Data Mining in Higher Education

Dina A.Aziz AlHammadi
Department of Information Sciences
College of Computer and Information Sciences
King Saudi University P.O.Box 2454, Riyadh 11451
Saudi Arabia
dina.alhammadi@gmail.com

Mehmet Sabih Aksoy
Department of Information Sciences
College of Computer and Information Sciences
King Saudi University P.O.Box 2454, Riyadh 11451
Saudi Arabia
msaksoy@ksu.edu.sa

Abstract
Data mining is slowly but surely making its way into the educational field after dominating the business fields. This paper focuses on the research completed in the area of data mining in the higher education sector: colleges and universities. We will look at the different implementation of data mining and to what extent was it utilized and benefited from.

Keywords: data mining, higher education, inductive learning, rules-3

1. Introduction

Data mining is now the tool to advance every education establishment in every country. And now many nations are using data mining to upgrade their education, and thus producing highly educated citizens who can elevate the other sectors, and upheave the economy and value of the nations.

Data mining is now considered a great asset to the educational sector in highlighting and scrutinizing many elements absent from educational managers. An experimental study was presented in [1]. Our scope is higher education, and the managers can be any professor, instructor, academic counselor, teacher, head, teaching assistant, etc. Providing certain data to these managers can help in early identification of students who are falling behind their peers, guiding students through the best college curricula path, deciding majors, identifying pass and fail rations, locating education related issues and the ability to sole these problems early on in the learning process.

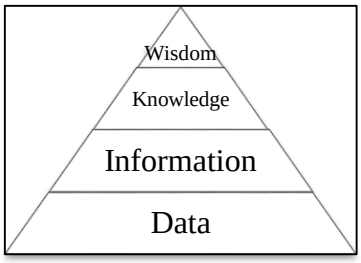


Figure 1. DIKW Pyramid

In Figure 1, Data is unprocessed and raw facts that hold no meaning, as we move to the next layer, the data is now information, this means that data has undergone some processing to produce meaningful information, the next layer is mining the information for knowledge, and moving to the top layer means we have managed to acquire wisdom. The main text of this study consists of three sections. The second section is the work covered in the area of educational data mining. The third section is the conclusion.

2. Related Work

In 2007, Vranice et al., in [2] explored data mining algorithms on a Croatian university students' data. The focus was whether future students of this course will succeed or fail. They have tested several algorithm and their result was similar, however, the authors indicated that their sample was small and perhaps future research would include more detailed student data.

Parack et al., in [3] presented a paper focusing on predicting academic trends and student patterns behavior. This has eased the process in grouping similar student profiles and identifying their learning patterns.

Very interesting paper by Nasiri and Minaei in [4] in higher education focused on two issues: GPA and academic dismissal in a Learning Management System (LMS). Both algorithms used for the data mining process indicated a weakness if there is a slight variation to the data, this will lead to different results, they solved it by adding association rules.

Shi et al., presented a paper [5] focusing on managing the university curricula based on data association mining technology. They reached the conclusion that if a student was successful in a certain course, then he will be

DOI: 10.21533/pen.v1i2.17

successful in similar courses as well. An example can be mathematics and physics.

Again in India, Bunkar et al., presented a paper where they applied data mining techniques to predict the performance improvement of graduate students using classification [6]. Several techniques were discussed and the authors were able to isolate students that are most likely to fail and provide proper counseling and guidance.

In Romania, Bresfelean et al., exploited the university academic failure issue in [7]. Their aim was to define an academic failure profile for students to be able to predict students' exam failure and success based on data mining techniques. They aimed to improve students learning methods and detect their weakness, and assist in managerial educational decision.

Offering high quality education means being able to predict student enrollment in courses, identify beneficial teaching methods, forecast student performance in end exams and identify drop out rates, and help those students during the semester. Their method described in [8] based on classification helped in the proper dividing of students, and paying special attention to students most likely to fail, and help in increasing the success in the success and failure ration.

In China, Wu had a different intake on higher education by using clustering to identify student course selection based on teachers [9]. Their goal was based on guiding the students and giving them most appropriate advice to succeed. The author identifies different categories of student-teacher selection, out of the three clusters, one was successful in selecting the teacher based on several important factors which has increased their class interactivity, discipline, behavior and led to their success. The other two groups, should receive proper guidance to achieve what the first group has achieved.

Japan had another take on the topic by focusing on the university curricula and built an EDM to reach an optimal learning success in terms of best possible Grade Point Average (GPA) [10]. They enforced their system by including an individual learner profile that includes pre-university educational data and grows as the student progresses through the university. This helps in grouping similar profiles and inferring success patterns that can't be identified through conventional student analyzing methods.

A recent study in 2012 in [11] focused on predicting the drop-out rate of students from universities, colleges, and institutions in developing countries. Mustafa et al., used classification and regression to classify successful and unsuccessful students based on gender, financial condition, ethnicity, work-status, disability, and study environment. The search was based on background information. And they identified the most important classification factors were: financial support, age group, and gender.

In 2005, a paper presented by Delavari et al., developed an educational data mining model to be tested in a Malaysian University [12]. They had three targets, the first was understanding the course enrollment pattern in a course, and identifying which students are successful in passing the course, and who will fail. This leads to their next target, faculty, who take proper action in guiding these students, and choose direct or indirect methods to provide the necessary class skills, and financial aid. The final target of the system, are the decision makers, the system enhances the education quality and provide quality management, improve policy making, and setting new strategies and goals.

More on the topic of student failure in college courses can be found in [13]. The author used association rules mining algorithm to find the factors that lead to student course failures.

In Spain, a research presented by Tovar and Soto to improve their predicting model [14]. Instructors can locate students having problems with the course and help them. It also identifies student who have the capabilities to pass the course but fail. Their research is based more on statistics than data mining techniques.

A paper by Knauf et al., [15] presented their storyboard model which students at the Tokyo Denki University were using for progressing through the curriculum. Then they used data mining techniques to build an individual student profile to record their personal properties, talents, weaknesses, and preferences. This model presents students with suggested courses where they will be successful.

Ningning presented a paper focusing on data warehousing and data mining [16]. His model identifies students most likely to drop out a course. This can aid business managers in pinpointing students in need of help and guidance.

3. Case Study

An important question to ask higher education institutes is "How much do you know about your students?" And if the answer is the number of enrolled students and percentages of those who passed and failed, then this the wring answer. The right answer would be acknowledging the student's choices in curriculum and college life. This includes finance, achievements, success, and tailored guidance by their academic counselors.

Higher education institutes are in great need of embedding data mining tools as they are essential for progressing academically.

We will explore three different cases where data mining is essential and we will show its benefits.

3.1. Case 1: Students' Level

Students are the basic building block of any educational institute. In any higher education institute students are somewhat dependent on themselves to chooses their own

courses, professors, make important decisions such as finances, dropping from the university, changing majors, and transfers.

Sometimes, students choose courses based on their peers, past experiences with certain professors, and other related or directly unrelated issues. This has lead to the fact that students make wrong or missed choices.

The suggested solution would be incorporating data mining in the educational gate which provides many services to students at King Saud University such as but not limited to: course registration, dropping course, appealing for postponing a semester, and finance.

If based on students' grades, past courses passing or failing, personal background, we may find a pattern where other students were successful, and thus recommend courses and decisions that are likely to help the students achieve much more than he already gaining from his university education. Students will be able to make decisions based on reliable information.

3.2. Academic Guidance Level

Going up the ladder, academic counselors need certain background information to help aid their students. Such as those who are failing courses, or barely passing, constantly transferring between majors, amongst many other issues. If counselors have social, financial and academic background knowledge about the students then they can make a well-informed guiding decision supported by similar past students experiences. Data mining, if incorporated, in the educational gate system of King Saud University will benefit academic counselor by providing essential factors affecting student's success or failure, and will be able to suggest best academic and financial, social paths based on similar patterns by previous students.

3.3. Decision Makers Level

The top of the pyramid is as important in any educational institute as the previous building blocks if not more essential for the success of any higher education institute. Deans, rectors, vice rectors, deanships, college deans, and administrative departments, are all required to make important decisions at any given time. Therefore, information must be delivered to them accurately and in a timely manner. Until recently, decisions are made based on past experiences, however, the institutions can't wait for every decision maker to gain experiences of 10 years or more before having the ability to make the right judgment or gain the ability to assess the institutes capabilities and opportunities. This leads us to incorporating data mining as tools providing not only simple enrollment, transferring or dropping students' data as well as courses and professors. The tool should be able to assess and present the leaders with related data, important factors leading to success or failure regarding students, courses, and professors. Leaders will have an eagle-eye view of the institute, highlighting problem areas, and significant success areas as well. This will be based on past data of all important building blocks and this will help suffice for the experience young leaders are lacking, as identified by the management.

4. Discussion

As portrayed in the three cases, we are not looking at a pyramid, but a cycle where elements extracted from each phase affects the out come of the next phase.

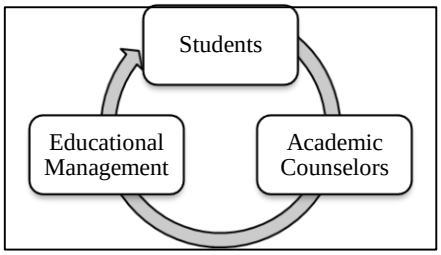


Figure 2: Proposed Data Mining Effect Cycle in Education

Figure 2 clearly explains it. This can be described simply with an example. If we have a student named Ahmad, he is a first year students at the College of Computer and Information Sciences at King Saud University. His first step is registering courses like his peers. The Edu Gate should display to him the required core and elective courses, however, based on his Prepatory Year Program (PYP) courses and other factors where selected patterns are found between him and previous first years students, courses where his success rate is estimated to be high will be suggested to him. If Ahmad registers these suggested courses, and as the system predicted, passes all courses except one. Then we have fed into the system new facts about Ahmed and perhaps a new pattern was introduced. The next year, Ahmad may or may not visit his academic counselor, still, all counselors will have certain facts about troubled students who didn't fair well the previous semester presented in the Edu Gate, alerting the counselor to take notice and act upon the data presented. He may also feed into the system new facts about Ahmed, his courses, or his professors. And because Ahmed failed one course, the system will alarm the counselor to take action based on his PYP data, and passed courses, amongst other important factors highlighted by data mining. And any decision made by the counselor will affect future academic choices presented and decided for Ahmed.

The third phase is laying the most important information to management. Looking more at the big picture rather than unrelated details. Data mining tool in this case will represent problem areas or areas requiring attention, as well as areas that are improving since last assessment. The college dean will notice that there is a course with a high percentage of failure regardless of which students failed. The system will present similar patterns shared by those students that may be a factor. For example; if they are all

taking a course by a single professor, or they have similar weakness area such as: low math grades in PYP, financial issues, or social related issues. Management will be able to focus on those problematic areas and solve the issues instead of making unknowledgeable decisions or neglect those problematical areas. This way the education management will focus their time and effort to improve the institute from all perspectives: students, courses, and faculty. The management will make several changes to all levels, thus feeding back to the students. This is why it should be considered a cycle instead of a pyramid.

5. Conclusions

The aim of this paper is to discuss the need for data mining in all levels of any educational institute. An effective data mining cycle was presented to clarify the process. We will take this research further by incorporating RULES-3 developed by Pham and Aksoy in [17] and later improved by Mathkour in [18], in the Edu Gate System and develop a tool for management to make their insightful decisions. Moreover, we will also focus on data selection, and preparation phases when the system is tested for accuracy, validation, and verification.

Acknowledgements

The authors would like to thank Research Center in the College of Computer and Information Sciences, King Saud University for its support to complete this study.

References

[1] D. A. Alhammadi and M. S. Aksoy, "Data Mining in Education – An Experimental Study," *International Journal of Computer Applications*, vol. 62, no. 15, pp. 31-34, 2013.

[2] M. Vranic, D. Pintar and Z. Skocir, "The use of data mining in education environment," IEEE *Telecommunications ConTel 2007. 9th International Conference*, pp. 243-250, 2007.

[3] S. Parack, Z. Zahid and F. Merchant, "Applications of Data Mining in Educational Databases for Predicting Academic Trends and Patterns," *IEEE International Conference on Technology Enhanced Education (ICTEE)*, pp. 1-4, 2012.

[4] M. Nasiri and B. Minaei, "Predicting GPA and Academic Dismissal in LMS sing Educational Data Mining: A Case Mining, " IEEE *Third International Conference on E-Learning and E-Teaching (ICELET)*, pp. 53-58, 2012.

[5] F. Shi, Q. Miao and D. Mei, "The Application of Data Association Mining Technology in University Curriculum Management," *IEEE Symposium on Robotics and Applications (ISRA)*, pp. 521-524, 2012.

[6] K. Bunkar, U. K. Singh, B. Pandya and R. Bunkar, "Data Mining: Prediction for Performance Improvement of Graduate Students using Classification," *IEEE Wireless and Optical Communications Networks (WOCN)*, pp. 1-5, 2012.

[7] V. P. Bresfelean, M. Bresfelean, N. Ghisoiu and C.A. Comes, "Determining Students' Academic Failure Profile Founded on Data Mining Methods," *IEEE 30th International Conference on Information Technology Interfaces*, pp. 317-322, 2008.

[8] S. K. Yadav, B.K. Bharadwaj and S. Pal, "Data Mining Applications: A comparative study for Predicting Student's Performance", *International Journal of Innovative Technology and Creative Engineering (IJITCE)*, vol. 1, no. 12, pp. 13-19, 2011.

[9] F. Wu, "Apply Data Mining to Students' Choosing Teachers Under Complete Credit Hour," *IEEE Education Technology and Computer Science (ETCS)*, vol. 1, pp. 606-609, 2010.

[10] Y. Sakurai, K. Takada, S. Tsuruta and R. Knauf, "A Case Study on Using Data Mining for University Curricula," *IEEE 12th International Conference on Advanced Learning Technologies (ICALT)*, pp. 3-4, 2012.

[11] M. N. Mustafa, L. Chowdhury and M. Kamal, "Students Dropout Prediction for Intelligent System from Tertiary Level in Developing Country," *International Conference Informatics, Electronics & Vision (ICIEV)*, pp. 113-118, 2012.

[12] N. Delavari, M. R. Beikzadeh and S. Phon-Amnuaisuk, "Application of Enhanced Analysis Model for Data Mining Processes in Higher Educational System," *IEEE International Conference on Information Technology Based Higher Education and Training (ITHET)*, pp. F4B/1-F4B/6, 2005.

[13] Z. Zhang, "Study and Analysis of Data Mining Technology in College Courses Students Failed," *International Conference on Intelligent Computing and Integrated Systems (ICISS)*, pp. 800-802, 2010.

[14] E. Tovar and O. Soto, "The Use of Competences Assessment to Predict the Performance of First Year Students," *IEEE Frontiers in Education Conference (FIE)*, pp. F3J-1, 2010.

[15] R. Knauf, Y. Sakurai, K. Takada and S. Tsuruta, "Personalizing Learning Processes by Data Mining," *IEEE International Conference on Advanced Learning Technologies (ICALT)*, pp. 488-492, 2010.

[16] G. Ningning, "Proposing Data Warehouse and Data Mining in Teaching Management Research," *International Forum on Information Technology and Applications (IFITA)*, vol. 1, pp. 436-439, 2010.

[17] D. T. Pham and M. S.Aksoy, "RULES: A Simple Rule Extraction System," *Expert Systems with Applications*, vol. *8*, no. 1, pp. 59-65, 1995.

[18] H. Mathkour, "RULES3-EXT: Improvements of RULES3 Induction Algorithm," *Mathematical and Computational Applications*, vol. 15, no. 3, pp. 318-324, 2010.

ISSN 2303-4521

PERIODICALS OF ENGINEERING AND NATURAL SCIENCES
Vol. 1 No. 2 (2013)
Available online at: http://pen.ius.edu.ba

Boosted Networks for the Diagnosis of Cardiovascular Diseases

Mehmet Can

International University of Sarajevo

Faculty of Engineering and Natural Sciences

Hrasnicka Cesta 15, 71000 Sarajevo

Bosnia and Herzegovina

Abstract
A boosting by filtering technique for neural network systems with back propagation together with a majority voting scheme is presented in this paper. Previous research with regards to predict the presence of cardiovascular diseases has shown accuracy rates up to 72.9%. Using a boosting by filtering technique prediction accuracy increased over 80%. The designed neural network system in this article presents a significant increase of robustness and it is shown that by majority voting of the parallel networks, recognition rates reach to > 90 in the V.A. Medical Center, Long Beach and Cleveland Clinic Foundation data set.

Keywords—Machine learning, Parallel neural networks, boosting by filtering, cardiovascular diseases

1. Introduction

Cardiovascular disease, also called heart disease, is a class of diseases that involve the heart or blood vessels (arteries, capillaries and veins).[1]

Cardiovascular disease refers to any disease that affects the cardiovascular system, principally cardiac disease, vascular diseases of the brain and kidney, and peripheral arterial disease.[2] The causes of cardiovascular disease are diverse but atherosclerosis and/or hypertension are the most common. Additionally, with aging come a number of physiological and morphological changes that alter cardiovascular function and lead to subsequently increased risk of cardiovascular disease, even in healthy asymptomatic individuals.[3]

Cardiovascular disease is the leading cause of deaths worldwide, though since the 1970s, cardiovascular mortality rates have declined in many high-income countries. [4] At the same time, cardiovascular deaths and disease have increased at a fast rate in low- and middle-income countries.[5] Although cardiovascular disease usually affects older adults, the antecedents of cardiovascular disease, notably atherosclerosis, begin in early life, making primary prevention efforts necessary from childhood.[6] There is therefore increased emphasis on preventing atherosclerosis by modifying risk factors, such as healthy eating, exercise, and avoidance of smoking.

1.1 Types of cardiovascular diseases

- Coronary heart disease (also ischaemic heart disease or coronary artery disease)
- Cardiomyopathy - diseases of cardiac muscle
- Hypertensive heart disease - diseases of the heart secondary to high blood pressure
- Heart failure
- Cor pulmonale - a failure of the right side of the heart
- Cardiac dysrhythmias - abnormalities of heart rhythm
- Inflammatory heart disease
- Endocarditis – inflammation of the inner layer of the heart, the endocardium. The structures most commonly involved are the heart valves.
- Inflammatory cardiomegaly
- Myocarditis – inflammation of the myocardium, the muscular part of the heart.
- Valvular heart disease
- Cerebrovascular disease - disease of blood vessels that supplies to the brain such as stroke
- Peripheral arterial disease - disease of blood vessels that supplies to the arms and legs
- Congenital heart disease - heart structure malformations existing at birth

DOI: 10.21533/pen.v1i2.18

- Rheumatic heart disease - heart muscles and valves damage due to rheumatic fever caused by streptococcal bacteria infections

1.2 Risk factors

Epidemiology suggests a number of risk factors for heart disease: age, gender, high blood pressure, high serum cholesterol levels, tobacco smoking, excessive alcohol consumption, family history, obesity, lack of physical activity, psychosocial factors, diabetes mellitus, and air pollution. While the individual contribution of each risk factor varies between different communities or ethnic groups the consistency of the overall contribution of these risk factors to epidemiological studies is remarkably strong. Some of these risk factors, such as age, gender or family history, are immutable; however, many important cardiovascular risk factors are modifiable by lifestyle change, drug treatment or social change.

Age: Age is an important risk factor in developing cardiovascular diseases. It is estimated that 87 percent of people who die of coronary heart disease are 60 and older. At the same time, the risk of stroke doubles every decade after age 55. Multiple explanations have been proposed to explain why age increases the risk of cardiovascular diseases. One of them is related to serum cholesterol level [7]. In most populations, the serum total cholesterol level increases as age increases. In men, this increase levels off around age 45 to 50 years. In women, the increase continues sharply until age 60 to 65 years. Aging is also associated with changes in the mechanical and structural properties of the vascular wall, which leads to the loss of arterial elasticity and reduced arterial compliance and may subsequently lead to coronary artery disease.

Sex: Men are at greater risk of heart disease than pre-menopausal women.[8] However, once past menopause, a woman's risk is similar to a man's. Among middle-aged people, coronary heart disease is 2 to 5 times more common in men than in women. In a study done by the World Health Organization, sex contributes to approximately 40% of the variation in the sex ratios of coronary heart disease mortality. Another study reports similar results that gender difference explains nearly half of the risk associated with cardiovascular diseases. One of the proposed explanations for the gender difference in cardiovascular disease is hormonal difference. Among women, estrogen is the predominant sex hormone. Estrogen may have protective effects through glucose metabolism and hemostatic system, and it may have a direct effect on improving endothelial cell function. The production of estrogen decreases after menopause, and may change the female lipid metabolism toward a more atherogenic form by decreasing the HDL cholesterol level and by increasing LDL and total cholesterol levels. Women who have experienced early menopause, either naturally or because they have had a hysterectomy, are twice as likely to develop heart disease as women of the same age group who have not yet gone through menopause.

Among men and women, there are differences in body weight, height, body fat distribution, heart rate, stroke volume, and arterial compliance. In the very elderly, age related large artery pulsatility and stiffness is more pronounced in women. This may be caused by the smaller body size and arterial dimensions independent of menopause [9].

Air pollution: Particulate matter has been studied for its short- and long-term exposure effects on cardiovascular disease. Currently, PM2.5 is the major focus, in which gradients are used to determine CVD risk. For every 10 μg/m3 of PM2.5 long-term exposure, there was an estimated 8-18% CVD mortality risk [10]. Women had a higher relative risk (RR) (1.42) for PM2.5 induced coronary artery disease than men (0.90) did. Overall, long-term PM exposure increased rate of atherosclerosis and inflammation. In regards to short-term exposure (2 hours), every 25 μg/m3 of PM2.5 resulted in a 48% increase of CVD mortality risk. Additionally, after only 5 days of exposure, a rise in systolic (2.8 mmHg) and diastolic (2.7 mmHg) blood pressure occurred for every 10.5 μg/m3 of PM2.5. Other research has implicated PM2.5 in irregular heart rhythm, reduced heart rate variability (decreased vagal tone), and most notably heart failure. PM2.5 is also linked to carotid artery thickening and increased risk of acute myocardial infarction.

Pathophysiology: Population based studies show that atherosclerosis the major precursor of cardiovascular disease begins in childhood. The Pathobiological Determinants of Atherosclerosis in Youth Study demonstrated that intimal lesions appear in all the aortas and more than half of the right coronary arteries of youths aged 7–9 years.

This is extremely important considering that 1 in 3 people will die from complications attributable to atherosclerosis. In order to stem the tide education and awareness that cardiovascular disease poses the greatest threat and measures to prevent or reverse this disease must be taken.

Obesity and diabetes mellitus are often linked to cardiovascular disease, as are a history of chronic kidney disease and hypercholesterolaemia. In fact, cardiovascular disease is the most life threatening of the diabetic complications and diabetics are two- to four-fold more likely to die of cardiovascular-related causes than nondiabetics [11].

Screening

Screening ECGs (either at rest or with exercise) are not recommended in those without symptoms who are at low risk. In those at higher risk the evidence for screening with ECGs is inconclusive [12].

Some biomarkers may add to conventional cardiovascular risk factors in predicting the risk of future cardiovascular disease; however, the clinical value of some biomarkers is still questionable [13]. Currently, biomarkers which may reflect a higher risk of cardiovascular disease include:

- Coronary artery calcification
- Carotid intima-media thickness
- Carotid total plaque area
- Higher fibrinogen and PAI-1 blood concentrations
- Elevated homocysteine
- Elevated blood levels of asymmetric dimethylarginine
- Inflammation as measured by C-reactive protein
- Elevated blood levels of brain natriuretic peptide (also known as B-type) (BNP)

Prevention

Currently practiced measures to prevent cardiovascular disease include:

- A low-fat, high-fiber diet including whole grains and plenty of fresh fruit and vegetables (at least five portions a day)
- Tobacco cessation and avoidance of second-hand smoke;
- No alcohol consumption; excessive alcohol intake increases the risk of cardiovascular disease.
- Lower blood pressures, if elevated, through the use of antihypertensive medications;
- Decrease body fat (BMI) if overweight or obese;
- Increase daily activity to 30 minutes of vigorous exercise per day at least five times per week;
- Decrease psychosocial stress. Stress however plays a relatively minor role in hypertension. Specific relaxation therapies are not supported by the evidence.

Routine counseling of adults to advise them to improve their diet and increase their physical activity has not been found to significantly alter behavior, and thus is not recommended.

Diet

Evidence suggests that the Mediterranean diet may improve cardiovascular outcomes. On February 25, 2013, medical researchers at the University of Barcelona, based on a five year study of 7,447 people, reported in the New England Journal of Medicine that the Mediterranean diet reduced the risk of heart disease in people at high risk by "about 30 percent" [14].

In clinical trials the DASH diet (high in fruits and vegetables, low in sweets, red meat and fat) has been shown to reduce blood pressure, lower total and low density lipoprotein cholesterol [15]and improve metabolic syndrome;[16] but the long term benefits outside the context of a clinical trial have been questioned.

The link between saturated fat intake and cardiovascular disease is controversial (see Saturated fat and cardiovascular disease controversy) and scientific studies, both observational and clinical, show conflicting results [17]. Dietary substitution of polyunsaturated fats for saturated fats may reduce risk, substitution with carbohydrates does not change or may increase risk. Increased dietary intake of Trans fatty acids significantly increases the risk of cardiovascular disease [18].

The effect of a low salt diet is unclear with any benefit in either hypertensive or normal tensive people being small if present. A low salt diet may be harmful in those with congestive heart failure [19].

Supplements

Evidence to support omega-3 fatty acid supplementation is lacking [20]. As is evidence to support antioxidants and vitamins [21].

Medication

Aspirin has not been found to be of benefit overall in those at low risk of heart disease as the risk of serious bleeding is equal to the benefit with respect to cardiovascular problems[22]. Statins are effective in preventing further cardiovascular disease in those with a history of cardiovascular disease [23]. A decreased risk of death however seems to only occur in men.

Management

Cardiovascular disease is treatable with initial treatment primarily focused on diet and lifestyle interventions. Medication may also be useful for prevention.

Mortality

According to the World Health Organization, cardiovascular diseases are the leading cause of death. In 2008, 30% of all global death is attributed to cardiovascular diseases. Death caused by cardiovascular diseases are also higher in low and middle-income countries as over 80% of all global death caused by cardiovascular diseases occurred in

those countries. It is also estimated that by 2030, over 23 million people will die from cardiovascular diseases annually.

Research

The first studies on cardiovascular health were performed in 1949 by Jerry Morris using occupational health data and were published in 1958 [24]. The causes, prevention, and/or treatment of all forms of cardiovascular disease remain active fields of biomedical research, with hundreds of scientific studies being published on a weekly basis. A trend has emerged, particularly in the early 2000s, in which numerous studies have revealed a link between fast food and an increase in heart disease. These studies include those conducted by the Ryan Mackey Memorial Research Institute, Harvard University and the Sydney Center for Cardiovascular Health. Many major fast food chains, particularly McDonald's, have protested the methods used in these studies and have responded with healthier menu options.

A fairly recent emphasis is on the link between low-grade inflammation that hallmarks atherosclerosis and its possible interventions. C-reactive protein (CRP) is a common inflammatory marker that has been found to be present in increased levels in patients at risk for cardiovascular disease [25]. Also osteoprotegerin which involved with regulation of a key inflammatory transcription factor called NF-κB has been found to be a risk factor of cardiovascular disease and mortality.

Some areas currently being researched include possible links between infection with Chlamydophila pneumoniae (a major cause of pneumonia) and coronary artery disease. The Chlamydia link has become less plausible with the absence of improvement after antibiotic use [26].

Several research also investigated the benefits of melatonin on cardiovascular diseases prevention and cure. Melatonin is a pineal gland secretion and it is shown to be able to lower total cholesterol, very low density and low density lipoprotein cholesterol levels in the blood plasma of rats. Reduction of blood pressure is also observed when pharmacological doses are applied. Thus, it is deemed to be a plausible treatment for hypertension. However, further research needs to be conducted to investigate the side effects, optimal dosage and etc. before it can be licensed for use [27].

1.3 Neural networks for complex medical diagnosis

In this article, an artificial intelligence alternative to the medical diagnosis is proposed. Neural networks are the tools that should be recalled for any classification job. They are developed enormously since the first attempts made modeling the perceptron architecture six decades ago [28].

The massive parallel computational structure of neural networks is what has contributed to its success in predictive tasks. It has been shown that the approach of using parallel networks is successful with respect to increasing the predictive accuracy of neural networks in robotics [29] and in disease diagnosis.

This work presents a parallel networks system which is bound together with a majority voting system in order to further increase the predictive accuracy of a cardiovascular diseases disease data set based on clinic recordings (reference).

For the proposed system it is shown with a case study of cardiovascular diseases. The type of network used is the standard feed forward back-propagation neural network, since they have proven useful in biomedical classification tasks [30]. The performance of the trained neural networks is evaluated according to the true positive, and true negative rate of the prediction task. Furthermore the area under the receiver operating characteristic curve and the Mean Squared Error are used as statistical measurements to compare the success of the different models.

The paper is organized as follows; first, the data used in this work is introduced in section 2. The neural network that is boosted by filtering is illustrated in section 3. Results of the research are shown in section 4 which followed by a conclusion.

2. Data Set of Cardiovascular Diseases

Source of Data: Cleveland Clinic

The Cleveland Clinic, formally known as the Cleveland Clinic Foundation, is a multispecialty academic medical center located in Cleveland, Ohio, United States. The Cleveland Clinic was established in 1921 by four physicians for the purpose of providing patient care, research, and medical education in an ideal medical setting.

The Cleveland Clinic Lerner Research Institute is home to all laboratory-based, translational and clinical research at Cleveland Clinic. A new medical school, the Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, was opened in 2004. The program's curriculum was devised by Cleveland Clinic staff physicians to train and mentor a new generation of physician-investigators.

Database contains 302 data with 76 attributes for each of them, but all published experiments refer to using a subset of 14 of them. The "goal" field refers to the presence of heart disease in the patient. It is integer valued from 0 no presence) to 4.

Table 1: Table describing the 14 attributes that are not used

1.	age
2.	sex
3.	cp
4.	trestbps
5.	chol
6.	fbs
7.	restecg
8.	thalach
9.	exang
10.	oldpeak
11.	slope
12.	ca
13.	thal
14.	num (the predicted attribute)

Experiments with the Cleveland database have concentrated on simply attempting to distinguish presence values 1, 2, 3, 4) from absence value 0) [31].

3. Principal Component Analysis

Principle component analysis (PCA) finds the linear combination of attributes that best accounts for the variations in the data. Two-dimensional plots of the first two principal components supply us with a means to inspect visually for trends, which occur as clusters of points. Later, cluster analysis may follow this step.

This simple but effective method continues to be used today, partly because of the ease with which the results are communicated and interpreted.

3.1 Theory of Principal component Analysis

Multivariate statistics deals with the relation between several random variables. The sets of observations of the random variables are represented by a multivariate data matrix $\boldsymbol{X}$,

Multivariate statistics deals with the relation between several random variables. The sets of observations of the random variables are represented by a multivariate data matrix $\boldsymbol{X}$,

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ x_{31} & x_{32} & \cdots & x_{3p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}. \quad (1)$$

Each column vector $\boldsymbol{u}_k$ represents the data for a different variable. If $\boldsymbol{c}$ is an $p \times 1$ matrix, then

$$\boldsymbol{Xc} = c_1 \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \\ \vdots \\ x_{n1} \end{bmatrix} + c_2 \begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \\ \vdots \\ x_{n2} \end{bmatrix} + \cdots + c_p \begin{bmatrix} x_{1p} \\ x_{2p} \\ x_{3p} \\ \vdots \\ x_{np} \end{bmatrix} \quad (2)$$

is a linear combinations of the set of observations.

Descriptive statistics can also be applied to a multivariate data matrix $\boldsymbol{X}$, the sample mean of the kth variable is

$$\bar{x}_k = \frac{1}{n}\sum_{i=1}^{n} x_{ik}, k = 1,2,\dots,p, \quad (3)$$

the sample variance is defined by

$$s_k^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{ik} - \bar{x}_k)^2, k = 1,2,\dots,p. \quad (4)$$

Next we introduce a matrix that contains statistics that relate pairs of variables (x_i, x_k), sample covariance s_{ik}:

$$s_{ik} = \frac{1}{n}\sum_{j=1}^{n}(x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k), \quad i = 1,2,\dots,p, k = 1,2,\dots,p. \quad (5)$$

It follows that $s_{ik} = s_{ki}$ and $s_{ii} = s_i^2$, the sample variance.

Matrix of sample covariances

$$\boldsymbol{S_n} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ s_{31} & s_{32} & \cdots & s_{3p} \\ \vdots & \vdots & \vdots & \vdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{bmatrix} \quad (6)$$

is symmetric.

THEOREM Let $\boldsymbol{S_n}$ be the $p \times p$ covariance matrix related to the multivariate data matrix $\boldsymbol{X}$. Let eigenvalues of $\boldsymbol{S_n}$ be $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, and corresponding orthonormal eigenvectors be $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_p}$. Then ith principal component $\boldsymbol{y}_i$ is given by the linear combination of the original variables in the data matrix $\boldsymbol{X}$:

$$\boldsymbol{y}_i = \boldsymbol{X}\boldsymbol{u_i}, i = 1,2,..,p. \quad (7)$$

The variance of $\boldsymbol{y}_i$ is λ_i, and $\text{cov}(\boldsymbol{y}_i, \boldsymbol{y}_j) = 0, i \neq j$. The total variance of the data in $\boldsymbol{X}$ is equal to the sum of eigenvalues:

$$\sum_{j=1}^{p} s_{jj} = \sum_{j=1}^{p} \lambda_j. \quad (8)$$

Proportion of the total variance covered by the "kth principal component"

$$= \frac{\lambda_k}{\sum_{j=1}^{p} \lambda_j}. \quad (9)$$

If a large percentage of the total variance can be attributed to the first few components, then these new variables can replace the original variables without significant loss of information. Thus we can achieve significant reduction in data.

4. Principal Components of Cardiovascular Diseases Data

The information in the covariance matrix is used to define a set of new variables as a linear combination of the original variables in the data matrices**.** The new variables are derived in a decreasing order of importance. The first of them is called first principal component and accounts for as much as possible of the variation in the original data. The second of them is called second principal component and accounts for another, but smaller portion of the variation, and so on.

If there are *p* variables, to cover all of the variation in the original data, one needs *p* components, but often much of the variation is covered by a smaller number of components. Thus PCA has as its goals the interpretation of the variation and data reduction.

In fact PCA is nothing but the spectral decomposition of the covariance matrix. The fourteenth component of the data which is related to the absence and presence of the disease removed, and the principal transformation of the data is realized. The first two principal components of these two types of data are intermingled as seen in Figure 1. In our classification perceptron, first five principal components are found to be satisfactory.

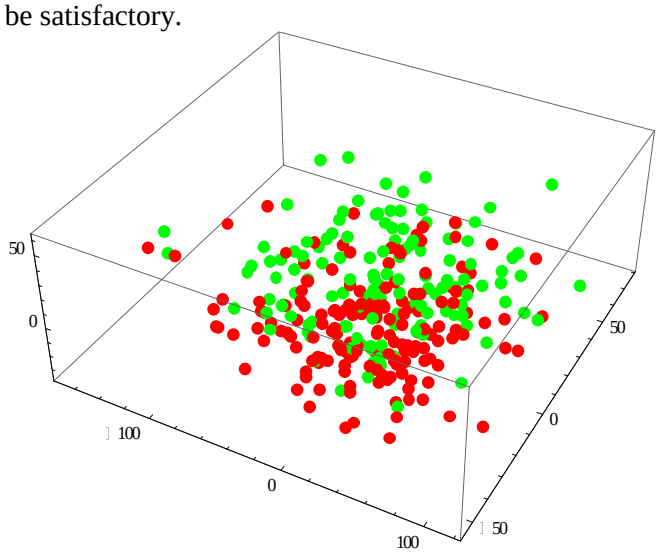


Fig. 1. The distribution obtained by the use of first three principal components of absence and presence data

5. Artificial Neural Networks

Nervous systems existing in biological organism for years have been the subject of studies for mathematicians who tried to develop some models describing such systems and all their complexities. Artificial Neural Networks emerged as generalizations of these concepts with mathematical model of artificial neuron due to McCuloch and Pitts [32] described in 1943 definition of unsupervised learning rule by Hebb [33] in 1949, and the first ever implementation of Rosenblatt's perceptron [34] in 1958. The efficiency and applicability of artificial neural networks to computational tasks have been questioned many times, especially at the very beginning of their history the book "Perceptrons" by Minsky and Papert [35], published in 1969, caused dissipation of initial interest and enthusiasm in applications of neural networks.

It was not until 1970s and 80s, when the back propagation algorithm for supervised learning was documented that artificial neural networks regained their status and proved beyond doubt to be sufficiently good approach to many problems. Artificial Neural Network can be looked upon as a parallel computing system comprised of some number of rather simple processing units (neurons) and their interconnections. They follow inherent organizational principles such as the ability to learn and adapt, generalization, distributed knowledge representation, and fault tolerance. Neural network specification comprises definitions of the set of neurons (not only their number but also their organization), activation states for all neurons expressed by their activation functions and offsets specifying when they fire, connections between neurons which by their weights determine the effect the output signal of a neuron has on other neurons it is connected with, and a method for gathering information by the network that is its learning (or training) rule.

5.1. Architecture

From architecture point of view neural networks can be divided into two categories: feed-forward and recurrent networks. In feed-forward networks the flow of data is strictly from input to output cells that can be grouped into layers but no feedback interconnections can exist. On the other hand, recurrent networks contain feedback loops and their dynamical properties are very important.

The most popularly used type of neural networks employed in pattern classification tasks is the feedforward network which is constructed from layers and possesses unidirectional weighted connections between neurons. The common examples of this category are Multilayer Perceptron or Radial Basis Function networks, and committee machines.

Multilayer perceptron type is more closely defined by establishing the number of neurons from which it is built, and this process can be divided into three parts, the two of which, finding the number of input and

output units, are quite simple, whereas the third, specification of the number of hidden neurons can become crucial to accuracy of obtained classification results.

The number of input and output neurons can be actually seen as external specification of the network and these parameters are rather found in a task specification. For classification purposes as many distinct features are defined for objects which are analyzed that many input nodes are required. The only way to better adapt the network to the problem is in consideration of chosen data types for each of selected features. For example instead of using the absolute value of some feature for each sample it can be more advantageous to calculate its change as this relative value should be smaller than the whole range of possible values and thus variations could be more easily picked up by Artificial Neural Network. The number of network outputs typically reflects the number of classification classes.

The third factor in specification of the Multilayer Perceptron is the number of hidden neurons and layers and it is essential to classification ability and accuracy. With no hidden layer the network is able to properly solve only linearly separable problems with the output neuron dividing the input space by a hyperplane. Since not many problems to be solved are within this category, usually some hidden layer is necessary.

With a single hidden layer the network can classify objects in the input space that are sometimes and not quite formally referred to as simplexes, single convex objects that can be created by partitioning out from the space by some number of hyperplanes, whereas with two hidden layers the network can classify any objects since they can always be represented as a sum or difference of some such simplexes classified by the second hidden layer.

Apart from the number of layers there is another issue of the number of neurons in these layers. When the number of neurons is unnecessarily high the network easily learns but poorly generalizes on new data. This situation reminds auto-associative property: too many neurons keep too much information about training set rather "remembering" than "learning" its characteristics. This is not enough to ensure good generalization that is needed.

On the other hand, when there are too few hidden neurons the network may never learn the relationships amongst the input data. Since there is no precise indicator how many neurons should be used in the construction of a network, it is a common practice to build a network with some initial number of units and when it trains poorly this number is either increased or decreased as required. Obtained solutions are usually task-dependant.

5.2 Activation Functions

Activation or transfer function of a neuron is a rule that defines how it reacts to data received through its inputs that all have certain weights.

Among the most frequently used activation functions are linear or semi-linear function, a hard limiting thresh-old function or a smoothly limiting threshold such as a sigmoid or a hyperbolic tangent. Due to their inherent properties, whether they are linear, continuous or differentiable, different activation functions perform with different efficiency in task-specific solutions.

For classification tasks antisymmetric sigmoid tangent hyperbolic function is the most popularly used activation function:

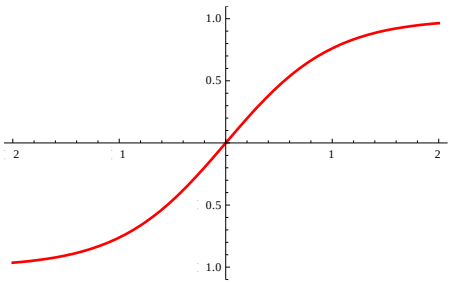


Fig. 2. Antisymmetric sigmoid tangent hyperbolic activation function

5.3 Learning Rules

In order to produce the desired set of output states whenever a set of inputs is presented to a neural network it has to be configured by setting the strengths of the interconnections and this step corresponds to the network learning procedure. Learning rules are roughly divided into three categories of supervised, unsupervised and reinforcement learning methods.

The term supervised indicates an external teacher who provides information about the desired answer for each input sample. Thus in case of supervised learning the training data is specified in forms of pairs of input values and expected outputs. By comparing the expected outcomes with the ones actually obtained from the network the error function is calculated and its minimization leads to modification of connection weights in such a way as to obtain the output values closest to expected for each training sample and to the whole training set.

In unsupervised learning no answer is specified as expected of the neural network and it is left somewhat to itself to discover such self-organization which yields the same values at an output neuron for new samples as there are for the nearest sample of the training set.

Reinforcement learning relies on constant interaction between the network and its environment.

The network has no indication what is expected of it but it can induce it by discovering which actions bring the highest reward even if this reward is not immediate but delayed. Basing on these rewards it performs such re-organization that is most advantageous in the long run [34].

The modification of weights associated with network interconnections can be performed either after each of the training samples or after finished iteration of the whole training set.

The important factor in this algorithm is the learning rate η whose value when too high can cause oscillations around the local minima of the error function and when too low results in slow convergence. This locality is considered the drawback of the backpropagation method but its universality is the advantage.

5.4 Architecture of artificial neural networks, Committee Machines

As the base topology of artificial neural network committee machines [24] with the feed-forward multilayer perceptron with sigmoid activation function trained by backpropagation algorithm is used.

In committee machines approach, a complex computational task is solved by dividing it into a number of computationally simple tasks and then combining the solutions to those tasks. In supervised learning, computational simplicity is achieved by distributing the learning task among a number of experts, which in turn divides the input space into a set of subspaces. The combination of experts is said to constitute a committee machine. Basically, it fuses knowledge acquired by experts to arrive at an overall decision that is supposedly superior to that attainable by anyone of them acting alone. The idea of a committee machine may be traced back to Nilsson [36] (1965); the network structure considered therein consisted of a layer of elementary perceptrons followed by a vote-taking perceptron in the second layer.

Committee machines are universal approximators. They may be classified into two major categories:

1. Static structures. In this class of committee machines, the responses of several predictors (experts) are combined by means of a mechanism that does not involve the input signal, hence the designation "static." This category includes the following methods:

• Ensemble averaging, where the outputs of different predictors are linearly combined to produce an overall output.

• Boosting, where a weak learning algorithm is converted into one that achieves arbitrarily high accuracy.

2. Dynamic structures. In this second class of committee machines, the input signal is directly involved in actuating the mechanism that integrates the outputs of the individual experts into an overall output, hence the designation "dynamic."

5.5 Boosting

Boosting is a method that belongs to the "static" class of committee machines. Boosting is quite different from ensemble averaging. In a committee machine based on ensemble averaging, all the experts in the machine are trained on the same data set; they may differ from each other in the choice of initial conditions used in network training. By contrast, in a boosting machine the experts are trained on data sets with entirely different distributions; it is a general method that can be used to improve the performance of any learning algorithm.

Boosting' can be implemented in three fundament-ally different ways:

1. Boosting by filtering. This approach involves filtering the training examples by different versions of a weak learning algorithm. It assumes the availability of a large (in theory, infinite) source of examples, with the examples being either discarded or kept during training. An advantage of this approach is that it allows for a small memory requirement compared to the other two approaches.

2. Boosting by subsampling. This second approach works with a training sample of fixed size. The examples are "resampled" according to a given probability distribution during training. The error is calculated with respect to the fixed training sample.

3. Boosting by reweighting. This third approach also works with a fixed training sample, but it assumes that the weak learning algorithm can receive "weighted" examples. The error is calculated with respect to the weighted examples.

In this paper Boosting by filtering is used. This algorithm is due to Schapire [36] (1990). The original idea of boosting described in Schapire (1990) is rooted in a distribution free or probably approximately correct (PAC) model of learning. To be more specific, the goal of the learning machine is to find a hypothesis or prediction rule with an error rate of at most ε, for arbitrarily small positive values of ε, and this should hold uniformly for all input distributions.

In boosting by filtering, the committee machine consists of three experts or subhypotheses. The algorithm used to train them is called a boosting algorithm. The three experts are arbitrarily labeled "first," "second," and "third." The three experts are individually trained as follows:

1. The first expert is trained on a set consisting of N_2 examples.

2. The trained first expert is used to filter another set of examples by proceeding in the following manner:

Flip a fair coin; this in effect simulates a random guess. If the result is heads, pass new patterns through the first expert and discard correctly classified patterns until a pattern is misclassified. That misclassified pattern is added to the training set for the second expert.

If the result is tails, do the opposite. Specifically, pass new patterns through the first expert and discard incorrectly classified patterns until a pattern is classified correctly. That correctly classified pattern is added to the training set for the second expert.

Continue this process until a total of N_1 examples has been filtered by the first expert. This set of filtered examples constitutes the training set for the second expert.

By following this coin flipping procedure, it is ensured that if the first expert is tested on the second set of examples, it would have an error rate of 1/2. In other words, the second set of N_1 examples available for training the second expert has a distribution entirely different from the first set of N_2 examples used to train the first expert. In this way the second expert is forced to learn a distribution different from that learned by the first expert [37].

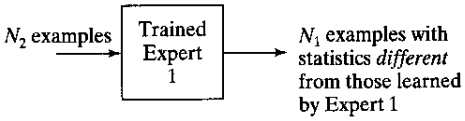
3. Once the second expert has been trained in the usual way, a third training set is formed for the third expert by proceeding in the following manner:

• Pass a new pattern through both the first and second experts. If the two experts agree in their decisions, discard that pattern. If, on the other hand, they disagree, the pattern is added to the training set for the third expert.

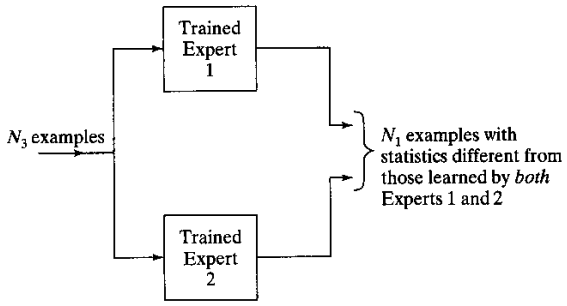
• Continue with this process until a total of N_1 examples have been filtered jointly by the first and second experts. This set of jointly filtered examples constitutes the training set for the third expert.

The third expert is then trained in the usual way, and the training of the entire committee machine is thereby completed.

Let N_2 denote the number of examples that must be filtered by the first expert to obtain the training set of N1 examples for the second expert. Note that N_1 is fixed, and N_2 depends on the generalization error rate of the first expert. Let N_3 denote the number of examples that must be jointly filtered by the first and second experts to obtain the training set of N_1 examples for the third expert [38].



a.Filtering of data by Expert 1



b. Filtering of data by Expert 2 and 3

Fig. 3.The three-point filtering procedure

With N_1 examples also needed to train the first expert, the total size of data set needed to train the entire committee machine is $N = N_1 + N_2 + N_3$. However, the computational cost is based on 3N1 examples because N1 is the number of examples actually used to train each of the three experts. We may therefore say that the boosting algorithm described herein is indeed "smart" in the sense that the committee machine requires a large set of examples for its operation, but only a subset of that data set is used to perform the actual training.

Table 2. Number of samples used at each stage of the training-testing processes.

	N_1	N_2	N_3	Test
Sick	47	50	37	50
Healthy	53	50	39	50

Another noteworthy point is that the filtering operation performed by the first expert and the joint filtering operation performed by the first and second experts make the second and third experts, respectively, focus on "hard-to-learn" parts of the distribution.

During the training stage, the performances of committee members, are shown in Table 3.

Table 3. Positives, and Negatives in training stage.

	First	Second	Third
True positive %	98	94	97
True negative %	98	96	97
False positive %	2	4	3
False negative %	2	6	3

In the theoretical derivation of the boosting algorithm originally presented in Schapire (1990)[36], simple voting was used to evaluate the performance of the committee machine on test patterns not seen before. Specifically, a test pattern is presented to the committee machine. If the first and second experts in the committee machine agree in their respective

decisions, that class label is used. Otherwise, the class label discovered by the third expert is used. However, in experimental work presented in Drucker et al.[39-40] (1993,1994), it has been determined that addition of the respective outputs of the three experts yields a better performance than voting. For example, in the optical character recognition (OCR) problem, the addition operation is performed simply by adding the "digit 0" outputs of the three experts, and likewise for the other nine digit outputs.

The number of input terminals equaled the number of attributes in the human voice data, thus it is eleven. There are two hidden layers with eleven neurons within each of three neural networks in the committee machine for preserving generalization properties but achieving convergence during training with tolerance at most 0.14 for all training samples recognized properly.

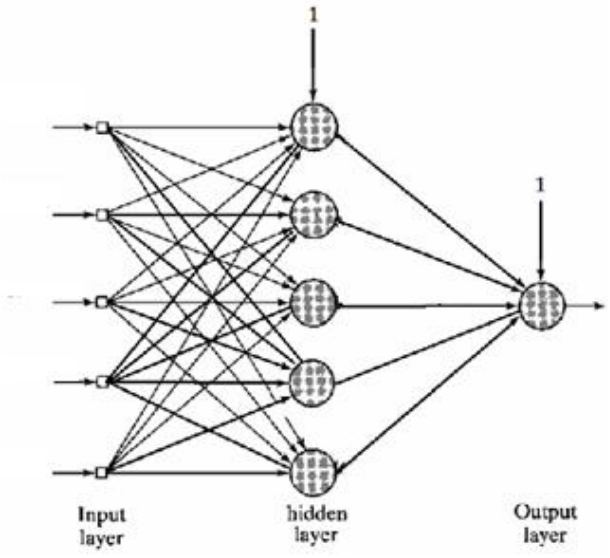


Fig. 4. Signal flow graph of each of the three expert machines with one hidden layer.

For all structures of artificial neural networks, only one output is produced. Actually, it was possible to use a single output and by interpretation of its active state as one class and inactive output state the second class the task would have been solved as well, but with such approach the text is attributed to either one or another author and classification is binary. Algorithm results in a decision about attribution of paragraphs whose textual description entered as inputs.

6. Results and Discussion

To perform the boosting by filtering technique, we the training data are chosen in a special way described in Section 3.5. A balanced set of 50-50 positive and negative members are chosen from available data for testing. During the testing stage, the performances of committee members and success in the final decision are shown in Table 4.

Table 4. False positives, and false negatives in testing stage.

	First	Second	Third	Majority
True positive %	74	80	88	82
True negative %	88	81	72	84
False positive %	12	19	28	16
False negative %	26	20	12	18

It has been shown that parallel neural networks, when boosted by filtering, in combination with a majority voting increase performance of true recognition rates in an imbalanced data set.

The data set is very unbalanced with regard to the class distribution. This, in combination with the small sample size, makes it difficult to train any type of classifier to predict the presence of Cardiovascular disease.

Out of 302 samples, 46% are cardiovascular disease type and the remainder is of healthy character.

False positive rates up to 25 - 40% of the positive class have been reported [41-50] in the literature. It has been demonstrated in this study that a true negative rate up to 90% can be achieved by using three parallel networks, and majority voting. This is a significant improvement compared to previously demonstrated results

7. Conclusions

A system has been presented consisting of parallel distributed neural networks with one hidden layer, boosted by the use of filtering, and a majority voting system. The different expertise of the committee members increases the robustness of the system. An empirical investigation demonstrates that it is possible to achieve >90% true positive rate for each class in a Cardiovascular disease data set.

References

[1] M. Anthea, Hopkins J., C. W. McLaughlin, S. Johnson, M. Q. Warner, D. LaHart, J. D. Wright (1993) *Human Biology and Health,* Englewood Cliffs, New Jersey: Prentice Hall.

[2] B. B. Kelly (2010) *Promoting Cardiovascular Health in the Developing World: A Critical Challenge to Achieve Global Health.* Washington, D.C: National Academies Press.

[3] A. P. Dantas, F. Jimenez-Altayo, E.Vila (2012) Vascular aging: facts and factors. *Frontiers in Vascular Physiology* 3 (325): 1–2.

[4] B. B. Kelly, editor (2010) *Promoting cardiovascular health in the developing world: a critical challenge to achieve global health*. Washington, D.C.: National Academies Press. Chapter 2.

[5] S. Mendis, P. Puska, and B. Norrving (editors) (2011) *Global Atlas on cardiovascular disease prevention and control*.

[6] H.C. McGill, C.A. McMahan, S.S. Gidding (2008). Preventing heart disease in the 21st century: implications of the Pathobiological Determinants of Atherosclerosis in Youth (PDAY) study, *Circulation* 117 (9) pp. 1216–27.

[7] J. Vartiainen, T. Puska (1999) Sex, age, cardiovascular risk factors, and coronary heart disease, *Circulation* 99: 1165–1172.

[8]http://www.world-heart-federation.org/ cardiovascular-health/cardiovascular-disease-risk-factors

[9] J. B, Rajkumar (2006) Ageing and vascular ageing, *Postgrad Med J* 82: 357–362.

[10]. M. Khallaf, (2011) The Impact of Air Pollution on Health, Economy, Environment and Agricultural Sources. *InTech*. pp. 69–92.

[11] E. Kvan, K.I. Pettersen, L. Sandvik, and A. Reikvam (2007) High mortality in diabetic patient with acute myocardial infarction: cardiovascular co-morbidities contribute most to the high risk, *Int J Cardiol* 121: 184–188.

[12] V.A. Moyer, U.S. Preventive Services Task, Force (2012) Screening for coronary heart disease with electrocardiography: U.S. Preventive Services Task Force recommendation statement. *Annals of internal medicine*157 (7): 512-8.

[13] T.J. Wang, P. Gona, M.G. Larson, G.H. Tofler, D. Levy, C. Newton-Cheh, P.F. Jacques, N. Rifai, J. Selhub, S.J. Robins, E.J. Benjamin, R.B. D'Agostino, and R.S. Vasan (2006) Multiple biomarkers for the prediction of first major cardiovascular events and death, *N. Engl. J. Med*. 355 (25): 2631–9.

[14] R. Estruch, et al (2013) Primary Prevention of Cardiovascular Disease with a Mediterranean Diet. *New England Journal of Medicine.*

[15] E. Obarzanek, F.M. Sacks, W.M. Vollmer et al. (2001) Effects on blood lipids of a blood pressure-lowering diet: the Dietary Approaches to Stop Hypertension (DASH) Trial, *Am. J. Clin. Nutr.* 74 (1): 80–9.

[16] L. Azadbakht, P. Mirmiran, A. Esmaillzadeh, T. Azizi, and F. Azizi (2005) Beneficial effects of a Dietary Approaches to Stop Hypertension eating plan on features of the metabolic syndrome. *Diabetes Care* 28 (12): 2823–31.

[17] P.W. Siri-Tarino, Sun Qi, Hu, B. Frank, and M. Krauss Ronald (2010) Saturated fat, carbohydrate, and cardiovascular disease, *American Journal of Clinical Nutrition* 91 (3): 502–509.

[18] D. Mozaffarian, A. Aro. W.C. Willett (2009) Health effects of trans-fatty acids: experimental and observational evidence. *European journal of clinical nutrition 63 Suppl 2*: S5-21.

[19] R.S. Taylor, K.E. Ashton, T. Moxham, L. Hooper, and S. Ebrahim, (2011) Reduced dietary salt for the prevention of cardiovascular disease, *Cochrane database of systematic reviews* (Online) (7)

[20] S.M. Kwak, S.K. Myung, Y.J. Lee, H.G. Seo, for the Korean Meta-analysis Study, Group (2012) Efficacy of Omega-3 Fatty Acid Supplements (Eicosapentaenoic Acid and Docosahexaenoic Acid) in the Secondary Prevention of Cardiovascular Disease: A Meta-analysis of Randomized, Double-blind, Placebo-Controlled Trials, *Archives of Internal Medicine.*

[21] S.K. Myung, W. Ju, B. Cho, S.W. Oh, S.M. Park, B.K. Koo, and B.J. Park, for the Korean Meta-Analysis (KORMA) Study, Group (2013) Efficacy of vitamin and antioxidant supplements in prevention of cardiovascular disease: systematic review and meta-analysis of randomised controlled trials, *BMJ (Clinical research ed.)* 346.

[22] J.S. Berger, A. Lala, M.J. Krantz, G.S. Baker, and W.R. Hiatt, (2011) Aspirin for the prevention of cardiovascular events in patients without clinical cardiovascular disease: a meta-analysis of randomized trials, *American heart journal* 162(1): 115–24.

[23] J. Gutierrez, G. Ramirez, T. Rundek, and R.L. Sacco (2012) Statin Therapy in the Prevention of Recurrent Cardiovascular Events: A Sex-Based Meta-analysis Statin Therapy to Prevent Recurrent CV Events, *Archives of Internal Medicine* 172 (12): 909–19.

[24] J.N. Morris, M.D. Crawford (1958) Coronary Heart Disease and Physical Activity of Work. *British Medical Journal* 2 (5111): 1485–1496.

[25] M. Karakas, and W. Koenig (2009) CRP in cardiovascular disease, *Herz* 34 (8): 607–13.

[26] R. Andraws, J.S. Berger, D.L. Brown (2005) Effects of antibiotic therapy on outcomes of patients with coronary artery disease: a meta-analysis of randomized controlled trials. *JAMA* 293 (21): 2641–7.

[27] A. Dominguez-Rodriguez (2012) Melatonin and Cardiovascular Disease: Myth or Reality?. *Rev Esp Cardiol* 65: 215-218.

[28] M. L. Minsky, and S. A. Papert, (1988) *Perceptrons*, Expanded Edition. Cambridge, MA: MIT Press. Original edition, 1969.

[29] R. Koker (2008) Reliability-based approach to the inverse kinematics solution of robots using Elman's networks, *Engineering Applications of Artificial Intelligence* (18) 685 – 693.

[30] A. J. Lee (1997) Parallel neural networks for speech recognition, *Neural Networks*,1997., International Conference on 4 (9-12) 2093–2097.

[31] R. Detrano ,V.A. Medical Center, Long Beach and Cleveland Clinic Foundation.

http://archive.ics.uci.edu/ml/machine-learning-databases/heart-disease/heart-disease.names

[32] W. S. Mcculloch, and W. Pitt's. (1943) A Logical Calculus of the Ideas Immanent in Nervous Activity, *Bulletin of Mathematical Biophysics*, 5:115-133. Reprinted in Anderson& Rosenfeld [1988], pp. 18-28.

[33] B. O. Hebb, (1949). *The Organization of Behavior*. New York: John Wiley & Sons. Introduction and Chapter 4 reprinted in Anderson & Rosenfeld 1988, pp. 45-56.

[34] E, Rosenblatt, (1958) The Perceptron: A probabilistic model for information storage and organization in the brain, *Psychological Review*, vol. 65, pp. 386-408.

[35] M. L. Minsky, and S. A. Papert, (1988) *Perceptrons*, Expanded Edition. Cambridge, MA: MIT Press. Original edition, 1969.

[36] R. E. Schapire (1990) The strength of weak learnability, *Machine Learning*, vol. 5, pp.197-227.

[37]R. E. Schapire (1997) Using output codes to boost multiclass learning problems, Machine Learning: *Proceedings of the Fourteenth International Conference*, Nashville, TN.

[38] Simon Haykin. Neural Network A Com-prehensive Foundation, Pearson Education, Inc, 1999,ch.4, pp.156-226.

[39] H. Drucker, C. Cortes, L.D. Jackel, and Y. LeCun, Boosting and other ensemble methods. Neural Computation, vol. 6, 1994, pp.1289-1301

[40] H. Drucker, R.E. Schapire, and P. Simard, Improving performance in neural networks using a boosting algorithm, Advances in Neural Information Processing Systems, vol. 5, 1993, pp. 42-49, Cambridge, MA: MIT Press.

[41] W. David Aha, and Dennis Kibler, "Instance-based prediction of heart disease presence with the Cleveland database"

[42] J. W. Everhart, J. E. Dickson, W. C. Knowler, W. C. Johannes (1988) Using the ADAP learning algorithm to forecast the onset of diabetes mellitus, Proc. Symp. on Computer Applications and Medical Care, pp. 261–5.

[43] R. Setiono (1998) Techniques for extracting rules from artificial neural networks, 5th International Conference on Soft Computing and Information Systems, Iizuka, Japan.

[44] W. H. Wolberg, and O.L. Mangasarian (1990) Multisurface method of pattern separation for medical diagnosis applied to breast cytology, Proceedings of the National Academy of Sciences, USA, Volume 87, pp 9193-9196.

[45] L. Prechelt (1994) Prolem1 - A Set of Neural Network Benchmark Problems and Benchmarking Rules, Technical Report 21/94, 76128 Karlsruhe, Germany.

[46] R. Parekh, J. Yang, and V. Honavar (2000) Constructive Neural Network Learning Algorithms for Pattern Classification, *IEEE Transactions On Neural Networks*, Vol. 11,No. 2.

[47] R. Setiono, and L.C.K. Hui(1995) Use of quasi-Newton method in a feedforward neural network construction algorithm, *IEEE Transactions on Neural Networks*, vol. 6, no. 1, pp. 273-277.

[48] Rüdiger W. Brause,"Medical Analysis and Diagnosis by Neural Networks", J.W. Goethe-University, Computer Science Dept., Frankfurt a. M., Germany.

[49] R. Setiono and H. Liu (1996) Improving Back-propagation Learning with Feature selection", *Applied Intelligence*, Vol. 6, No. 2, pp. 129-140.

[50] S. M. Kamruzzaman, A.R. Hasan, A.B. Siddiquee, and Md. E. H. Mazumder (2004) Medical diagnosis using neural network, 3rd International Conference on Electrical & Computer Engineering ICECE 2004, 28-30 December 2004, Dhaka, Bangladesh.

ISSN 2303-4521
PERIODICALS OF ENGINEERING AND NATURAL SCIENCES
Vol. 1 No. 2 (2013)
Available online at: http://pen.ius.edu.ba

Analyzing Classified Listings at an E-Commerce Site by Using Survival Analysis

Ayhan Demiriz
Dept. of Industrial Engineering
Sakarya University
Sakarya, Turkey
ademiriz@gmail.com

Sümeyye Şen
Dept. of Computer Science
Sakarya University
Sakarya, Turkey
sumeyye@sakarya.edu.tr

Furkan Gökçegöz
Production Planning
Timay & Tempo
Merzifon, Turkey
furkan.gokcegoz@timay-tempo.co

Abstract:

Sahibinden.com is a leading e-commerce site in Turkey where sellers (buyers) may advertise their goods (needs) with or without a fee. Since it generates a large volume of traffic to the classified car listings, the site plays an important role for determining the market value of the used cars. In this study, we first randomly selected 200 car classifieds from 950 new classified ads on the day of February 22, 2012. We then observed these listings on a daily basis for a month to determine the possible updates and deletions of the ads. We assume that if an ad is taken out it means that the car has been sold. In addition to the cars' features, we observed the posted price and the number of daily views of the ads throughout the data collection. Therefore one can construct survival models to study the effects of the features and price of a car on the life of the ad. In other words, it is possible to study that what features and price levels expedite the sales of used cars.

Keywords: Survival Analysis, Used Car Sales, e-Commerce, Used Car Price Elasticity, Censored Data.

1. INTRODUCTION

Automotive industry is one of the leading contributors to GDP in the developed countries. Considering that automobiles form 70% of this industry (in terms of number of units), the importance of the automobile trade is obvious (Onat, 2007).

In addition, second hand (used) car (Wikipedia, 2013) sales have far exceeded the new car sales in many countries which shows the importance of the used car sales in world economy (Asilkan, 2009). For example, used car sales have a volume of over twice as much as the new car sales in the U.S.A (Lee, 2006).

As Internet has become very important medium for the second-hand car sales market in developed countries since the most people have access to the internet in these countries and second-hand car dealers and buyers can reach the other party readily on this environment. As a developing country, the amount of purchases made over the Internet is rapidly increasing in Turkey. For instance, there has been an increase of almost 20 percent in January-Februay period of 2013 compared to same period of previous year which has brought the annual monetary volume of total internet sales in Turkey to 5.2 billion Turkish lira (Dünya, 2013).

Sahibinden.com is one of the leading e-commerce sites in Turkey with a number of more than 2 million ads. According to the data from Sahibinden.com, the number of vehicles sold or rented within the first three months of 2013 over the same period of the previous year increased by 17 percent to be around 347,000 vehicles.

Considering that one vehicle is sold or rented every 23 seconds, the volume of the used vehicle listings (advertisements or ad) at Sahibinden.com is very significant. Indeed, the automobile listings in the category of Vehicles have a 59% share in terms of number of listings among 16 different types of vehicle categories and the number of listings in the category of vehicles within all categories has a 42% share of listings at Sahibinden.com e-commerce site.

A particular car listing may stay active for a number of days. Indeed, the time that the listing stays active can be considered as a random variable. There might be various reasons that an ad can be taken from the web site. However, it might be acceptable to assume that the particular car might have been sold by the time the ad is removed. The main purpose of this study is to determine the important automobile characteristics that affect directly the automobile sales. To achieve this, we collected data (Gökçegöz, 2012) by observing a random sample of automobile listings at Sahibinden.com for 30 days. We then analyzed the data to determine the impacts of the various car features on the sales. In Section 2, we give the background information on survival analysis. We then give details of the data used in this paper in Section 3. Statistical analysis of the data is given in Section 4. We then conclude the paper in Section 5.

2. SURVIVAL ANALYSIS

Survival Analysis is used for in the analysis of the data which are obtained at the realization of a predetermined event (such as death, failure etc.) at any time. The main challenge encountered in the analysis of survival data is that by the time predetermined

DOI: 10.21533/pen.v1i2.19

event has occurred we may no longer observe the object to collect the data. In other words, object may survive for a longer period that observations are no longer collected. These cases are called as right censored observations and mostly have longer survival times (Nelson, 1982). Analysis of such data has been one of the main problems of the statisticians. Like the rest of the data, censored observations should be used correctly to achieve better results.

There are various approaches for solving problems related to the survival analysis. In one of these approaches, survival analysis is conducted by using a variety of parametric survival distributions. Another approach is based on the nonparametric distribution analysis which can be used without any prior statistical distribution assumptions. In this study, the outcomes of the analyses are presented by both parametric nonparametric approaches.

Because there are two major analysis methods, the analysis of censored survival data leads to the problem of choices. Of these, the advantages of the non-parametric method of analysis are simple calculations and understandability of the outcomes. In nonparametric analysis method, Kaplan-Meier (Kaplan & Meier, 1958) is one of the commonly used calculation methods. On the other hand, parametric models are unbiased even if underlying distribution hypothesis is no longer valid as they are robust methods.

Parametric modeling will yield superior results when the preferred parametric distribution matches with the data. However, censored data particularly may result in poor outcomes when used in conjunction with the parametric methods. In short, best suitable survival analysis methods have been utilized in this paper to overcome problems that stem from real life data.

2.1. Survival Function

Survival time is the time interval for a person who is exposed to a specific disease until he heals or dies. Survival time of the individual or the system, indicated by *T*, is a random variable. The probability of an individual to live more than a certain time *t* is called the survival function. The survival function is given by following equation (Wang et al., 2002),

$$S(t) = P(T > t) = 1 - F(t) = \int_t^{\infty} f(u)du, 0 \leq t < \infty \qquad (1)$$

where F(t) is the cumulative distribution function of survival time and f(u) is the probability distribution function. Equivalently, hazard function can be defined based on survival function as,

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt}\ln\big(S(t)\big), \qquad (2)$$

which specifies the instantaneous failure rate at time t (Wang et al., 2002). Both survival and hazard functions are used extensively for analyzing the survival data in practice.

3. DATA COLLECTION

The data used in this paper were gathered from Sahibinden.com web site. Initially, 200 automobile listings were selected among 950 automobile listings posted on 22 February. Selected advertisements were observed for 30 days from the date of February 22, 2012 until 22 March 2012. The raw data obtained were cleaned and prepared for the analyses.

Automobile listings at Sahibinden.com are presented with car features that indicate used or new automobile, price, brand, model, type, mileage, color, engine capacity, engine power, fuel type, gear type, body type, transmission, warranty status, trade-in options, and the responsible party (owner or dealer). Number of page view (i.e. seen by site visitors) is also shown on listing pages. All of these features were collected by a special software developed to fetch the web pages of 200 random listings used in this paper. The cleaned data were stored in an Excel file for further processing.

The data gathering software was run every day to collect data for each listing to detect price changes, number of viewers and the status of the listing i.e. whether it was removed or not since the previous day. If a particular listing was no longer accessible, it was then assumed that the car was sold. Then the variable representing the death (failure) is assigned 1. The death time, *t*, was noted for that particular listing. The removal of the listing corresponds to the death or failure in our survival analysis approach.

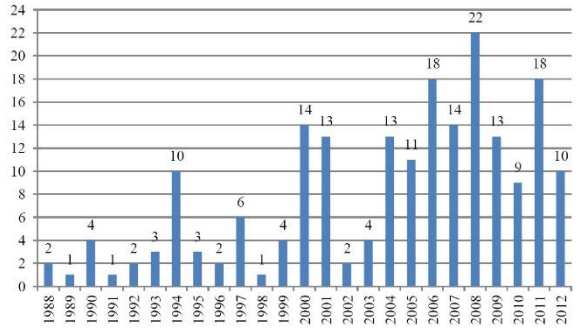


Figure 1. Distribution of Cars by Model Year

Once the 30-day long data collection task was completed, the dataset was preprocessed to convert prices given in foreign currencies to Turkish Lira (TL) based on the exchange rates on the day of original data collection. Some other minor discrepancies were also resolved during the data preprocessing step.

We give the summary charts in *Figure 1* and *Figure 2* to depict the content of the dataset. *Figure 1* summarizes model year of the cars in the dataset. Basically most of the cars are used less than ten years. *Figure 2* summarizes the composition of the dataset from the brand point of view. Again "1" represents the sold cars and "0" represents the unsold cars in *Figure 2*. It is easier to see the distribution of the cars by the brand for the dataset collected.

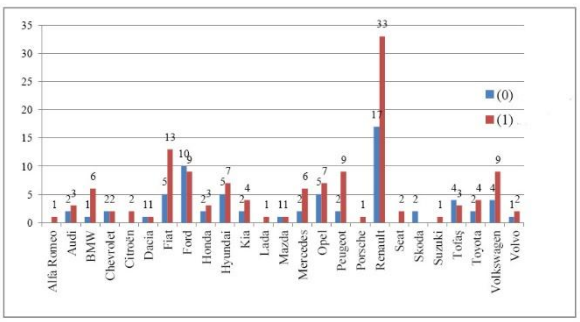


Figure 2. Summary of Data by Car Brand

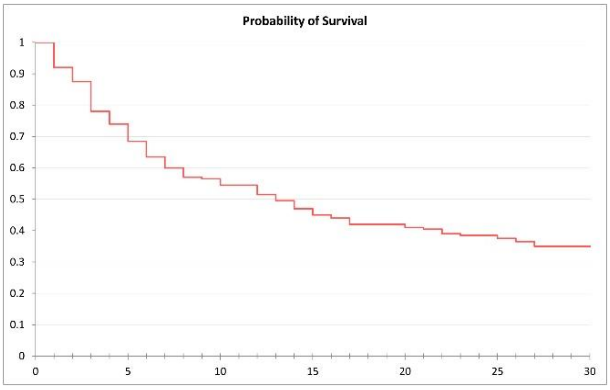


Figure 3. Nonparametric Survival Function

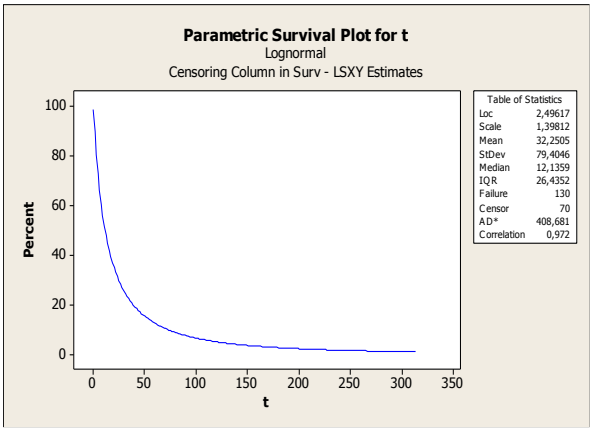


Figure 4. Parametric Lognormal Survival Function

4. ANALYZING THE DATA

In this section we report our results on analyzing the data by survival and regression analyses. Figure 3 depicts the empirical and nonparametric survival function based on 200 observations.

As a result of the work carried out to determine the distribution analysis on the dataset with Minitab, it is found that the lognormal distribution is the most suitable distribution for the available data. Therefore, parametric analysis was performed by using lognormal distribution. *Figure **4*** depicts lognormal survival function. *Figure **5*** shows nonparametric survival function generated by Kaplan-Meier method.

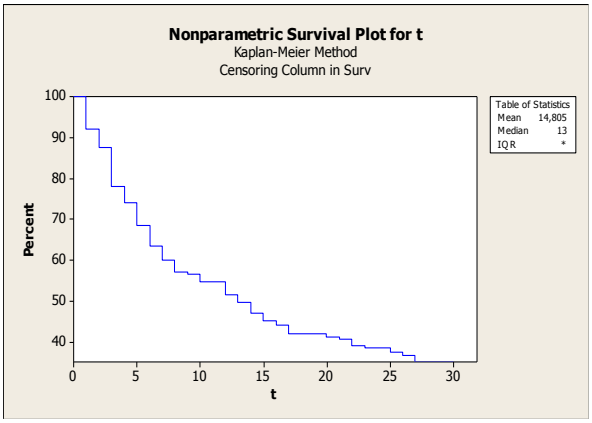
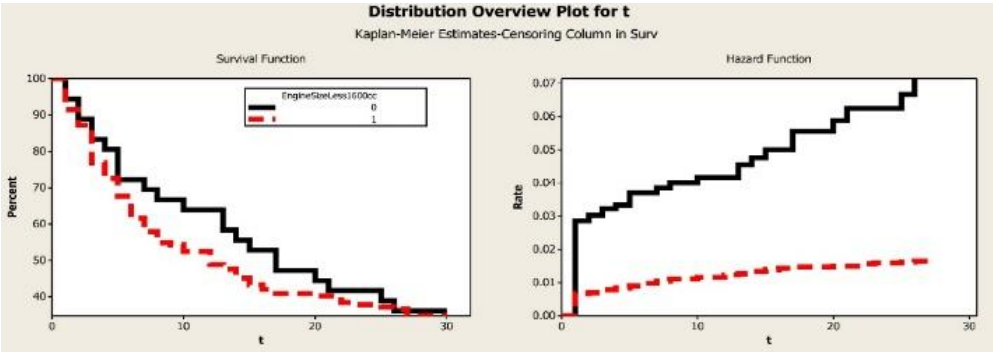


Figure 5. Nonparametric Survival Function Computed by Kaplan-Meier Method



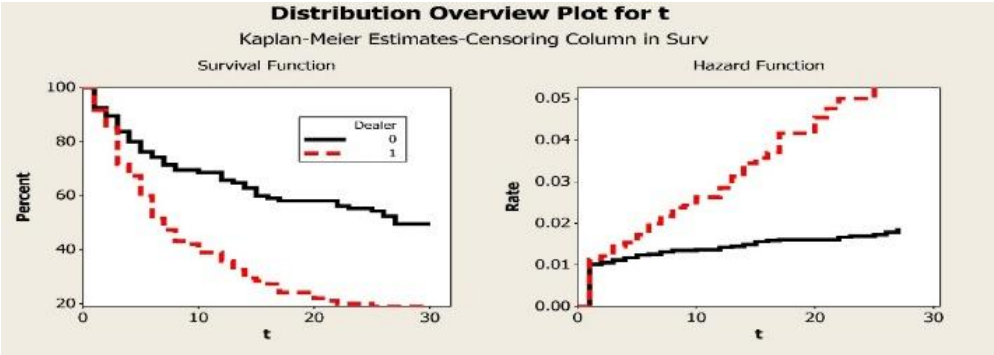


Figure 6. Nonparametric Survival Functions by Engine Size and Seller Type

Table 1. Summary of Regression Models

Model #	Model	R-Sq	F-Val	p Val
I	t = - 0.00590 ViewRatio + 0.0102 Year + 0.000008 KM - 3.3 Gasoline - 5.7 LPG - 2.1 Diesel - 1.16 ManShift - 0.12 AutoShift + 4.30 EngSizeLess1200cc + 4.79 EngSize1201-1400cc + 3.23 EngSize1401-1600cc + 7.48 EngSize1601-1800cc + 8.23 EngSize1801-2000cc + 17.3 EngSize2001-2500cc + 9.13 EngSize2501-3000cc - 1.70 Hatchback_5 + 0.61 Saloon + 3.50 Station Wagon - 9.32 Dealer	0.712	23.55	0.0001
II	t = 1.47 Gasoline - 0.06 LPG + 1.74 Diesel + 2.54 ManShift + 2.85 AutoShift + 18.0 EngSizeLess1200cc + 15.1 EngSize1201-1400cc + 12.5 EngSize1401-1600cc + 15.2 EngSize1601-1800cc + 16.8 EngSize1801-2000cc + 23.8 EngSize2001-2500cc + 15.5 EngSize2501-3000cc - 0.6 EngSize3001-3500cc - 2.39 Hatchback 5 - 1.21 Saloon + 1.59 Station Wagon	0.6537	21.7	0.0001
III	t = - 0.00544 ViewRatio + 0.048 Year + 0.000014 KM + 23.3 Gasoline + 20.6 LPG + 25.0 Diesel - 9.08 Dealer - 4.35 EngSizeLess1200cc - 3.97 EngSize1201-1400cc - 5.00 EngSize1401-1600cc	0.6946	43.2	0.00001
IV	t = - 0.00366 ViewRatio + 0.092 Year - 0.000009 KM - 163 Gasoline - 162 LPG - 161 Diesel + 0.07 ManShift - 0.64 AutoShift - 7.94 EngSizeLess1200cc - 12.5 EngSize1201-1400cc - 10.6 EngSize1401-1600cc - 5.85 EngSize1601-1800cc - 6.63 EngSize1801-2000cc + 5.3 EngSize2001-2500cc - 4.24 EngSize2501-3000cc - 14.2 EngSize3001-3500cc - 2.72 Dealer	0.6515	12.43	0.0001
V	t = 9.64 Gasoline + 9.74 LPG + 12.0 Diesel + 0.57 ManShift - 0.08 AutoShift - 0.95 EngSizeLess1200cc - 4.95 EngSize1201-1400- 2.98 EngSize1401-1600 + 1.36 EngSize1601-1800-0.91 EngSize1801-2000 + 12.4 EngSize2001-2500+ 3.05 EngSize2501-3000 - 3.28 Dealer + 2.14 Station Wagon + 2.09 Hatchback 5 + 1.67 Saloon	0.6313	12.2	0.0001
VI	t = - 0.00390 ViewRatio - 0.105 Year - 0.000008 KM + 15.4 Gasoline + 16.3 LPG + 17.5 Diesel - 2.49 EngSizeLess1200cc - 6.81 EngSize1201-1400cc - 4.91 EngSize1401-1600cc - 2.70 Dealer	0.6354	20.91	0.00001

We also analyzed the effects of the listing (car) features on the survival function. *Figure 6* depicts the effect of engine size (whether less than 1600 cc or not) in the top plot and the type of the seller (whether dealer or not) in the bottom plot. The cars with less than 1600 cc are sold quicker than the larger engine sizes. Somehow the car listings posted by dealers were removed earlier than listings by owners which may indicate faster sales by the dealers.

The effects of the engine types on the survival functions are depicted in *Figure 7*. The top plot shows the effect of diesel engines that first 20 days non-diesel cars are sold quickly. In the last 10 days of the period, diesel cars are sold much faster. Notice that there are also cars with LPG engines. We see very little difference in survival functions of gasoline vs. non-gasoline engines at the bottom plot of *Figure 7*.

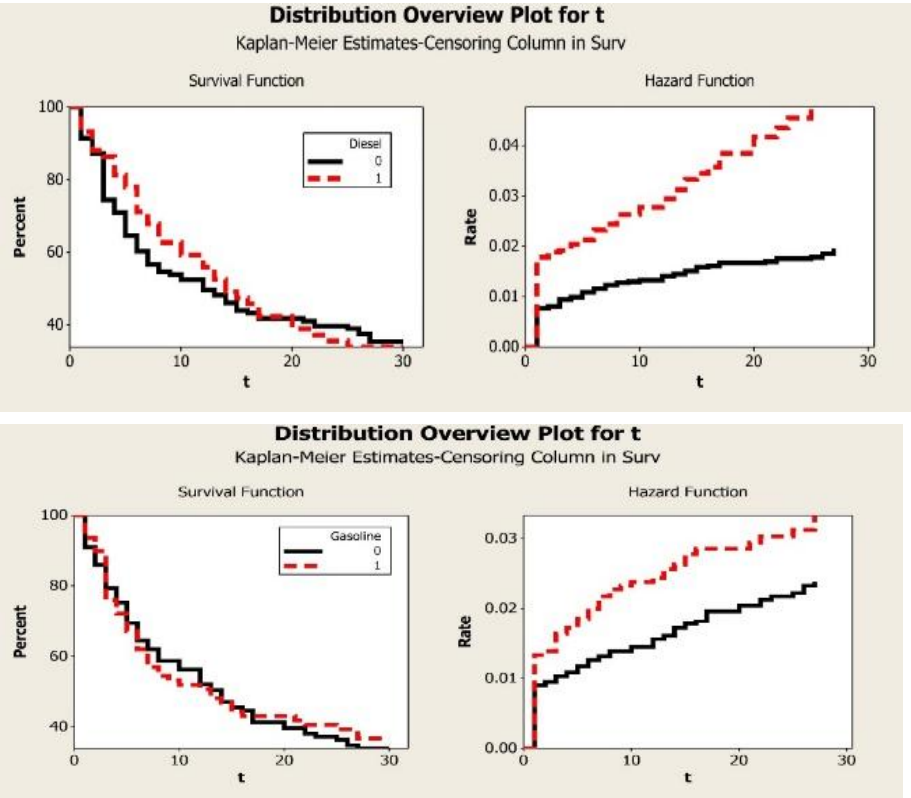


Figure 7. Nonparametric Survival Functions by Engine Types

4.1. Regression Models

Any analysis without proper models to determine the important factors that affect the time to death (in our case time to sell the car) will not be complete (Kleinbaum & Klein, 2005). For this reason we conducted multivariate regression analysis by creating related dummy (indicator) variables that indicate whether a particular car has certain features or not. Notice that car listings already have data of some car features as continuous variables such as mileage (km), price and the number of total listing views.

The results of multivariate regression models are reported in **Error! Reference source not found.** Models I, II, and III comprise the regression models based on full data i.e. they include also the censored observation at 30 days. Models IV, V, and VI, on the other had were constructed by only using the data from sold cars within 30 days. Regression analysis of the all automobiles sold and unsold are compared with regression analysis of the only sold car. Notice that intercept was not fitted in any of the models reported in **Error! Reference source not found.**

The variable `ViewRatio` is the average number of views per day for a given listing. It is calculated by total number of views divided by *t* (i.e. time to sell the car or the censor time which is 30 at most). The variable `KM` represents the mileage in kilometers. The variable `Year` presents the age of the car in years. The remaining variables in Table 1 are dummy variables to indicate whether the cars have the corresponding features or not. For example, the variable `Diesel` represents whether the engine type is diesel or not.

R-Squared values are reasonable for all the regression models and *p-values* that they are significant. The partial regression coefficients are in line with the expectations. For example, the variable `ViewRatio` has negative signs in the models which indicate the more a car listing is seen on average per day, the sooner it will be sold. Partial regression coefficients of `Diesel` are higher than the other engine types which indicate that it takes longer time to sell diesel cars on the average. As the sign of the partial regression coefficients of `Dealer` are negative, it takes less time to sell the cars listed by the dealers. The variable `KM` has both positive and negative signs in these six models presented in Table 1 which realistically determines that `KM` may not be significant at all. Again small size engines have negative signs which indicates that it is easier to sell cars with smaller engines.

5. DISCUSSION AND CONCLUSION

We presented the statistical analysis results of data collected from an e-commerce site about car listings. We successfully implemented methods from survival analysis to analyze such data. We then implemented regression models to analyze the factors that affect the time to sell the car (or remove the car listing). Survival functions and the regression models agree with each other's outcomes.

Since price data are also collected in our study, we can easily determine the price elasticity of the used cars as prices may vary from day to day for a given car listing. We can also construct classification models to predict that a certain vehicle will be sold within a specified time period or not. We plan to conduct such studies in out forthcoming paper.

REFERENCES

[1] Asilkan Ö., İkinci El Otomobillerin Gelecekteki Fiyatlarının Yapay Sinir Ağları ile Tahmin Edilmesi, Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi Y.2009, C.14, S.2 S.375-391.

[2] Dünya, http://www.dunya.com/e-ticaret-hacmi-yuzde-20-artti-187007h.htm, Accessed on June 6, 2013.

[3] Esen E., Sağ Kalım Analizinde Parametre Tahmin Problemlerine Katkılar, Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, Yüksek Lisans Tezi, Samsun, 2005

[4] Gökçegöz, F., "*Bir İnternet Sitesindeki İkinci El Otomobil İlanları İle Sağ Kalım Analizi Uygulaması*", Senior Thesis, Dept. of Industrial Engineering, Sakarya University. 2012

[5] Kaplan, E., and Meier, P., "Nonparametric Estimation From Incomplete Observations," J. Amer. Statist. Assoc. 53, 457-481, 1958.

[6] Kleinbaum, D.G. and Klein, M., "Survival Analysis A Self Learning Text,", 2nd Ed. Springer, 2005.

[7] Nelson W., Applied Life Data Analysis Canada: John Wiley & Sons, 1982.

[8] Wang, L., Puskorius, G., Nance, B., and Salmeen, I., "Application of Survival Analysis for Modeling the Effects of Vehicle Features on Days-on-Lot". In Proceedings of Joint Statistical Meeting, pp. 3585-3590, NY, USA, 2002.

[9] Wikipedia, en.wikipedia.org/wiki/Used_car, Accessed on June 6, 2013

ISSN 2303-4521
PERIODICALS OF ENGINEERING AND NATURAL SCIENCES
Vol. 1 No. 2 (2013)
Available online at: http://pen.ius.edu.ba

Development of PIM Components for Robot Surgery

Dong Yong Park
Department of Mechanical Engineering POSTECH, 77 Cheongam-ro Pohang, South Korea

Fehim findik
Division of Engineering and Natural Sciences International University of Sarajevo, Sarajevo Bosnia-Herzegovina

Seong Jin Park
Department of Mechanical Engineering POSTECH, 77 Cheongam-ro Pohang, South Korea

Abstract
In this study, the micro forcep of end-effector for robot surgery was produced by powder injection molding. The 17-4PH stainless steel powder and binder system based on a wax-polymer were mixed to fabricate the feedstock. The optimum solid loading (vol. %) was determined by the torque rheometer experiments. After injection molding, debinding and sintering were carried out, final product having small and regular patterns was produced.

Keywords: μ-PIM, end-effector, and robot surgery

1. Introduction

Micro powder injection molding (μ-PIM) technology combined by plastic injection molding and conventional powder metallurgy has been known as a near net shaping technology suitable for the fabrication of micro-components [1]. This technology has many advantages including shape complexity, tight tolerances, and material selection of metals or ceramics. Once desired materials, mold geometries, and process parameters are decided, PIM is an appropriate process for the mass production. The PIM process consists of four steps; (i) mixing – producing the pelletized feedstock of the powder and organic binders, (ii) molding – injecting the feedstock melt into the mold cavity, similar with thermoplastics; (iii) debinding – extracting or removing the organic binders out of injection molded part via solvents or the thermal energy, (iv) sintering – densifying the debound part from the low initial density to the high final density, close to the full density [2]. Figure 1 shows a schematic diagram of the PIM process.

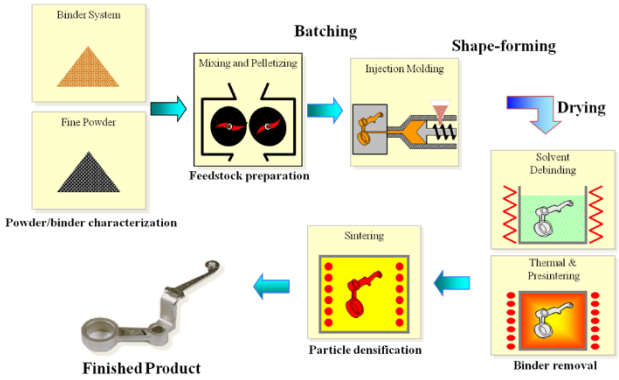


Figure 1. schematic diagram of powder injection molding [3].

In this regards, μ-PIM technology is employed to produce medical devices such as an endo-tip for the dental application and an end-effector for the robot surgery. Figure 2 shows the endo-tip for dental application made by CetaTech Inc.

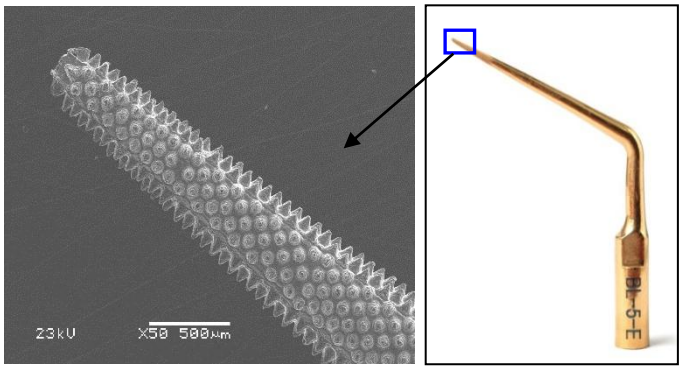


Figure 2. Dental application (Courtesy by CetaTech Inc., Korea).

2. Experimental Procedures

The water-atomized stainless steel powders of type 17-4PH were prepared with PF-15F (8.30 μm) produced by Atmix Inc. The particle characteristics of SUS17-4PH are given in Table 1. The powder was mixed with wax-polymer binder system including wax, polypropylene (PP), polyethylene (PE), and stearic acid (SA). The mechanical properties of the binder system is summarized in Table 2. The optimum solid loading (vol. %) was determined by the torque rheometer experiment. Mixing was carried out to produce feedstock by the twin-extruder type of mixer (CetaTech Inc.). All samples were debound to remove the binder. After solvent debinding, all samples were thermally debound and pre-sintered via following thermal cycles; temperature has been increased by

DOI: 10.21533/pen.v1i2.20

Table 1. The characteristics of powder

Powder	particle size (μm)			Distribution slope parameter (S_w)	Apparent density (g/cm^3)	Tap density (g/cm^3)	True density (g/cm^3)
	D_{10}	D_{50}	D_{90}				
PF-15F	3.22	8.30	19.47	3.28	3.08	4.20	7.75

Table 2 *Material properties of binder ingredients*

	Wax	PP	PE	SA
Density (g/cm^3)	0.90	0.90	0.92	0.94
Melting point (˚C)	42 - 62	110-150	60 - 130	74 - 83
Decomposition temperature (˚C)	180 - 320	350 - 470	420 - 480	263 - 306

ramping from 30 °C to 900 °C at 2 °C/min with intermediate 2-hour holds at 250 °C, 450 °C, and 700 °C using the tube furnace. The injection molded samples were debound at 60 °C for 10 hours in a *n*-Hexane solution. The debound samples were sintered in a H_2 atmosphere. In order to predict the densification behavior during sintering, dilatometry experiments were conducted. Figure 3 shows the thermal history of dilatometry experiments. The densification behavior was analyzed.

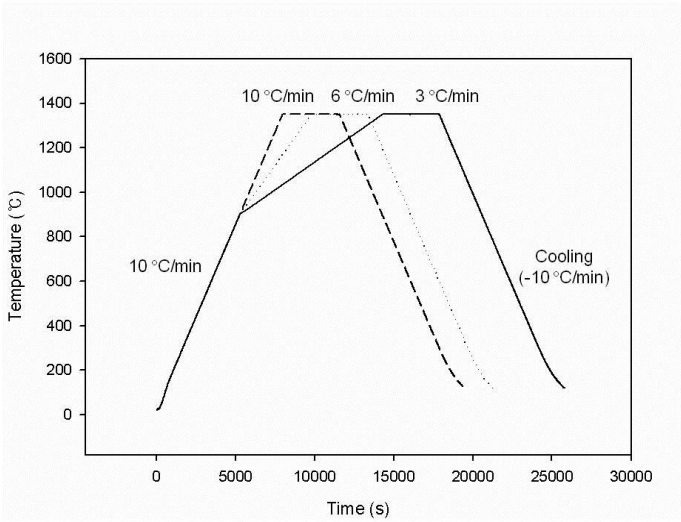


Figure 3. Thermal history of dilatometry experiments.

3. Results and Discussions

The optimum solid loading (vol. %) which is a balanced mixture of powder and binder can determine success or failure of subsequent PIM processes. To determine the optimum solid load, torque rheometer experiments were carried out. The torque was measured by adding 1 % of solid loading for each step at the temperature of 150 °C. The critical solid loading where the particles are tightly packed and remaining space filled with binder was determined at the solid loading of 63 % as shown in Figure 4. The optimum solid loading is considered as approximately 2 – 5 lower value than critical solid loading. In this study, we determined the solid loading set by 59 vol. %.

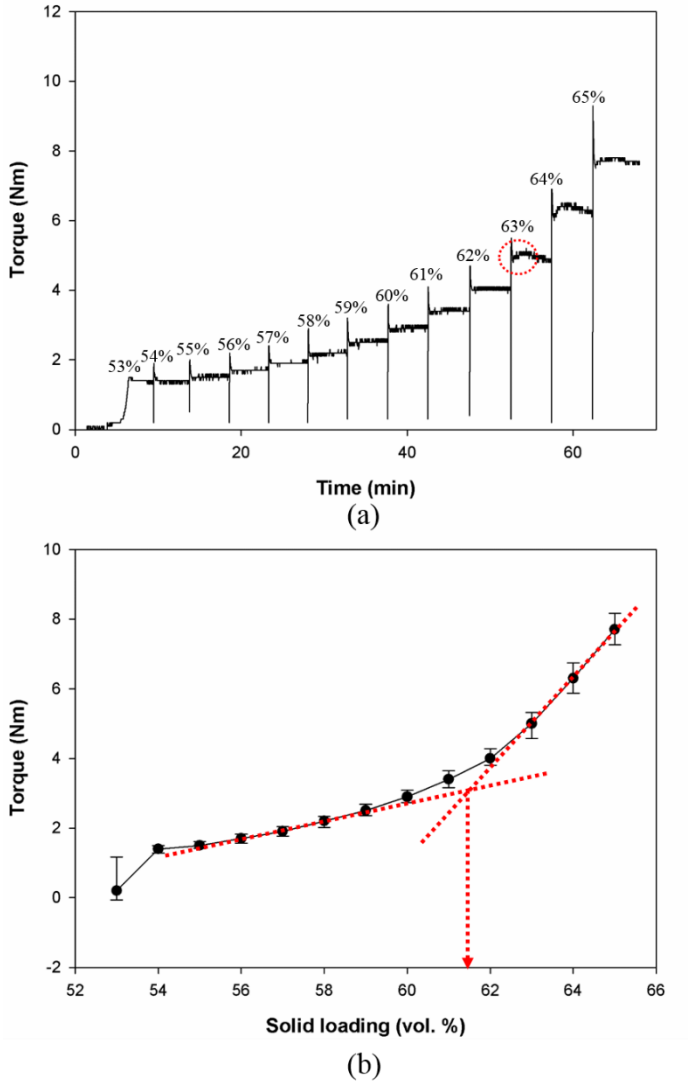


Figure 4. Mixing torque, (a) as a function of time, and (b) as a function of solid loading.

A forcep for robot surgery was designed using Solid Works. As shown in Figure 5, end-effector for robot surgery has the small and complex shape. Small components with complex geometries can be more economically produced by PIM technique. In order to

design the mold for injection molding, the shrinkage during sintering has to be considered to obtain the precise dimension of final components.

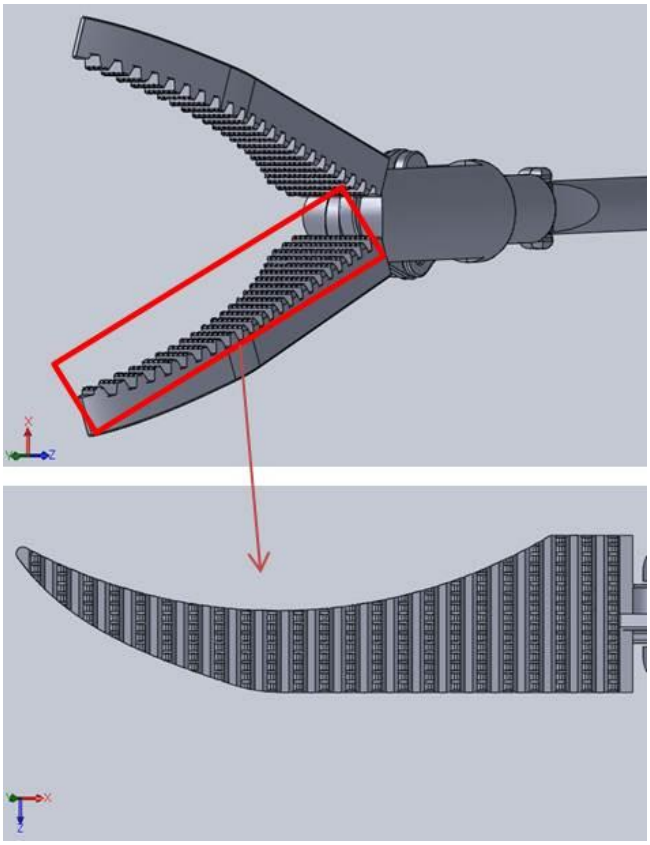


Figure 5. designed end-effector for robot surgery.

In this regard, dilatometry experiments were conducted to measure the *in situ* shrinkage during sintering. Figure 6 shows the shrinkage behavior for given sintering conditions.

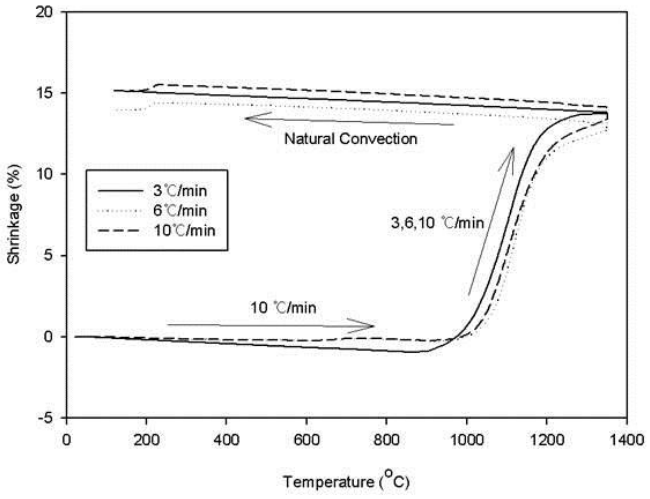


Figure 6. Shrinkage behavior for given sintering conditions.

The shrinkage, after sintering, is approximately 15 %. The feedstock was injection molded into the designed mold. Figure 7 shows the green, brown, and sintered parts for end-effector.

Figure 7. Green, brown, and sintered parts for the forcep of end-effector.

The injection molded samples (green part) were debound at 60 °C for 10 hours in a *n*-Hexane solution to remove the binder. After solvent debinding, thermal debinding and pre-sintering were carried out. Finally, debound samples (brown part) were sintered in H_2 atmosphere. The dimension of sintered components were shrunk as expected.

Conclusions

The end-effector for robot surgery was produced by PIM technique. Considering that the forcep of end-effector has the complicated shape, the PIM is a suitable method to produce the medical device with complex geometry.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2011-0030075).

References

[1] German, R.M. and A. Bose, "*Injection molding of metals and ceramics.*" Metal Powder Industries Federation1997.

[2] Liu, L., N.H. Loh, B.Y. Tay, and S.B. Tor, "Microstructure evolution of 316L stainless steel micro components prepared by micro powder injection molding." *Powder Technology*, vol. 206, no. 3, pp. 246-251, 2011.

[3] Park, D.Y., S.W. Lee, S.J. Park, Y.-S. Kwon, and I. Otsuka, "Effects of Particle Sizes on Sintering Behavior of 316L Stainless Steel Powder." *Metallurgical and Materials Transactions A*, vol. 44, no. 3, pp. 1508-18, 2013.

ISSN 2303-4521

PERIODICALS OF ENGINEERING AND NATURAL SCIENCES
Vol. 1 No. 2 (2013)
Available online at: http://pen.ius.edu.ba

Mechanical Properties and Corrosion Behaviors in 3.5% NaCl Solution of Grade-A and Dual-Phase Steels Welded by FCAW

Hüseyin Uzun

Sakarya University, Technology Faculty, Metallurgy and Materials Department, Esentepe Campus, Sakarya, Turkey

huzun@sakarya.edu.tr

Emrah Önal

Yüksekova Technical and Vocational High School, Van yolu üzeri İpekyolu caddesi Belediye Karşısı, Hakkari, Turkey

eonal@hotmail.com

Abstract:

The first aim of this study is to demonstrate the transformation of Grade-A steel into dual-phase steel by heat-treatment at the chosen 730°C and 800°C intercritical temperature for 60 min followed by cold water quenching. The second aim of this study is to investigate the mechanical properties and the corrosion behaviours in 3.5% NaCl solution of Grade-A and dual-phase steels welded by flux-cored arc welding (FCAW). It was carried out the mass loss measurements and investigated the micro and macrostructures of welding zone using by optical microscope.

The present study has demonstrated that Grade-A ship steel consisting of ferrite and pearlite can be successfully transformed to the dual-phase steel at the chosen intercritical temperatures. The microstructure of transformed dual-phase steel consists of island of martensite in a ferrite matrix. The Grade-A and transformed dual-phase steels were successfully joined by FCAW. The mechanical results showed that the tensile strength of the transformed dual-phase steel were higher, but lower elongation than that of the Grade-A steel. The results from the immersion test in 3.5% NaCl solution showed that the corrosion behaviours of the dual-phase and Grade-A steels depend upon the morphology of the phase constituents. The corrosion rate of the transformed dual-phase steel with ferrite-martensite structure was lower than that of the Grade-A steel with ferrit- pearlite structure. The corrosion has preferential started at fusion line and progressed towards the heat affected zone (HAZ) of the welded dual-phase and Grade-A steels.

Keywords: Grade-A steel, flux-cored arc welding (FCAW), corrosion behaviour of dual-phase steel, 3.5% NaCl solution.

1. Introduction

The shipbuilding industry has promoted high performance of ships and improved productivity in construction. It is very important to develop new steel with high the specific strength, good weldability and high corrosion resistance for construction of ship having lower construction weight, improving the fuel consumption economy and allowing greater quantities and sizes of loads. So far, the Grade-A steel was mostly employed to construct a great region of ship frame due to the being cheaper, having good weldability and high mechanical properties.

It was reported by Hayat et al. that Grade-A ship steel has been employed to transform into dual-phase (DP) steel accomplished by intercritically annealing at intercritical temperatures ranges of 730 ^{o}C and 800^{o}C, and following by water quenching (Hayat, 2012). Thus, the microstructure of Grade-A consists of ferrite and pearlite, while the transformed dual-phase steels consist of ferrite and martensite. Mechanical properties of the transformed dual-phase steels are higher than the Grade-A steel. Therefore, the transformed dual-phase steels are promising candidate materials for the ship building industry because of their good weldability, high strength, low yield-to-tensile strength ratio and high strength-to-weight ratio (Hayat, 2012). It was investigated the weldability of transformed dual-phase steels joined by flux-cored arc welding (FCAW) process at the same work.

Nowadays, it was reported the corrosion performance of dual-phase steel embedded in concrete by Ismail et al. (Ismail, 2009) and Keleştemur et al. (Keleştemur, 2009), the mechanical properties and corrosion susceptibility of dual-

DOI: 10.21533/pen.v1i2.21

phase steel in concrete by Trejo et al. (Trejo, 1994), microstructural influence on the electrochemical corrosion behaviour steels in 3.5% NaCl solution by Sarkar et al. (Sarkar, 2005) and mechanical and corrosion behaviours of plain low carbon dual-phase steels by Bhagavathi et al. (Bhagavathi, 2011). It was no any report the corrosion behaviours in 3.5% NaCl solution of the transformed dual-phase steels welded by FCAW.

The first aim of this study is to demonstrate the transformation of Grade-A steel into dual-phase steel by heat-treatment at the chosen 730°C and 800°C intercritical temperature for 60 min followed by cold water quenching. The second aim of this study is to investigate the mechanical properties and the corrosion behaviours in 3.5% NaCl solution of Grade-A and dual-phase steels welded by FCAW. It was carried out the mass loss measurements and investigated the micro and macrostructures of welding zone using by optical microscope. The pitting corrosion mechanism was described and compared with the Grade-A steel and the transformed dual-phase steel joints.

2. Experimental Procedure

The as-received hot-rolled Grade-A ship steel plates was employed to transform to the dual-phase steel. Table 1 shows the chemical compositions and mechanical properties of the Grade-A ship steel.

The Ac_1 and Ac_3 temperature limits (711-845ºC) were computed using the Andrews formula from the chemical composition of the Grade-A steel (Andrews, 1965). It was carried out the transformation of Grade-A steel into dual-phase steel by heat-treatment at the chosen relatively lower 730 °C and upper 800 °C intercritical temperatures for 60 min followed by cold water quenching. 730DP and 800DP represent the base metal at different annealing temperatures of 730 °C and 800 °C, respectively. The martensite volume fractions (MVF) of the transformed dual-phase steels are calculated by using the Clemex Vision Lite image analysis program by Nikon Eclipse L150A optical microscope.

Similar the Grade-A and the transformed dual-phase steel plates (400x150x8 mm) were joined on three different passes in the V-groove butt form by FCAW process with AWS/ASME SFA-5.29 E81 T1-Ni1 flux-cored welding wire. The schematic the single V-groove butt joint with ceramic backing is illustrated in Fig. 1.

The corrosion behaviours of Grade A and dual-phase steels non-welded and welded by flux-cored arc welding (FCAW) were studied by using immersion test in 3.5% NaCl solution. The ASTM established recommended procedure for immersion tests as covered by designation G-31 (ASTM, 2004) was employed. They were weighed before the start of the test and after completion of the test. The non-welded and welded specimens were immersed in 3.5% NaCl solution for 7, 15, 30, 45, 60, 75 and 90 days at room temperature. It was kept the corrosion solution with a constant pH value of 8 during the immersion test. The corrosion rate was calculated by weight loss method. The weight loss is converted to an average corrosion rate using the following formula:

$$CR = \frac{K \times W}{A \times T \times D} \quad (1)$$

Where CR is average corrosion rate (mm/year), K is constant (8.76×10^4), W is weight lost by the specimen during the test (g), A is = total surface area of the test specimen (cm^2), T is duration of exposure (h) and D is density of the specimen material (g/cm^3).

The tensile tests of non-welded and welded specimens were prepared according to TS EN ISO 6892-1 and TS EN ISO 4136:2011 specifications, respectively. The Vickers microhardness measurements were performed using a load of 200 g and a dwell time of 10 s.

3. Results And Discussion

3.1. Microstructures of Grade-A steel, dual-phase steel and weld metals

The microstructure of the as-received Grade-A steel shown in Fig. 2 reveals that the microstructure consists of uniformly distributed pearlite (dark) colonies in an equiaxed ferrite matrix (light). Figure 3 shows the microstructures of the base dual-phase steels which are transformed from the Grade-A steel. The microstructure of base dual-phase steel developed following intercritical annealing at lower temperature of 730°C is composed of light gray coloured uniform fibrous ferrite and dark coloured martensite, which is commonly termed as a fibrous structure (Fig. 3 (a)). The microstructure of base dual-phase steel developed following intercritical annealing at higher temperature of 800°C consists of islands of martensite in a ferrite matrix (Fig. 3 (b)). It is concluded that the higher and the lower intercritical annealing temperatures are influenced on the martensite morphology and distribution of the phase constituents. Similar observations were also reported in the related literatures (Sarkar, 2005 and Bhagavathi, 2011).

The image analysis exhibits that the martensite volume fractions (MVF) in the base dual-phase steels coded 730DP and 800DP are 16±5 % and 55±6 %, respectively. It shows that the MVF of the dual-phase steels increases with increasing the intercritical annealing temperatures. Similar observations were also made in the literature (Hayat, 2012).

The similar Grade-A and transformed dual-phase steels were successfully joined by FCAW process. The representative macrostructure of the weld was shown in Fig. 4. It was measured the weld face width of 12-14 mm and the weld face height of 5-7 mm.

Fig. 5 shows the microstructure of weld metal in the Grade- A and dual-phase steels joined by FCAW. It was confirmed by TEM analysis that the microstructure of weld metal was consisted of the primary ferrite phase, grain boundary ferrite (allotriomorphic ferrite), polygonal ferrite, Widmanstatten ferrite, acicular ferrite and martensite at a lower quantity (Hayat, 2012).

3.2. Mechanical properties

The microhardness profiles along the top and bottom in cross sectional horizontal to the weld direction in the similar Grade - A and dual-phase steels with different annealing temperatures of 730 °C and 800 °C welded by FCAW are shown in Fig. 6. The hardness profiles of the dual phase steels reported by Hayat et al. are also shown in the same Figure for comparison purposes (Hayat, 2012). The lowest hardness value was found in Grade-A steel, approximately 142±5 Hv. The hardness values in dual-phase steels with the tempering temperature of 730 ^{o}C and 800 ^{o}C were 200 ± 5 Hv and 258 ± 5, respectively.

The tensile test results of Grade-A and dual-phase steels welded by FCAW are presented in Table 2. The lowest tensile strength is exhibited by the Grade-A steel as compared the dual-phase steels. The tensile strength of dual-phase steels increases and the percentage of elongation decreases with increasing martensite volume fraction.

3.3. Corrosion properties

Fig. 7 shows the macrostructures of weld region of Grade-A and dual-phase steels welded by FCAW after immersion corrosion test in 3.5% NaCl solution. It is estimated that the galvanic couple triggered pitting reaction on the corrosion of Grade-A and dual-phase steel joints. It is observed that the corrosion failure at the weld region preferentially starts in the partially melted zone between weld metal and heat affected zone (HAZ), and prolonged towards the HAZ. Thus the partially melted zone may corrode faster than the weld metal and/or HAZ or HAZ may corrode faster than the weld metal. This is attributed that two galvanic couples are formed either between the weld metal and the partially melted zone or between HAZ and the partially melted zone. In galvanic couple between the weld metal and the partially melted zone, the weld metal acts as cathode and the partially melted zone acts anode. In galvanic couple between HAZ and the partially melted zone, HAZ acts as cathode and the partially melted zone acts anode.

In the corrosion of Grade-A and dual-phase steels in 3.5% NaCl solution, following corrosion reactions occur simultaneously.

Anodic reaction: $Fe \rightarrow Fe^{2+} + 2e^{-}$ (2)

Cathodic reaction: $Fe^{+2} + 2OH^{-} \rightarrow Fe(OH)_2$ (3)

In 3.5% NaCl solution, sodium and chlorine ions are decomposed and then reacted with ferrous ions, thus the pH of the solution from a value of 8 to 5 decreases. Therefore, the corrosion failure of steels increases. In the present study, the pH of 3.5% NaCl solution is arranged the constant value of pH 8 in order to evaluate the performance of steels.

$$NaCl \rightarrow Na^{+} + Cl^{-} \quad (4)$$

$$Fe^{+2} + 2Cl^{-} \rightarrow FeCl_2 \quad (5)$$

$$Fe^{+3} + 3Cl^{-} \rightarrow FeCl_3 \quad (6)$$

$FeCl_2$ and $FeCl_3$ compounds occurred at the end of the reactions are hydrolyzed. As a result of the reactions, hydrochloric acid (HCl) occurs.

$$FeCl_2 + 2H_2O \rightarrow Fe(OH)_2 + 2HCl \quad (7)$$

$$FeCl_3 + 2H_2O \rightarrow Fe(OH)_3 + 3HCl \quad (8)$$

The ferrous hydroxide is oxidized to ferric hydroxide ($Fe(OH)_3$) known as corrosion product.

$$Fe(OH)_2 + ½\ H_2O + ¼\ O_2 \rightarrow Fe(OH)_3 \quad (9)$$

Table 3 shows the results of the average corrosion rates of similar the Grade-A and the dual-phase steel joints obtained from immersion test in 3.5% NaCl solution. It can be seen from Table 3 that the average corrosion rate of similar Grade - A steel joint is higher than that for the similar dual-phase steel joints. This is attributed to the number of micro and macro corrosion couples. It was observed two micro corrosion couples and two macro corrosion couples at the Grade-A steel joint. Macro corrosion couples are formed (1) between the weld metal (cathode) and the partially melted zone (anode), (2) between HAZ (cathode) and the partially melted zone (anode). Micro corrosion couples are formed (1) between lamellae cementit in pearlite (cathode) and eutectoid ferrite in pearlite (anode), (2) between lamellae of pearlite (cathode) and proeutectoid ferrite (anode). Fig. 8 shows the schematic galvanic corrosion mechanisms of similar Grade-A steel joints.

On the other hand, it was observed a micro corrosion couple and two macro corrosion couples at the dual-phase steel joints. Macro corrosion couples are formed (1) between the weld metal (cathode) and the partially melted zone (anode), (2) between HAZ (cathode) and the partially melted zone (anode). Micro corrosion couple is formed between ferrite (anode) and martensite (cathode). Fig. 9 illustrates the schematic galvanic corrosion mechanisms of similar dual-phase steel joints.

It can be seen from Table 3 that the average corrosion rate of similar dual-phase steel joint coded 730DP developed following intercritical annealing at lower temperature of 730^{o}C is higher than that for the similar dual-phase steel joints coded 800DP developed following intercritical annealing at higher temperature of 800^{o}C. This is attributed the martensite morphology and martensite volume fraction (MVF). The MVF of dual-phase steel coded 730DP is higher than that for the dual-phase steel coded 800DP.

It has been reported (Sarkar, 2005) that island-like morphology of martensite gave better corrosion resistance as compared to the network form surrounding the ferrite grains. In present study, martensite phase has similar island-like morphology in the matrix of ferrite in dual phase steel developed following intercritical annealing at higher temperature of 800 ^{o}C. The microstructure of dual-phase steel coded 730DP developed following intercritical annealing at lower temperature of 730 °C is composed of uniform fibrous ferrite and martensite, which is commonly termed as a fibrous structure. Thus, the lower corrosion resistance is observed on the dual phase steel joint coded 730DP as compared to dual phase steel coded 800DP, in view of the structural and morphological considerations.

4. Conclusions

Grade-A ship steel was successfully transformed to the dual-phase steel at the chosen intercritical temperatures. The Grade-A and transformed dual-phase steels were also successfully welded by FCAW. The microstructure, microhardness, tensile properties and corrosion behaviours in 3.5% NaCl solution of joints have been studied in the present work. Following conclusions are drawn:

1- The microstructure incestigation shows that while the microstructure of Grade-A steel consists of proeutectoid ferrite and pearlite, the microstructure of transformed dual-phase steels consists of ferrite and martensite phases.

2- The microstructure of dual-phase steel developed following intercritical annealing at lower temperature of 730°C is composed of uniform fibrous ferrite and martensite. The microstructure of dual-phase steel developed following intercritical annealing at higher temperature of 800°C consists of islands of martensite in a ferrite matrix.

3- The tensile test results show that the tensile strength of Grade-A steel joint is lower than that of the transformed dual phase steel joints.

4- The corrosion resistance obtained from immersion tests of Grade-A steel joint were less than that for the dual phase steel joint. It is assumed two reasons. Firstly, the martensite formed in dual phase steel is structurally and compositionally closer to ferrite matrix phase. Therefore the galvanic couple formed between ferrite–martensite is weaker as compared to Grade-A steel wherein pearlite which consists of ferrite and cementite lamellae is structurally and compositionally inhomogeneous. Secondly, the number of micro and macro corrosion couples different from Grade-A steel joint and dual phase steel joint. It could be observed two macro and micro corrosion couples on the Grade-A steel joint (1) between the weld metal and the partially melted zone, (2) between HAZ and the partially melted zone, (3) between lamellae cementit in pearlite and eutectoid ferrite in pearlite, (4) between lamellae of pearlite and proeutectoid ferrite. Ont the other hand, it could be observed two macro and one micro corrosion couples on the dual phase steel joint (1) between the weld metal and the partially melted zone, (2) between HAZ and the partially melted zone, (3) between ferrite and martensite island.

5- The corrosion resistance of dual-phase steel increases depending on the increase in the volume fraction of martensite. In addition, the corrosion resistance of the dual-phase steel increases with intercritical annealing tempering heat treatment and the highest corrosion resistance take place in dual phase steel tempered at 800 °C. This is attributed to the depending upon the volume fraction and morphology of the phase constituents. The dual phase steel tempered at higher temperature of 800°C, which consists of island-like morphology in the matrix of ferrite, gives rise better corrosion resistance property than the dual phase steel tempered at lower temperature of 730°C, which consists of fibrous ferrite and martensite.

Acknowledgment

This work was supported by Sakarya University Scientific Research Foundation (Project number: BAP-2010-05-06-011).

References

[1] Hayat, F. and Uzun, H. (2012), ”Microstructural and Mechanical Properties of Dual-phase Steel Welded by GMAW with Solid and Flux-cored Welding Wires” *International Journal of Materials Research*, Vol. 103 (7), 828-837.

[2] Mohammad, I. (October 2010), “Corrosion Performance of Dual-phase Steel Embedded in Concrete”, *The Arabian Journal for Science and Engineering*, Vol. 35 (2B), 81-90.

[3]Keleştemus, O., Aksoy, M. and Yıldız, S.(2009),”Corrosion Behaviour of Tempered Dual-phase Steel Embedded in Concrete”, *International Journal of Minerals, Metallurgy and Materials*, Vol.16 (1),43-50.

[4] Trejo, D., Monteiro, P., Thomas, G. and Wang, X. (1994), “Mechanical Properties and Corrosion Susceptibility of Dual-phase Steel in Concrete”, *Cement and Concrete Research*, Vol. 24 (7), 1245–1254.

[5] Sarkar, P.P., Kumar, P., Manas Kumar Manna, Chakraborti, P.C. (2005), “Microstructural Influence on the Electrochemical Corrosion Behaviour of Dual-phase Steels in 3.5% NaCl Solution”, *Materials Letters,* Vo. 59, 2488 – 2491.

[6] Lakshmana, R. B, Chaudhari, G.P. and Nath, S.K. (2011), “Mechanical and Corrosion Behaviour of Plain Low Carbon Dual-phase Steels”, Materials and Design, Vol. 32, 433–440.

[7] Andrews, K.W. (July 1965), “Empirical Formulae for the Calculation of Some Transformation Temperatures*”. Journal of the Iron and Steel Institute*, Vol. 203 (7), 721-727.

[8] ASTM (2004), “Standard Practice for Laboratory Immersion Corrosion Testing of Metals”, G 31–72.

[9] Erdemir catalogue (2013), http://www.erdemir.com.tr/images/urun_hizmetler/katalog/sicak.html

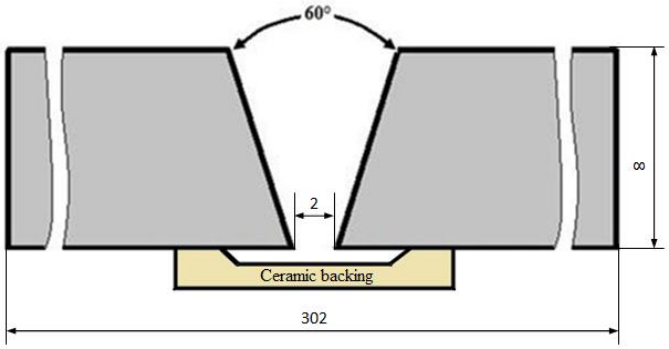


Figure 1. Schematic single V-groove butt joint with ceramic backing

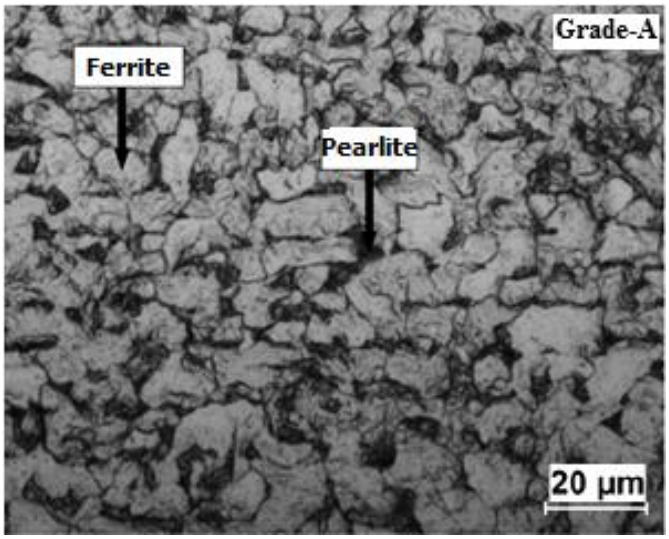


Figure 2. Microstructure of as-received Grade-A steel consisting of ferrite –pearlite structure

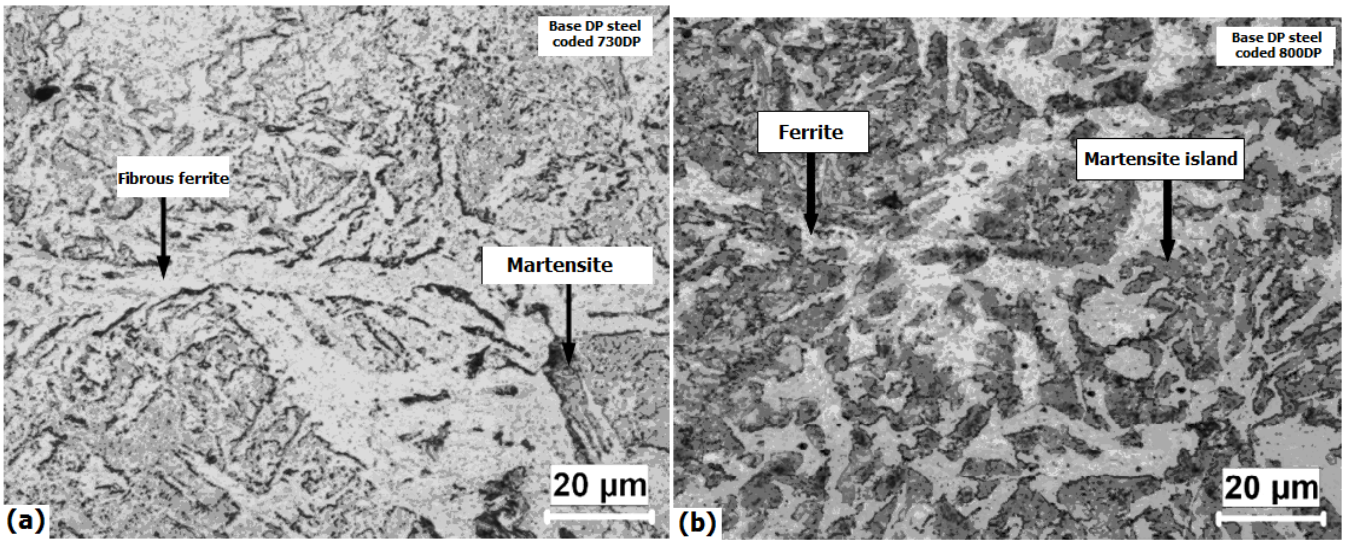


Figure 3. Microstructures of the base dual-phase steels (a) 730DP representing an annealing temperature of 730 °C and (b) 800DP representing an annealing temperature of 800 °C.

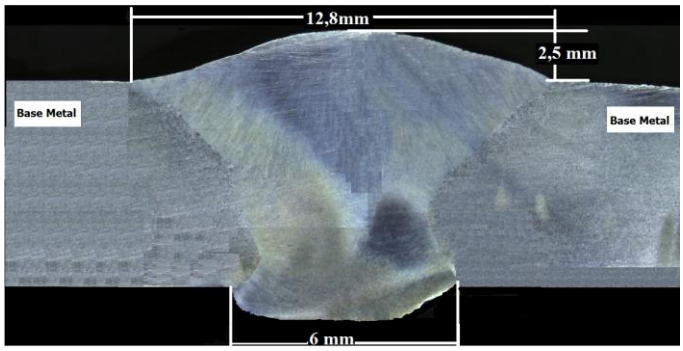


Figure 4. Representative macrostructure of the weld of Grade-A and dual-phase steel joints

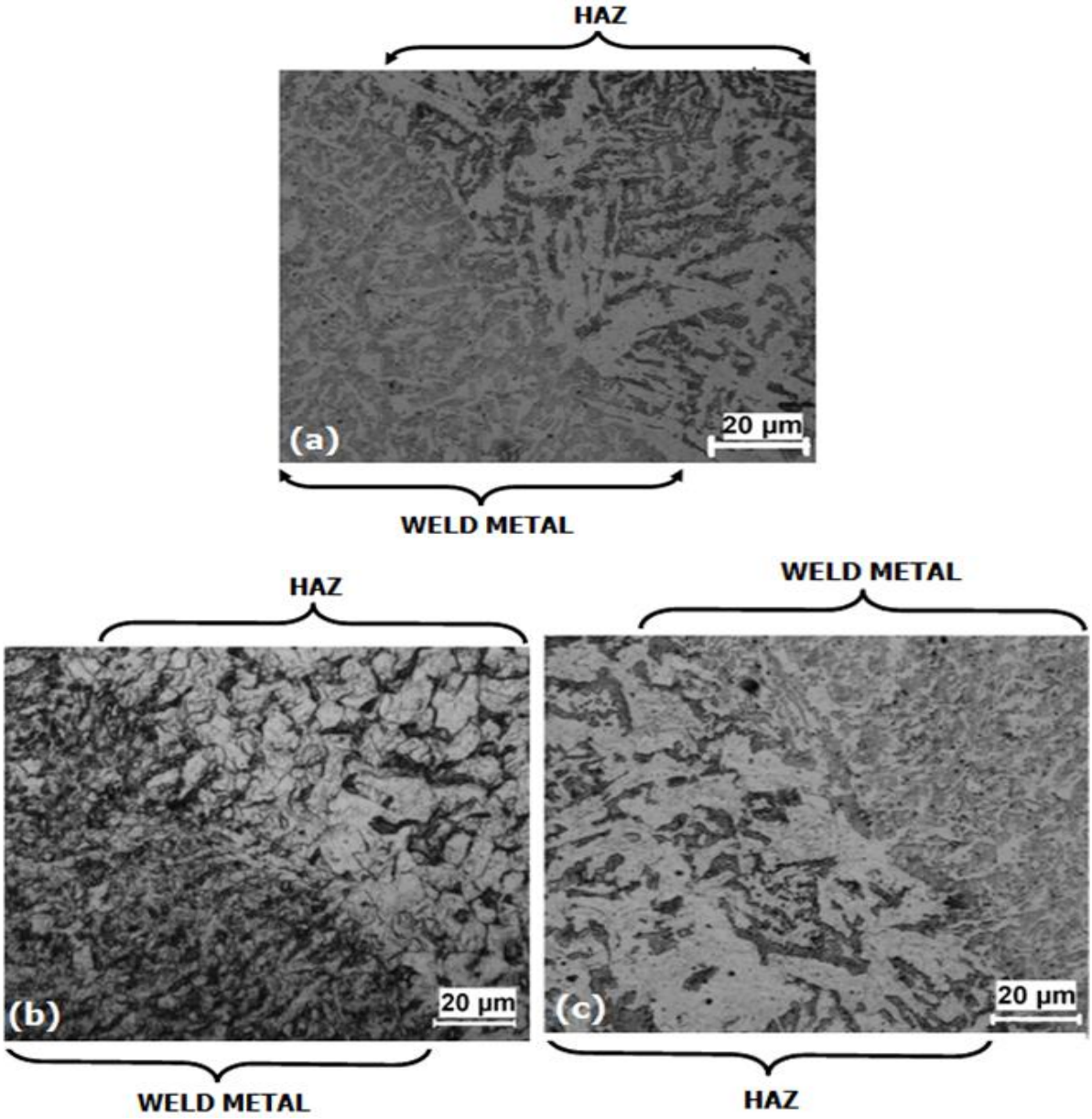


Figure 5. Weld metal and HAZ optical microstructure images of the similar Grade-A and the similar dual-phase steels welded by FCAW (a) Grade-A joint, (b) dual-phase steel joint coded 730DP and (c) dual-phase steel joint coded 800DP.

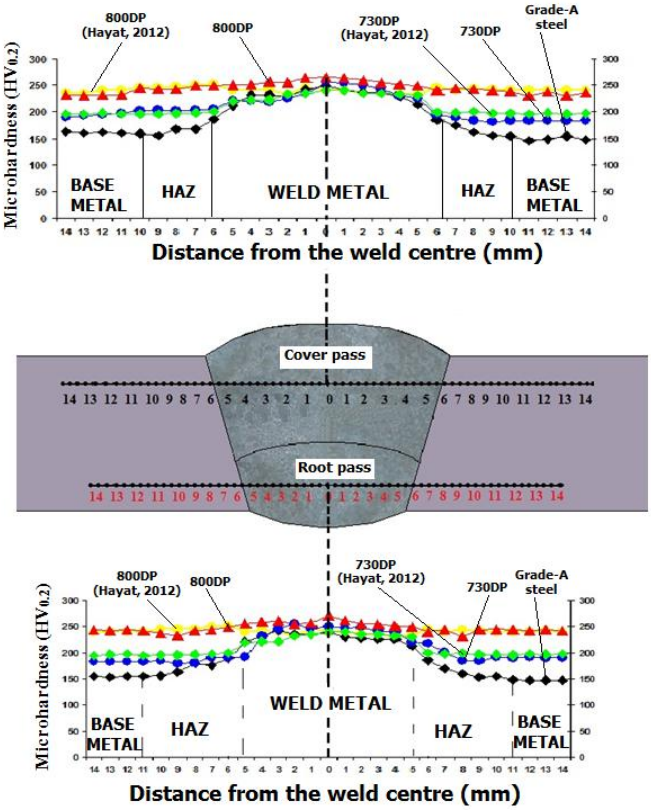


Figure 6. Hardness profiles along the top and bottom in cross-section of the Grade-A and dual phase steels

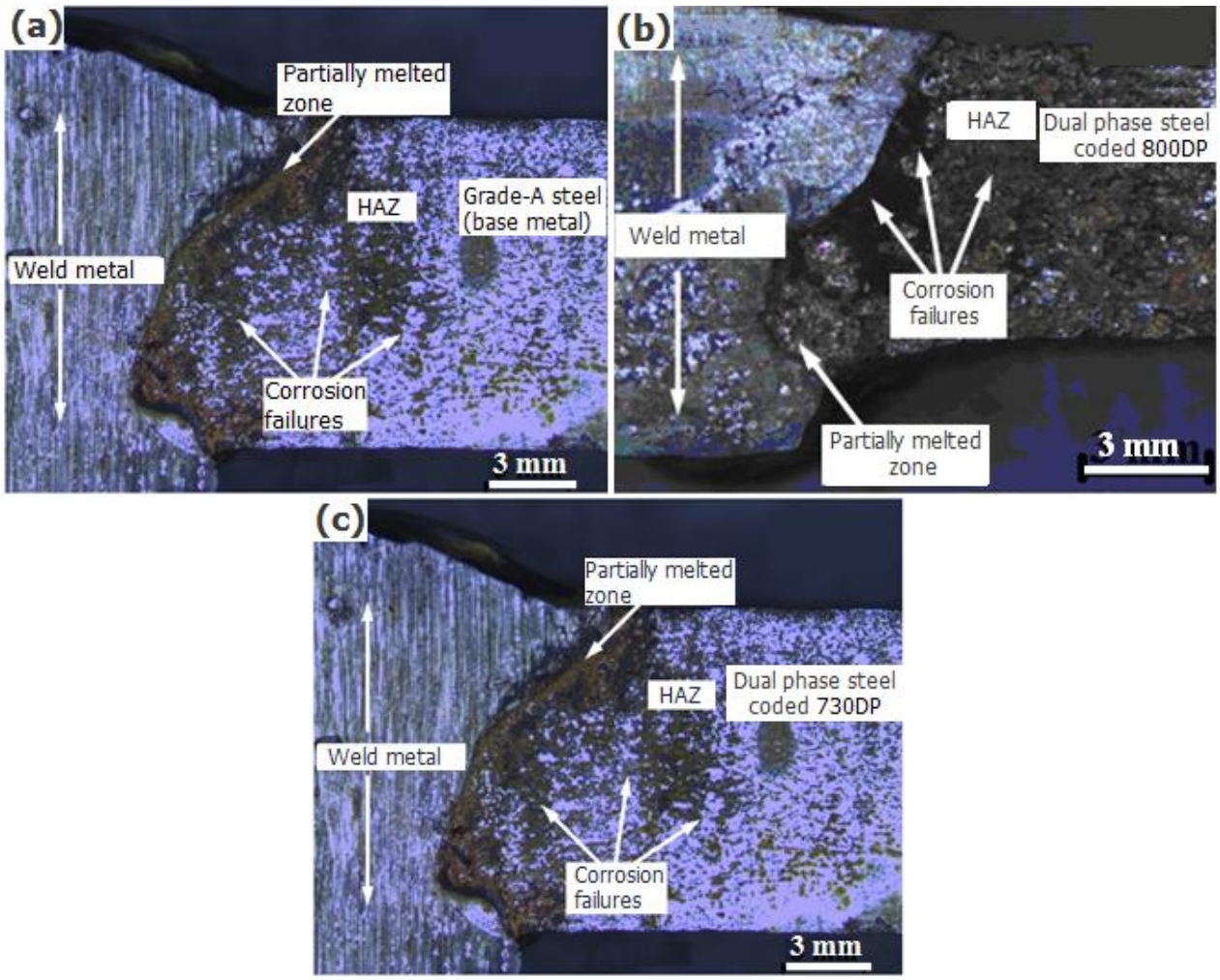


Figure 7. Macrostructures of weld region of Grade-A and dual-phase steels welded by FCAW after 90 days immersion corrosion test in 3.5% NaCl solution; ***(a)*** *Grade-A steel joint,* ***(b)*** *dual-phase steel joint coded 800DP and* ***(c)*** *dual-phase steel joint coded 730DP.*

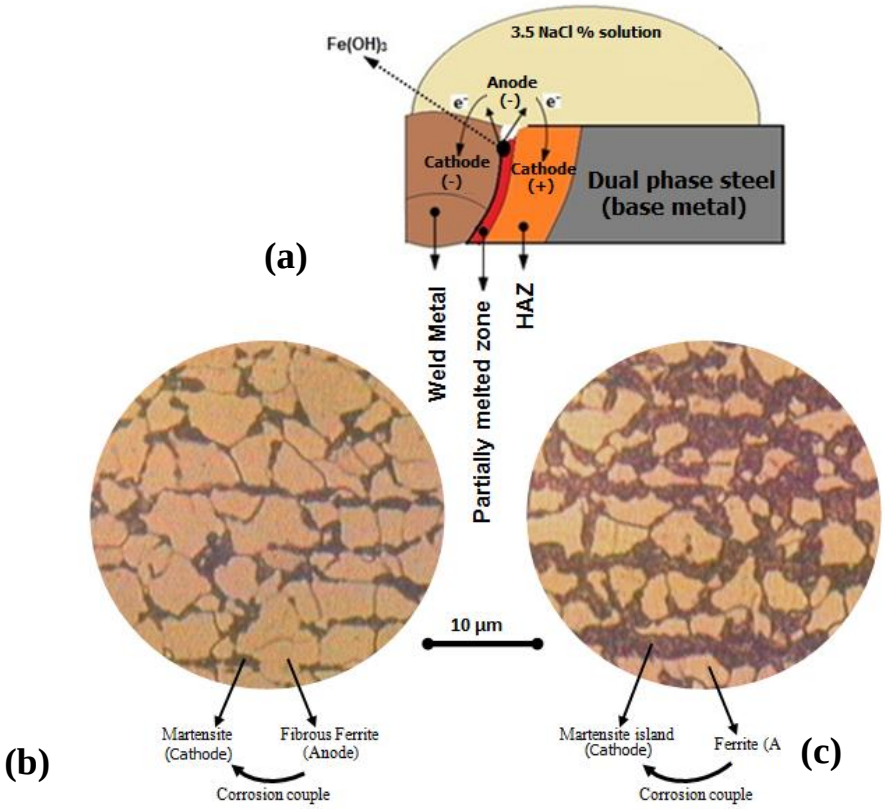


Figure 8. ***(a)*** *Schematic galvanic corrosion couples of dual phase steel welded by FCAW ((1) between the weld metal (cathode) and the partially melted zone (anode), (2) between HAZ (cathode) and the partially melted zone (anode)),* ***(b)*** *corrosion couple of base metal of dual phase steel joint coded 730DP developed following intercritical annealing at lower temperature of 730°C (between fibrous ferrite (anode) and martensite (cathode)) and* ***(c)*** *corrosion couple of base metal of dual phase steel joint coded 800DP developed following intercritical annealing at higher temperature of 800 °C (between ferrite (anode) and martensite island (cathode))*

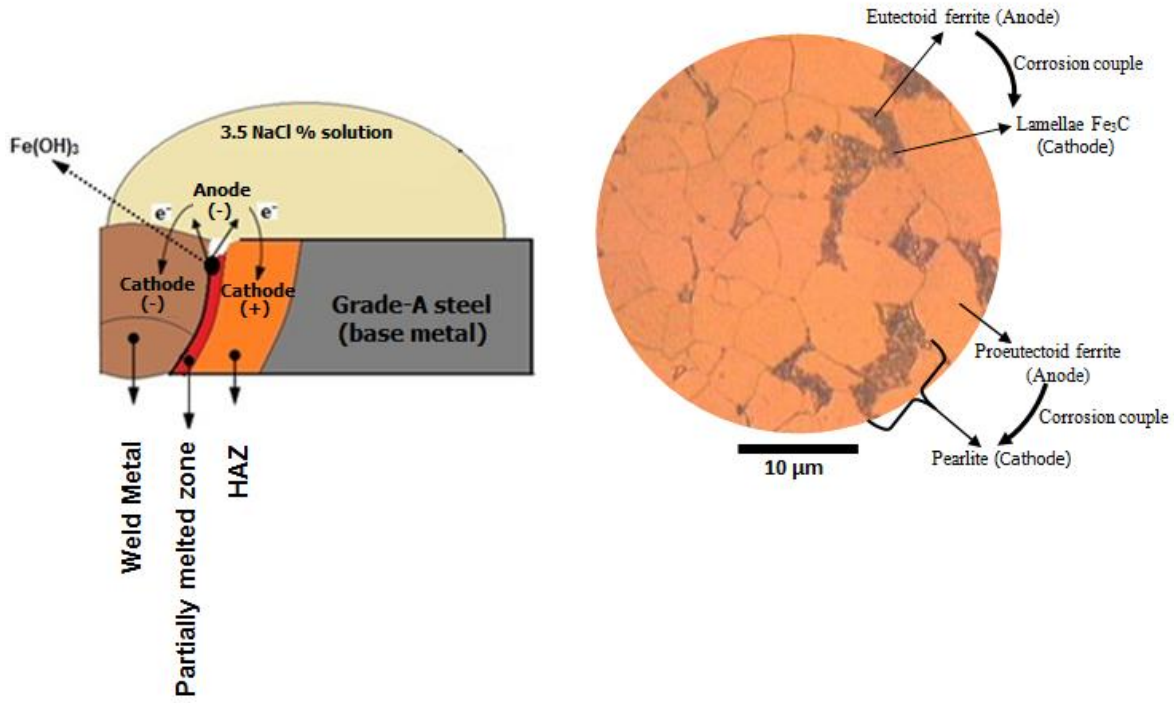


Figure 9. ***(a)*** *Schematic galvanic corrosion couples of Grade-A steel welded by FCAW ((1) between the weld metal (cathode) and the partially melted zone (anode), (2) between HAZ (cathode) and the partially melted zone (anode)),* ***(b)*** *corrosion couples of base metal of Grade-A steel joint ((1) between lamellae cementit in pearlite (cathode) and eutectoid ferrite in pearlite (anode), (2) between lamellae of pearlite (cathode) and proeutectoid ferrite (anode))*

Table 1. Chemical compositions and mechanical properties of the as-received hot rolled Grade-A steel

Materials	**Chemical composition (wt %)**					
	C	***Mn***	***Si***	***P***	***S***	***Fe***
Grade-A steel (ABS-P2)	0.12	0.71	0.14	0.016	0.015	Balanced
	Mechanical Properties (Erdemir, 2013)					
	Yield strength (MPa)		Tensile strength (MPa)		Elongation (%)	
	235		400 - 520		22	

Table 2. Tensile results of similar Grade-A and similar dual-phase steels welded by FCAW

Materials		**Yield strength (MPa)**	**Tensile strength (MPa)**	**Elongation (%)**	**MVF (%)**
Grade-A steel joint	Base of Grade-A steel	284±5	365±5	19	-
	Grade-A steel joint	317±5	410±5	16	
DP steel joint coded 730DP	Base of DP steel	371±6	429±4	16	16±5
	DP steel joint	490±5	580±7	13	
DP steel joint coded 800DP	Base of DP steel	345±5	437±5	14	55±6
	DP steel joint	450±5	534±5	11	

Table 3. Results of the average corrosion rates of similar the Grade-A and the dual-phase steel joints obtained from immersion corrosion test performed in 3.5% NaCl solution

Materials	Corrosion rate from immersion test (mm/year)	MVF (%)	Microstructure of base metal
Grade-A steel joint	1,334	-	Ferrite – Pearlite
DP steel joint coded 730DP	1,010	16±5	Uniform fibrous ferrite – Martensite
DP steel joint coded 800DP	0,861	52±6	Island martensite in continuous ferrite matrix

ISSN 2303-4521
PERIODICALS OF ENGINEERING AND NATURAL SCIENCES
Vol. 1 No. 2 (2013)
Available online at: http://pen.ius.edu.ba

Qualıty Management In Logıstıcs Sector: Prıncıples And Practıce

Sule Selcuk

Mechanical Engineering Programme
Faculty of Natural Sciences and Engineering
International University of Sarajevo
Hrasnicka Cesta 15, Ilidza, Sarajevo
Bosnia and Herzegovina
selcuk@ius.edu.ba

Abstract:
As it is the case for production and service sectors, one of the key ways for a logistics company to create sustainable competitive advantage is to differentiate itself from the rest with the quality that it offers. Researchers and practitioners have developed various quality improvement policies for logistics service providers and each of the policies addresses certain aspects of the business. This study individually evaluates the major quality approaches involved in an attempt to explore prospects of combining them in a way to maximize the benefit of their implementation. To promote applicability the quality approaches are analyzed from the perspective of customers as well as that of implementers of quality programmes.

Key Words : *Logistics, Quality Management, JIT Logistics, Green logistics, Value-added logistics*

1. Introductıon

Although originated from military efforts, logistics has now become a life-line for almost all sorts of activities in the world; from industrial and agricultural production, to service industry, from medicinal operations to humanitarian campaigns etc. Before 1960s it used to be considered merely as a "supporting" operation for the "main" job under consideration and, consequently, improvements recorded were bound to remain occasional only. In time, however, practioners and academicians have started treating logistics as a separate discipline in its own right after experiencing its complexity and also discovering many opportunities in it. Discrete improvements recorded through pratice have been transformed into principles, serving the industry systematically as bettering tools that are applicable to all other cases.

While Groover (Groover, 2007) names production, service, logistics and information as the four broad type of occupations, many treats logistics as a subset of service sector (Saura et al, 2008) and (Huang et al 2009). Regardless of their relative position, the three sectors, namely production, service and logistics, share many common concepts, approaches and tools of quality management. When developments in production and logistics are analyzed together, they seem to have followed somewhat parallel, even if not simultaneous, patterns. For example, in the case of production the overriding concern for the sector in earlier times was productivity and only subsequently quality has become an issue for the sector. Similar to the case of production, in logistics sector cost-cutting efforts were domiant first and importance of quality was appreciated in coming years (Table 1).

In addition to the parallelity between logistics and production with reference to quality, logistics sector also deploys quality concepts developed for service industry, which will be discussed in detail in the coming sections.

To stay competitive in the global market, Logistics Service Provider (LSP) companies have been substantially investing on the quality front and adopting any quality improvement available, regardless of the sector of origin, since there has been very little room for improvent in the cost-cutting front in these days due to globalization, which tends to equate all LSP in their access to energy and HR markets. This study compiles these approaches and analyzes particular contributions of each approach to the logistics sector.

2. Approaches Adopted For Logıstıcs Qualıty Management

Main goals of logistics are commonly and concisely expressed as 7Rs of logistics, denoting moving the right materials/products in the right quantitiy in the right condition at the right time to the right place at the right cost to the right customer. To achieve these goals various quality management approaches are adopted by the sector. It should be pointed out that since these approaches are not mutually exclusive alternatives, a given logistics company can adopt more than one of them at the same time, creating augmented improvements.

2.1. Servıce Qualıty Approach

As pointed out earlier, it is common to consider logistics as a subset of service industry and, therefore, utilizing service

DOI: 10.21533/pen.v1i2.22

quality approach for quality managements in logistics is also common. Service quality approch is based on the concept of customer-driven quality. As the name implies, some companies use service quality concept to assess their performance as percieved by the customer, because, after all, customer is at the receiving end of the service provided and the ultimate corporate goal is to satisfy the customer. The followings list is a selection of elements that are import to customers, regardless of them being an individual or a corporate customer (Filho and Souki, 2007), (Lambert and Burduroglu, 2000) and (Fugate et al, 2010):

- Time between order receipt and delivery
- Punctuality in the delivery (delivery in the agreed period of time)
- Absence of damage
- Resolution of complains
- Reliability of the information given by the personnel
- Information about the order's status
- Ease of placing order
- Flexibility
- Time in repairing faults
- Efficient communication
- Return policy
- Supplier absorbs cost of freight and handling on returns due to shipping damages or product shipped in error
- Quality/durability of packaging
- Product according to specification
- Delivery of the correct product
- The shipments rarely contains wrong items
- The shipment rarely contains incorrect quantities
- The shipments rarely contains substituted items
- Correct invoice
- The products are constantly in stock
- Advanced warning about delays
- Physical appearance of the employees, trucks and equipments
- Employees' capacity to solve problems
- Employees' cordiality
- Urgent and special services
- Availability of technical information

The above elements are what the customers take into consideration in evaluating quality and, consequently, they are the ultimate targets of logistics companies. To reach these ultimate targets, the management sets internal and interim operational targets for itself to reach, which are constantly monitored to evaluate company's performance. In other words, the list above tells what customers want, but the job of how to deliver them is another matter and it requires a set of performance metrics similar to the ones in Table 2. (Wojciechowska, 2011):

It should be noted that the metrics in the table are basically a kind of interim indications that give an idea about the performance of the company, rather than the data of standards related to what are being delivered to the customer. Consequently they can only be used internally as a feedback to figure out how to reach the ultimate target, which is customer satisfaction.

2.2. Benchmarkıng

Benchmarking is one of the common concepts that has been utilized in any quality management programme. It essentially involves determining the leading firm in a particular field as the reference against which the given company compares itself. The comparison leads to question performance of the company and then to identify the areas for improvement and finally conduct the improvement work. The following list is a selection of factors that are used for benchmarking (Rushton, 2010):

- lead time
- stock availability
- minimum delivery/order size policy
- order and delivery frequency
- full loads delivered on time

2.3. Just-In-Tıme (Jıt) Logıstıcs

Although defined differently by many, JIT is basically a management philosophy with some tools that focuses on waste reduction, where the word "waste" not only refers to scraps or lost time but also to any non-value-adding-activites/inventory/space/skills. It was originally developed for improving production sector, but later deployed by service industries although relatively slowly (Lai and Cheng, 2009). As stated by Lai and Cheng, "While JIT is a process-orientated waste elimination management approach, the principles are relevant and applicable to both production and service firms processes and systems to perform the tasks of production and delivery of products or services." (Lai and Cheng, 2009) .

The followings are some areas where JIT philosophy can record progress:

- Minimize idle capacity for transport through more efficient scheduling
- Maximize carriage capacity through more efficient routing
- Eliminate unnecessary inventory unless there is a considerable price discount to gain or a possibility of stock-out loss
- Streamlined information flow by integration between all parties involved

2.4. Value Added Logıstıcs

Another opportunity that a logistics firm can explore for differentiating itself with its quality is to offer value added logistics to its customers. Moving pre-sales production functions and after-sales services to an LSP transforms logistics function to value-added activities that increase customer satisfaction, among other benefits (Li, 2011). The followings are examples of value-adding activities that can be encountered in logistics (UNESCAP), (Zeybek and Kaynak, 2008) and (Gudehuss and Kotzab, 2012):

- Labelling, for example for customizing purposes
- (Re)Packaging
- Assembling
- Adding manuals
- Breaking bulk/
- Palletizing/unitizing

- Stretch-shrink-wrapping
- Final assembly
- Repairment
- Blending/mixing
- Cleaning
- Quality control
- Reverse logistics

2.5. Green Logıstıcs

The EU identified five areas of economic activity that may affect the environment, as follows (Rushton, 2010):

- Tourism
- Energy
- Transport
- Agriculture
- Industry

While the above list signifies gravity of logistics on the environment on the one hand, it also indicates opportunities of making logistics activities more environmentally friendly. Since more and more customers have concerns about the environment and perceive greener companies as more quality ones, even if they receive the same level of service, one of the key ways for a logistics provider to differentiate itself qualitywise is to opt for operations with less damage to the environment. Table 3 gives a summary of ecological objectives for logistics sector, along with the means to achieve them:

In order to evaluate their own performance, companies should monitor and keep records of the following data of their activities:

- Fuel consumption (lt/km)
- Used oil (lt/km)
- Percentage of empty kilometers run by vehicles
- Percentage of utilization of vehicle load space
- Targets for reducing waste packaging

Although corporate customer may be less sensitive to impact of logistics activities on the environment, individual customers take this issue into account in great deal in making their decisions. Especially in urban logistics and courier service businesses it is a must for businesses to provide environmentally friendly logistics services. Use of bicycle couriers is an obvious solution not only for a more environmentally friendly logistics but also for a faster service (Maes and Vanelslander, 2011).

2.6. Standards, Qualıty Certıfıcates And Awards

Similar to the early phases of quality movement in production, where quality concept was perceived simply a matter of detecting and discarding the products that do not comply with certain standards, logistics sector has also established certain standards that are used as criteria for assessing logistics quality. This approach, however, is not as straight forward as production as there are many intangible factors that determine logistics quality, setting definite standards. Nevertheless standards, such as family of ISO 9000s, ISO 14000s are available for quality improvement efforts, the latter governing ecological side of logistics operations.

Research on ISO 9000 implementation in 288 Spanish companies reveals (Casadesus and Gimenez, 2000) that 80 percent of the companies say that customer satisfaction has increased and 63 percent say that customer complaints have decreased.

Quality awards delivered by various national and international institutions can also be useful and important indications that should be taken into consideration since they, in a way, reflect voice of the customer.

2.7. Relatıonshıp-Based Logıstıcs

It is a well established fact that “maintaining existing customers and extending business with them is significantly less expensive than acquiring new customers” (Cahill, 2007). Since logistics sector typically offers services that a customer needs more frequently than many other necessities, such as buying a fridge, relationship-based logistics naturally attracts considerable amount of attention. Relationship-based logistics, needless to say, requires customer satisfaction and as Cahill stated “..quality is an antecedent of satisfaction”. Offering such a long-term relation is also statement of self-confidence, reinforcing the perception of being a company of quality. Cahill further states that “According to social exchange theory, the benefit and consequently the satisfaction perceived by the customer depend both on cost and on performance aspects, the latter being represented by service quality” (Cahill, 2007)

Many companies seek to establish relationship-based interactions as opposed to transactional ones and this can only be achieved through high quality service delivered consistently.

3. Conclusıons

As with production and service industries, logistics sector is also focusing on customer satisfaction in its quality undertaking and to achieve that it is, too, using regular quality management concepts such as JIT, benchmarking etc. However quality improvement efforts of logistics companies have far gone beyond meeting customers’ demand, which can be extremely changeable and not always easy to describe. Companies are well aware of the fact that they must not only satisfy their customers’ expectations but also create such quality that the customers have not asked for yet at their transactions. In this context, the concepts of value-added logistics and relationship-based logistics are all those efforts of offering quality that the customer has not experienced and, as a result not demanded yet. These concepts are being widely practised by many companies as a part of their quality management programmmes, distinguishing themselves from their rivals. In addition to these concepts of value-added logistics and relationship-based logistics, the concept of green logistics also inherently lends high quality logistics opportunities.

4. Tables

Table 1 - Overriding Concerns in Production and Logistics, Illustrating Their Parallel Evolution

	PAST	PRESENT
PRODUCTION	Productivity	Quality
LOGISTICS	Low-cost	Quality

Table 2. Some Internal Logistics Metrics Used to Evaluate Performance of an Logistics Company

METRIC'S NAME	METRIC'S FORMULA	UNIT
Orders delivery timeliness	Number of fulfilled orders / Number of orders to be fulfilled	[%]
Number of lost deliveries	Number of all lost deliveries for which company is responsible	[unit]
Number of transportation damages	Number of all transportation damages for which company is re-sponsible	[unit]
Employees being on call	Number of picking up phone (during 30 seconds) / All phone calls	[%]
Completeness of deliveries fulfilment process	Number of deliveries fulfilled completely /number of all deliveries fulfilled during a chosen time interval	[%]
Number of complaints per 1000 bill of ladings	Number of complaints raised in the chosen division / 1000 bill of ladings	[unit./1000 bill of ladings]

Table 3 : Ecological objectives for logistics and the means of attaining them

OBJECTIVES	MEANS
Minimum fuel and energy consumption	* vehicles with more efficient motors * vehicles with hybrid motor in urban logistics * well-maintained vehicles (tyres, brakes, clutches etc) * aerodynamic loading * bikes for courier service
Minimum emission of harmful gases	* vehicles with catalitic convertor
Waste reduction	* re-usable containers/packages * recycle packaging * recycle used oil * adopt predictive maintenance for optimum oil-change timing

5. References

[1] Cahill D. L., (2007), "Customer Loyalty in Third Party Logistics Relationships", ISBN-10 3-7908-1903-4 Physica-Verlag Heidelberg New York.

[2] Casadesús M and, Giménez G., (2000), "The benefits of the implementation of the ISO 9000 standard: empirical research in 288 Spanish companies", The TQM Magazine, Vol. 12 Iss: 6, pp. 432 – 441.

[3] Rushton A., (2010), "The Handbook of Logistics and Distribution Management, Fourth Edition.

[4] Fugate B. S. , Mentzer J. T. and . Stank T. P., (2010), "Logistics Performance: Efficiency, Effectiveness and Differentiation", Journal of Business Logistics, vol.31, No. 1, pp. 43-62.

[5] Gonçalves C. F.and Souki G. Q., (2007), "Measuring the Logistic Services Quality and the Retailers' Behavioral Intentions: An Empirical Study of the Soft Drink Industry", POMS 18th Annual Conference Dallas, Texas, U.S.A. May 4 to May 7.

[6] Groover M, (2007), Work Systems The Methods, Measurement & Management of Work: International Edition Pearson, ISBN13: 9780131355699.

[7] Gudehus T. and Kotzab H, (2012), "Comprehensive Logistics", 2nd ed., Springer-Verlag Berlin Heidelberg, ISBN 978-3-642-24366.

[8] Huang Y. K, Kuo Y.W. and Xu S.W., (2009), "Applying Importance-Performance Analysis to Evaluate Logistics Service Quality for Online Shopping Among Retailing Delivery", International Journal of Electronic Business Management, vol. 7, no.2, pp. 128-136.

[9] Lai K. H. and Cheng T.C.E., (2009), "Just-in-Time Logistics", ISBN: 978-0-566-08900-8, Gower Publishing Limited.

[10] Lambert D. M. and Burduroglu R., (2000), "Measuring and Selling the Value of Logistics", The International Journal of Logistics Management, Vol.11, Number 1, pp.1-17.

[11] Li L., (2011), "Assessing the Relational Benefits of Logistics Services Percieved by Manufacturing in Supply Chain", International Journal of Production Economics, 132, pp. 58-67.

[12] Maes M. J. and Vanelslander T, (2011), "The Use of Bycycle Messengers in the Logistics Chain, Concepts Further Revised", The Seventh International Conference on City Logistics.

[13] Saura I. G., Frances D.S. B. and Blasco M. F., (2008), "Logistics Service Quality : A New Way to Loyalty", Industrial Management and Data Systems, vol. 108, no. 5, pp. 650-668.

[14] UNESCAP, Value-Added Services Of Logistics Centres In Port Areas, Commercial Development of Regional Ports as Logistics Centres, United Nations

[15] Wojciechowska S. W., (2011), "On Logistics Service Quality Evaluation – Case Study", Logistics and Transport, No 2 (13).

[16] Zeybek H. and Kaynak M, (2008), "What Role for Turkish Ports in the Regional Logistics Supply Chains?", International Conference on Information Systems, Logistics and Supply Chain, Wisconsin, USA, May pp. 27-30.

ISSN 2303-4521

PERIODICALS OF ENGINEERING AND NATURAL SCIENCES
Vol. 1 No. 2 (2013)
Available online at: http://pen.ius.edu.ba

Production of Functionally Graded AlB2/Al-4%Mg Composite by Centrifugal Casting

Ömer Savaş
Yıldız Tecnical University, Naval Architect and Marine Engineering Faculty, 34349 Istanbul, Turkey

Ramazan Kayikci
Sakarya Üniversity, Faculty of Technology, Department of Mechanical Engineering, 54187, Sakarya, Turkey

Ferit Ficici
Sakarya Üniversity, Faculty of Technology, Department of Mechanical Engineering, 54187, Sakarya, Turkey

Sakip Köksal
Sakarya Üniversity, Faculty of Technology, Department of Mechanical Engineering, 54187, Sakarya, Turkey

Abstract:

In this study, production of functionally graded AlB_2 reinforced Al-4%Mg composite has been investigated. In-situ high-aspect-ratio flake AlB_2 reinforcement was obtained by the addition of boron oxide to the liquid Al-Mg at 1400 °C. It has been observed that there were two distinct regions without a smooth grading: one is rich of and the other is poor of AlB_2. The results showed that the highest AlB_2 content has been found to be around 10% in external zones, whereas the lowest AlB_2 content has been observed to be 0.02% in the internal zones. Additionally, results showed that, depending on the increase in the reinforcement phase in external zones, up to 20 % increase in the hardness of the composite has been achieved.

Key words: Aluminum boride, AlB_2, boron oxide, functionally graded composites.

1. Introduction

Metal matrix composite have been gaining importance in engineering applications, because of their high specific strength, high modulus and high wear resistance [1-6]. Functionally graded metal matrix composites have a great significance for the use in automobile, aerospace and defence industries. A functionally graded metal matrix composite can be defined as its reinforcement particle's volume fraction varies continuously from the inner to the outer sections of the part. Therefore, their mechanical properties are different from the inner region and outer region accordingly [7, 8].

Until recently, a number of studies on functionally graded metal matrix composites produced by centrifugal casting had SiC, Al_2O_3 and AlB_2 reinforcement particles in various aluminium alloys [7-9]. Production of functionally graded materials under centrifugal force is an effective method. Advantage of centrifugal casting is good mould filling combined with good microstructural control. Aluminium boron master alloys have been used commercially to scavenge transition-metal elements during the production of aluminium electrical wires [10] and is also used for grain refiner as Al-Ti-B in aluminium castings [11]. It is known that the crystal structure of borides is a hexagonal close packed (HCP). Depending on the cooling rate they can be in fine flake shaped structure having few micron thickness with aspect ratio ranging from 30 to 400 [12, 13].

AlB_2 flakes in liquid aluminium alloys can be precipitated during solidification. [14-17]. Thus, the reinforcement phase (AlB_2 flakes) can readily be produced by in-situ fabrication techniques. There are a few study regarding AlB_2/ Al composites for producing functionally graded metal matrix composites [9, 18, 19]. AlB_2 phase have high strength and high modulus and has higher density (3.19 g/cm^3) [18], than liquid aluminium (2.4 g/cm^3) at 700 °C.

In the present work, in-situ AlB_2 flake reinforced Al-Mg matrix composites were in-situ prepared by the reaction of boron oxide (B_2O_3) with liquid Al at a reaction temperature of 1400ºC. Centrifugal casting was applied to the composite to produce functionally graded material FGM.

2. Experimental Procedure

For the fabrication of AlB_2 flake reinforced Al-4%Mg composites were prepared by the reaction of boron oxide (B_2O_3) with Al at a reaction temperature of 1400ºC as explained in detail elsewhere [13]. The composites were placed in a die and the die was heated at 800°C in an

DOI: 10.21533/pen.v1i2.23

electric resistance furnace to bring the composites to a semi-solid state ($Al_{(liquid)}$ + $AlB_{2\,(solid)}$). Centrifugal action was then employed with the die containing semi-solid composite to drive the solid AlB_2 particles towards the outer region to produce a functionally graded AlB_2/Al-Mg composite with even higher volume per cent reinforcement. The centrifugation process was carried out under rotation speed of 600 rpm at 800 ºC. A schematic representation of the in-situ production and the following centrifugation process of the functionally graded AlB_2/Al-Mg composite were shown in Fig. 1.

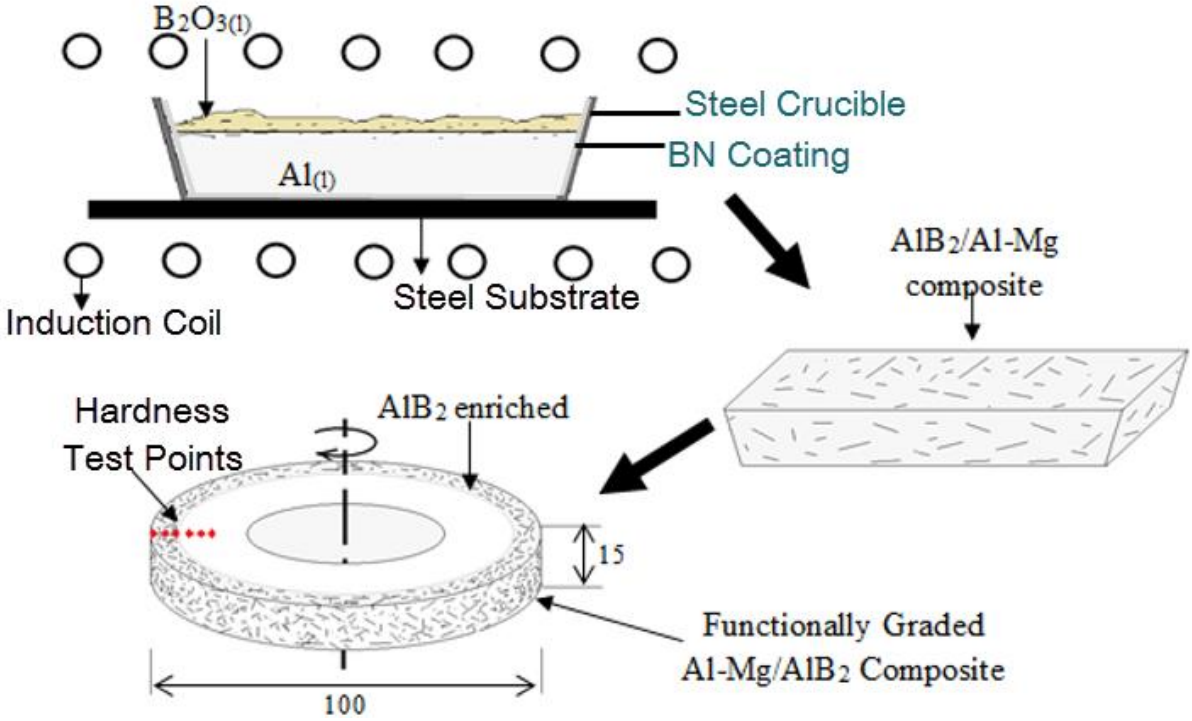


Figure 1. Schematic illustration of AlB_2 /Al-Mg composites production route.

Since boron cannot be quantified by Energy Dispersive Spectrometer (EDS), measurement of total weight per cent of boron in the AlB_2/Al-Mg composites were carried out by a wet chemical analysis method, as explained elsewhere [13]. For examination, Small samples were extracted from the AlB_2/Al-Mg composites. For the microstructure analysis the samples were ground up to 1200 grid by using SiC paper followed by polishing, using 0.2 µm diamond paste. For detailed evaluation of the boride structure, the samples of composites were deep etched using % 10 HCl solution followed by an examination with JEOL JSM 6060LV scanning electron microscope (SEM). Brinell hardness (BHN) of the composites was measured after polishing to a 1 µm finish. Hardness of composites was measured with 2 mm ranges on the composites from outside to inside as seen Fig 1.

The hardness values of the samples were measured using a 2.5 mm diameter ball at load of 31.25 kgf for 10 sec. The composites were solutionized at 540 °C for 4 h followed by water quenching at 60 °C, waiting at room temperature for 12 h, and aging was carried out at 190 °C for 10 h.

3. Results And Discussion

The mixture of boron oxide and Al-4%Mg alloys was melted and heated to 1400 °C for 1.5 h to maximize the boron solubility in the liquid Al-Mg alloy as explained in detail in a previous work [10]. During synthesizing a portion of boron oxide, chemical reaction of Equation (Eq.) 1 occurs at the interface between the molten alloy and boron oxide resulting in boron in liquid solution and aluminium oxide mixed with remaining (unreacted) boron oxide on top of the melt [11, 12]. In Eq. (1), α represents aluminium boron solid solution at 298 degree K. ΔG and ΔH are the Gibbs free energy and the enthalpy of the reaction respectively.

$$2Al + B_2O_3 \rightarrow \alpha - Al_2O_3 + 2[B] \qquad (1)$$

$$\Delta G^0_{298^\circ} = -416.9\ kJ/mol$$
$$\Delta H^0_{298^\circ} = -402.7\ kj/mol$$

Results from wet chemistry analyses showed that the amount of boron in the as-cast (unfiltered) casting was

1.2 wt % as all the boron atoms within the Al-B alloy can be considered to be boride compounds as reported in previous studies [9, 14]. The volume content of the boride phase within the matrix was calculated as 2.6 %. The microstructure of 2.6 vol. % AlB_2/Al-Mg composite material is given in Fig. 2.

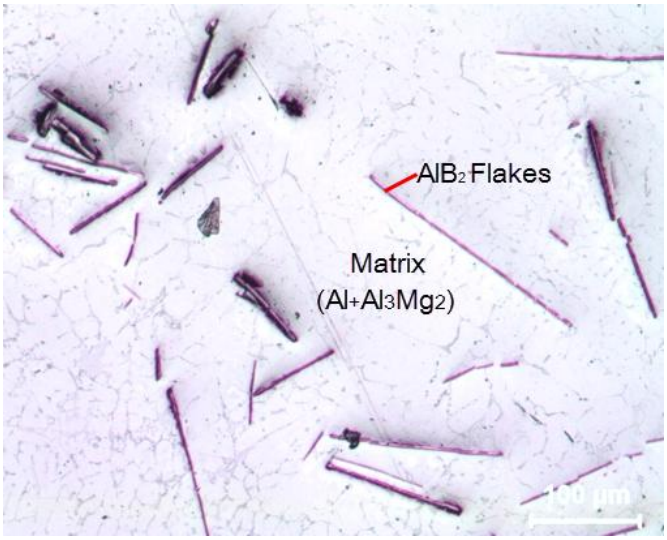


Figure 2. Microstructure of the in-situ 2.6 vol. % AlB_2/Al-Mg composite material

In Fig. 2, it can be observed that the in-situ AlB_2 boride particles with flake crystals and the β- Al_3Mg_2 phases with needle-shape form are distributed uniformly in the aluminium matrix.

A sample of 2.6 vol.% AlB_2 reinforced aluminium composite was deep-etched to extricate boride flakes from the aluminium matrix. Fig. 3 shows a SEM image of %2.6 vol. AlB_2/Al-Mg composite. As seen in the SEM image, AlB_2 structures have a fine flake shape and a high aspect ratio (width/thickness). This result is in good agreement with previous reports on AlB_2/Al [13, 14, 16].

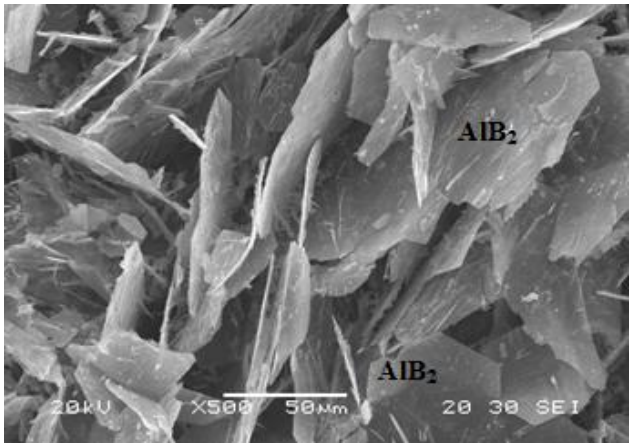


Figure 3. SEM image of the deep etched AlB_2/Al-Mg composite

The 2.6 % vol. AlB_2/Al-Mg composite was heated for centrifugal casting to 800 °C to produce functionally graded (FG) composites material. After the centrifugal casting process the composite was cooled to room temperature for examinations. A cross section of the FG composites after deep etching with 10% HCl solution for 30 sec is shown in Fig 4. It has been observed that there were two distinct regions; the darker region which is named the external zone, the lighter region which is named the internal zone.

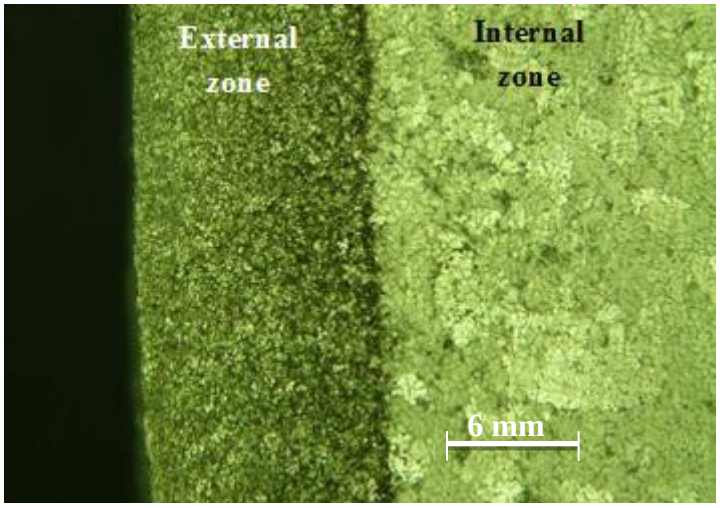


Figure 4. A cross section of the functionally graded AlB_2/Al-Mg composites

The microstructures of the darker and the lighter regions are shown in Fig 5. As the seen in Fig. 5a that there are numerous of AlB_2 reinforcement particles embedded in the aluminium matrix, whereas, there are almost no AlB_2 particles within the Al-Mg matrix in the internal zone as seen Fig 5b. These results show that all the AlB_2 flakes have been segregated by the centrifugal force at 800 °C as also reported in previous work [9,18].

Results from wet chemistry analyses showed that volume fraction of AlB_2 flake in the external zone increased significantly. The mean volume fraction of AlB_2 flake in this region has been measured as 10 vol. %. On the other hand in the internal region it was 0.02 vol. %.

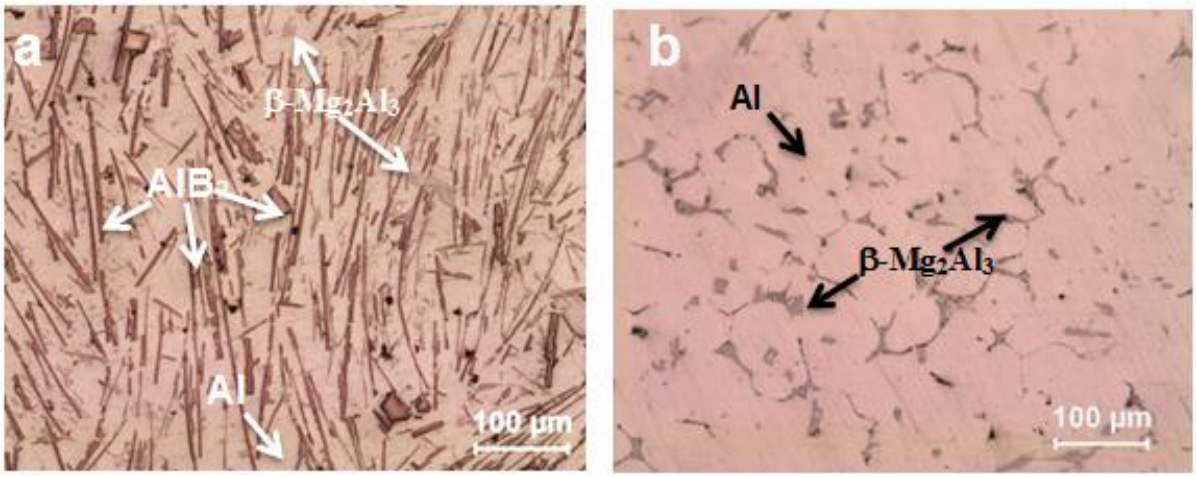


Figure 5. Microstructure of composite of a) external zone and b) internal zone.

Examination with optical microscope shows that the average thickness of AlB_2 is less than 1 µm. The measured average AlB_2 width distribution histogram is seen in Fig. 6. As seen from these histograms that, the average width of AlB_2 flakes is 170 µm, the width of the flakes vary between 40 µm and 350 µm and has demonstrated a normal (Gaussian) distribution within the matrix.

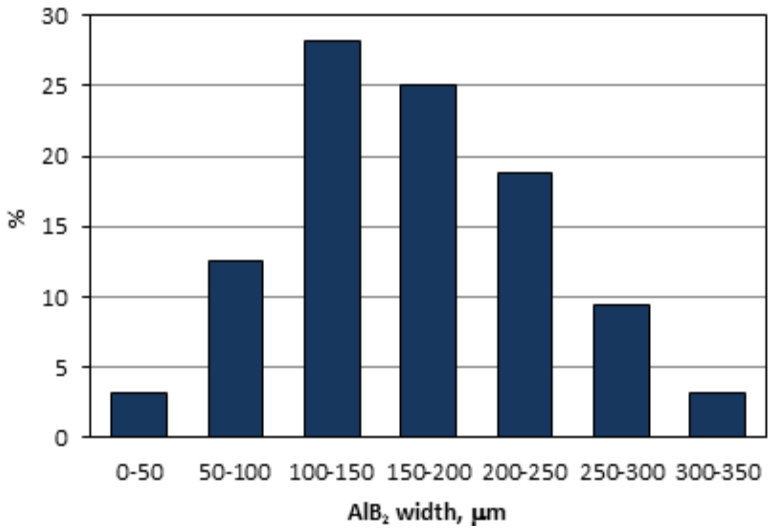


Figure 6. AlB_2 flake width histogram

Measured Brinell hardness as a function of radial distance along the centrifugally cast composite is given in Fig. 7 for the functionally graded AlB_2/Al-Mg composites. A significant increase in hardness of external zone was observed. It can be seen that the Brinell hardness value increased with solution heat treated composites. The functionally graded AlB_2/Al-Mg composites the external zone has the highest hardness with 90 HB. These results show that the hardness of Al-Mg alloys were significantly increased by the addition of 10 % of AlB_2 flake. This result is in good agreement in a number of experimental works reported for AlB_2/Al type composites [9,15,18-22].

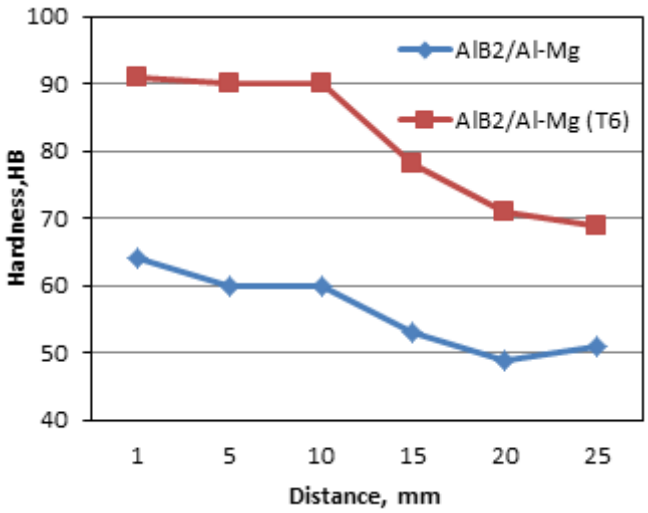


Figure 7. Measured Brinell hardness (HB) as a function of radial distance along the centrifugally casting composite

4. Conclusions

1- AlB_2 Reinforcing particles were produced by chemically reacting aluminium alloys with boron oxide at 1400 °C for 1.5 h. It was observed that the AlB_2 flakes in Al-Mg alloy have average 170 micron width, fine flake shape and volume fraction of AlB_2 flake is 2.6 vol.% in as cast condition.

2- Results showed that volume fraction of AlB_2 flakes in external zone significantly increased under centrifugal force. The average volume fraction of AlB_2 flakes in the external region was 10 vol.%. However in the other region was decreased to almost zero.

3- These results show that the functionally graded AlB_2/Al-Mg composites in external zone have the highest hardness with 90 HB. These results show that the Brinell hardness of Al-Mg alloys were significantly increased by the addition of 10 % of AlB_2 flake.

Acknowledgements

This work has been carried out at Sakarya University and supported financially by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under the grant number 107M023. Thus, authors gratefully acknowledge Sakarya University and TÜBİTAK.

References

[1] Taya M., Arsenault R.J., 1989, "Metal Matrix Composites: Thermo-Mechanical Behavior", Pergamon Press.

[2] Miracle D.B., (2005)," Metal matrix composites – From science to technological significance", Composites Science and Technology, 65:2526–2540.

[3] Rohatgi P.K., Cast Metal Matrix Compozites: Past, Present ve Future, American Foundry Society, Silver Anniversary Paper, Div. 2. AFS Transactions, 01-133 (2001) 1-25.

[4] Rawal Suraj P.,(2001)" Interface Structure in Graphite-fiber-reinforced Metal–matrix Composites", Surface And Interface Analysis,31: 692-700.

[5] Miracle, D. B., (2001-b) "Aeronautical Applications of Metal-matrix Composites", ASM International, 1043-1049.

[6] Prasad S.V., and R. Asthana (2004)," Aluminum Metal–matrix Composites for Automotive Applications: Tribological Considerations" Tribology Letters, 17 (3): 445-453.

[7]Attia, A.N. (2001), "New Phase Reinforcement for Composite Materials" Materials and Design, 22(6): 459-466.

[8] Leon L. Mishnaevsky Jr, (2006), "Functionally Gradient Metal Matrix Composites: Numerical Analysis of the Microstructure–strength Relationships" Composites Science and Technology 66: 1873–1887.

[9] Melgarejo Z.H., Suarez O.M., Sridharan K., (2006), "Wear Resistance of a Functionally-graded Aluminum Matrix Composite", Scripta Materials 55: 95-98.

[10] Cooper P.S, Kearns, MA., (1996), "Removal of Transition Metal Impurities in Aluminium Melts by Boron Additives", Materials Science Forum, 217–222:141–146.

[11] Sigworth,G.K., Kuhn,T. A., (2007), "Grain Refinement of Aluminium Casting Alloys", AFS Transaction, 115: 1−12.

[12] Hall A.C, (1995), "Pathways Family of Low Cost, High Performance, Metal Matrix Composites Based on AlB_2 in Aluminium", B.S Thesis., The University of Tulsa.

[13] Savas, O., Kayikci, R., Koksal, S., (2012) "Application of Taguchi Method to Investigate the Effect of Some Factors on In-situ Formed Flake Structures of Al/ AlB_2 Composite", Advanced Composite Letters, 21:44-50.

[14] Deppisch, G. Liu, J.K. Shang, J. Economy, (1997)," Processing and Mechanical Properties of AlB_2 Flake Reinforced Al-alloy Composites", Material Science Engineering A, 225:153-161.

[15] Ficici F., Koksal S., Kayikci R., Savas O., (2011), "Investigation of Unlubricated Sliding Wear Behaviours of In-Situ AlB_2/Al Metal Matrix Composite" ,Advanced Composites Letters , 20:109-116.

[16] Deppisch G., Liu A., Hall Y., Xu A., Zangvil J.K., Shang J. Economy, (1998)," The Crystallization and Growth of AlB_2 Single Crystal Flakes in Aluminum", Journal of Materials Research, 13(12): 3485-3497.

[17] El-Mahallawy, M. A. Taha, A.E.W. Jarfors, H.Fredriksson (1999)," On the Reaction Between Aluminium, K_2TiF_6 and KBF_4", Journal of Alloys and Compounds, 292:221–229.

[18]Melgarejo Z.H., Suárez O.M., Sridharan K., (2008), "Wear Behavior of Functionally Graded Al-Mg-B Composites Fabricated by Centrifugal Casting," Composites Part A: Applied Science & Manufacturing, 39(7):1150–1158.

[19] Duque N.B., Melgarejo Z.H., Suarez O.M., (2005), "Functionally Graded Aluminum Matrix Composites Produced by Centrifugal Casting", Materials Characterization, 55:167–171.

[20]Calderon, H. E., Suarez, O. M., (2008), " Thermomechanical Effects on Aluminum Matrix Composites Reinforced with AlB2 Particles", Journal of Composite Materials, 42: 2651-2672.

[21] Suarez, O. M., (2003), "Precipitation Hardening of a Novel Aluminum Matrix Composite", Materials Characterization 49:187–191.

[22]Kubota M., Sugamata M., (2008), "Reaction Milled and Spark Plasma Sintered Al-AlB_2 Composite Materials" Reviews on Advanced Materials Science, 18:269-275.

ISSN 2303-4521
PERIODICALS OF ENGINEERING AND NATURAL SCIENCES
Vol. 1 No. 2 (2013)
Available online at: http://pen.ius.edu.ba

A Conceptual Framework of a Cloud-Based Customer Analytics Tool for Retail SMEs

Abdulkadir HIZIROĞLU
School of Management
Yıldırım Beyazıt University
Ankara / Turkey
hiziroglu@ybsum.info

Halil İbrahim CEBECİ
Business Faculty
Sakarya University
Sakarya / Turkey
hcebeci@sakarya.edu.tr

Abstract
Since customers are seen as a strategic element in a company's downstream supply chain, many retail organizations have been employing a customer-centric business strategy and started investing into such technologies and solutions known as customer analytics that are capable of processing huge amount customer data for enhanced decision making. Customer analytics has been of significant importance in most developed economies around the world particularly for large organizations. The off-the-shelf analytics solutions provided by vendors are perceived to be unmanageable, risky and unaffordable especially for Small and Medium Enterprises (SMEs) operating in retail sector. This becomes more vital for the SMEs in developing countries especially in the Eastern part of Europe where they constitute a noteworthy part of the economy. The majority of the SMEs in these countries lack of facilities, infrastructure and abilities to perform such analytical applications. Not being able to extract strategic knowledge using customer data is a missing component for them to be competitive and sustainable in the market from relationship marketing point of view. The aim of this paper is to propose a conceptual model that addresses this problem by providing retail SMEs with a cloud-based open platform for customer data analytics and knowledge extraction. The platform will be able to connect with numerous apps already employed at the retail SMEs, acquire customer data and then perform customer analytics in order to produce a rich set of reports and knowledge.

Keywords: customer analytics, business intelligence, cloud-based computing, knowledge extraction, retail SMEs

1. Introduction

1.1. Use of word-processing software

Customer analytics is a set of technologies that enables companies to acquire the competency of making accurate, timely and effective decisions at all levels regarding customer management processes [1]. It refers to applying various analytics techniques to customer data which may be generated through the internal business processes (the operational data stores) or could be acquired through external and open data sources [2]. The term customer analytics was originated from analytical customer relationship management but it goes beyond its analytical capabilities that are applied in customer management processes. Customer analytics consists of more sophisticated tools and technologies which are capable of processing huge amount of customer data. The main functionalities concerning an analytical customer management solution are associated with four dimensions of customer relationship management, namely customer identification (selection of target customers, customer segmentation/profiling), customer attraction (direct marketing, campaign management), customer retention (customer loyalty including scoring models, one-to-one marketing, complaints management, customer churn modeling, behavioral and clickstream analysis), and customer development (customer lifetime value modeling, up/cross selling, market basket analysis) [1], [3], [4].

The structure of any customer analytics solution consists of four layers. On the bottom, there is the Data Infrastructure layer where customer data and information are stored. On top of this layer there is the Data Integration layer in which the data is turned into information by extraction and integration processes. The next layer is dedicated to more advance processing and reporting and as such it includes Online Analytical Processing cubes: a multi-dimensional view of customer information that allows users to drill data. The fourth and final level of the framework is the analytical environment, which is a domain of business users who use analytical tools to query, report, analyze, mine, visualize and most importantly act on the data [2], [5], [6], [7], [8], [9]

Customer analytics has been of significant importance in most developed economies around the world particularly for large organizations. The off-the-shelf analytics solutions provided by vendors are perceived to be unmanageable, risky and unaffordable especially for SMEs operating in retail sector [10], [11]. This becomes

more vital for the SMEs in developing countries especially in the Eastern part of Europe where they constitute a noteworthy part of the economy.

SMEs form the backbone of EU economy by accounting for 99.8 per cent of non-financial enterprises in 2012, which equates to 20.7 million businesses. In employment terms, SMEs also provide an estimated 67.4 per cent of jobs in the non-financial business economy in 2012. Especially, in wholesale and retail trade, SME employment and added-value are positive and higher than those of large businesses [12]. For example, in Turkey, SMEs corresponds for 78% of employment, 55% of total added-value, 65,5% of total sales, 50% total investments, (2008 data) 59% of total exports (2009 data). Trade and commerce is the SME sector with the highest export and added-value on aggregate in 2007 [13]. The retail sector (organized and unorganized) is expected to reach US$313 billion in 2012 and grow with 10% CAGR until 2016 [14]. The online trade volume in Turkey is booming too. The value of goods sold online has grown 61.1% from 2006 to 2011, predicting online sales in Turkey grow an additional 54% in real terms—that is excluding inflation—between 2012 and 2016 confirming that Turkey's e-commerce sector is one of the fastest growing in the world [15].

The majority of the retail SMEs in the above-mentioned countries lack of facilities, infrastructure and abilities to perform such analytical applications. Moreover, many SMEs lack the skills and the technical / economic capacity to attract the necessary funding in order to make the transformation to a more ICT oriented business cycle. Most of the SMEs in the retail sector rely on a back-office of simple PC-based apps; a few do employ business support systems such as ERPs or other type of information systems. In the majority of cases, data analytics and extracting strategic knowledge based on modern ICT systems is absent, especially within the context of customer management. Also, since most of the existing cloud-based analytics solutions do not fit to firm requirements [16], building a cloud solution specific to sectoral requirements can be of significant importance. The proposed model aims to address these problems by providing retail SMEs with a cloud-based open platform for customer data analytics and knowledge extraction.

The rest of the paper is organized as followings. In Section 2 and 3, recent work on similar frameworks pertaining analytics applications for SMEs as well as related literature regarding the current status of business/customer analytics market for SMEs are provided. Section 4 presents the proposed conceptual model and its components. Finally, Section 5 concludes the study by providing some discussions on potentials benefits of the proposed model.

2. Cloud-based Analytics for SMEs

In today's competitive environment, many SMEs started exploiting innovative BI solutions based cloud computing in order to gain competitive advantage [17]. Cloud BI, aka Software as a Services (SaaS) BI or BI services on-demand, provides IT supporting infrastructure and analytical abilities with excellent scalability, large scale storage, and high performance [18]. NIST (National Institute of Standards and Technology) defines cloud computing as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction" [19].

The related literature presents essential characteristics of cloud computing as followings [17], [19]:

- On-demand self-service (platform with minimal management or self-management),
- Broad network access (accessible via internet from any device),
- Resource pooling (shared, configurable, flexible, dynamic resources),
- Rapid elasticity (virtual, dynamic, scalable and massive infrastructure),
- Measured services (charging based on consumption).

Considering the above-mentioned attributes of cloud computing, it can be argued that cloud technology could be of potential use for SMEs in accessing cheaper analytical solutions. Several operational and financial benefits can be obtained through the use of cloud-based BI for SMEs some of which are [17], [20]:

- Time saving with speed of implementation and deployment
- Lower Total Cost of Ownership
- Elasticity / Flexibility / Agility
- Expertise support
- Focus on core strength
- Standardization of business models
- Quick access and On-demand availability
- Scalability

Although recent studies indicated that most of the organizations would consider cloud computing as a realistic technological option for future, current state of use of cloud-based solutions are in fact limited [17]. The main reason behind that is due to some potential risks or challenges associated with cloud computing which as follows [17], [20], [21] :

- Moving data to the cloud
- Data security and Privacy
- Speed of data access
- Integration with on premise data
- Lack of control
- Vendor maturity
- Reliability of service
- Internet resilience and bandwidth

Existing literature on analytics provide some discussions regarding the architectural and technological infrastructure required for and as well as concrete conceptual frameworks on web- or cloud-based analytics solutions for SMEs. Grabova et al. (2010) [16] made thorough discussions related to limitations of four main BI technologies and presented an architecture which includes relational on-line analytical processing (ROLAP) with java-based ETL. Xu et al. (2009) [18] suggested three layer cloud architecture (business application layer, big cloud-based parallel data mining layer, platform interface layer) for BI applications in telecommunications industry. Liyang et al. (2011) [22] provided more conventional approach to cloud-based BI architecture. Their SaaS BI architecture includes five layers, namely Infrastructure Layer, Data Service Layer, Business Service Layer, User Interface Service Layer, and Operational Service Layer. Wu et al. (2007) [23] also suggested five layer web-based BI architecture which is called Service Oriented Architecture for IT Performance Analytic (SOA-ITPA). The framework consists of data source layer, ETL layer, physical layer, logical layer and analytic application layer.

3. Assessment of Current Analytics Solutions From SMEs' Perspectives

The business analytics software market can be considered as an aggregation of several software tools. Table 1 provides a snapshot of the business analytics software market which has three primary segments: performance management and analytic applications, business intelligence and analytic tools, and data warehouse platform software [24]

Table 1. Business Analytics Market (Source: (Vesset et al., 2012)[24])

<table>
<tr><th colspan="2">PERFORMANCE MANAGEMENT AND ANALYTICS APPLICATIONS</th><th>BUSINESS INTELLIGENCE TOOLS</th></tr>
<tr><td>Financial performance, strategy management, and GRC applications
(Budgeting and planning, consolidation, profitability, and cross-functional GRC)</td><td>CRM analytic applications
(Sales, customer service, contact centre, marketing, web site analytics, and price optimization)</td><td>Query, reporting, and analysis tools
(Dashboards, production reporting, OLAP, and ad hoc query)</td></tr>
<tr><td>Supply chain analytic applications
(procurement, logistics, and manufacturing)</td><td>Service operations analytic applications
(Financial services, education, government, healthcare, communication services etc.)</td><td>Advanced analytics tools
(Data mining and statistics)</td></tr>
<tr><td rowspan="2">Workforce analytic applications</td><td rowspan="2">Production planning analytic applications
(Demand, supply, and production planning)</td><td>Content analysis tools</td></tr>
<tr><td>Spatial information analytics tools</td></tr>
<tr><th colspan="3">DATAWAREHOUSE MANAGEMENT PLATFORM</th></tr>
<tr><td colspan="3">Data warehouse management</td></tr>
<tr><td colspan="3">Data warehouse generation
(Data extraction, transformation, loading; data quality)</td></tr>
</table>

The business analytics market is comprised of more than 100 companies and their associated products. Some of these service providers offer a single application, while some others provide software that spans multiple market segments (12 different market segments). In addition to that, there is a range of business models (e.g., commercial software and open source software) and deployment options (e.g., on-premise and public and private cloud) among these vendors. Though the top 5 business analytics vendors account for approximately 62% of the software revenue, the remaining 38% of the market can be considered as large enough piece of the cake for the small players in this area as shown in Table 2 [24], [25].

Table 2. Worldwide Business Analytics Software Revenue by Leading Vendor, 2009–2011 (Source: (Vesset et al., 2012)[24])

	Revenue ($M)			Share (%)			Growth (%)	
	2009	**2010**	**2011**	**2009**	**2010**	**2011**	**2009-2010**	**2010-2011**
Oracle	4,563.5	5,194.9	6,117.4	18.3	18.7	19.3	13.8	17.8
SAP	3,472.4	3,990.5	4,600.6	14.0	14.4	14.5	14.9	15.3
IBM	3,458.2	3,826.6	4,369.3	13.9	13.8	13.8	10.7	14.2
Microsoft	1,870.7	2,110.6	2,349.7	7.5	7.6	7.4	12.8	11.3
SAS	1,898.2	2,007.5	2,263.2	7.6	7.2	7.1	5.8	12.7
Other	9,624.3	10,655.1	12,002.2	38.7	38.3	37.9	10.7	12.6

Total	24,887.3	27,785.2	31,702.4	100.0	100.0	100.0	11.6	14.1

Recently, organizations with less business analytics experience, particularly SMEs, have become interested in analytics solutions. The majority of SMEs perform their business decisions through the managers' gut feelings and based on a combination of static historical reports generated via spreadsheets [11]. As companies continue to grow or face fierce competition, the necessity of making decisions that are based on meaningful information has become an imperative [26]. However, penetration of analytic applications into SMEs market is low because of the following reasons [2], [5], [10], [26], [27], [28], [29], [30]:

- SMEs are being reluctant on giving up spreadsheets
- Leaders are not fully familiar-with/convinced-about the return on investment (ROI)
- Leaders are aware of ROI, but end-users are not familiar with right tools or end-users are unable to interpret the recommendations from the analytics and translate them into business decisions.
- Tight ICT budgets, high prices and Time implementation consideration
- Lack of sophistication (limited ICT expertise), personnel and infrastructure
- Low flexibility and limited number of solutions available

Due to the reasons stated above an external help to SMEs is certainly needed. But, the majority of the vendors and their solution partners usually use a time-and-materials approach to deploy analytics solutions to SMEs. This could be affordable and suitable for large enterprises but it does not fit SMEs with smaller budgets. Also, equally important is the fact that the buyers of analytics solutions are mainly less tech-savvy managers and they are looking for simpler, less technology-centred solutions. Pre-built solutions delivered by the players in the market via the traditional service-oriented approach could be considered an alternative for SMEs. However, such an approach is too expensive for SMEs and it is neither scalable nor repeatable [29]. For an SME, successful business analytics implementation relies on four basic stages: information/data; technology; intelligence; and implementation and communication [26]:

1. With regards to information/data, one could argue that the smaller the size of the company the less data is available. Compared to the data available in an ERP system, SMEs possess a basic entry-level accounting program that stores substantially less information. Most SMEs usually manage their key business data in spreadsheets, contact manager databases, payroll systems, and other home-grown databases.
2. In terms of the technology aspect, the software solutions that can deliver analytics applications are beyond the financial reach of SMEs. Therefore, they usually tend to go for very cheap and simple solutions in the form of spreadsheets. Ubiquitous spreadsheet packages like Excel are undoubtedly valuable and powerful tools, but this approach possesses some limitations and dangers (dependency, inflexibility, insecurity, complexity) [10]. Another technological solution for SMEs is the packaged solutions provided via cloud computing. Although some of the products delivered via the software-as-a-service model could be considered as an alternative solution to SMEs, but much cheaper versions of such technology should be made available to SMEs.
3. As far as the intelligence aspect is concerned, SMEs face exactly the same problem as large companies do in determining the relevant data and metrics. Being overwhelmed by the data and having difficulty to focus on what is truly important are also unavoidable for SMEs mainly due to the amount of available data in the operational data stores and not having the relevant trained staff. Therefore, an external help is needed in defining critical success factors and the data/metrics for a typical SME.
4. In the majority of analytics applications, the crucial element frequently missing is the communication aspect (explaining to the users what the analytics mean, how to interpret them, and what actions to take). Therefore, a decision support system platform could be of potential use in an analytics solutions tailored to SMEs.

The discussion above highlights that having such analytics applications accessible to SMEs should be within their financial capacity and horizon, while also the solutions provided should overcome the difficulties of adapting the analytics pertaining to SMEs by taking into account the following critical factors, including aligning business and IT, time concern, ease of use, reducing total ownership cost, flexibility and adaptability [5]. In conclusion, it is believed that the business analytics market is still dominated by traditional on-premises solutions, therefore, utilizing key forces like cloud computing in an adaptive and cheaper way will play a key role in increased adoption of the analytics for SMEs.

4. Proposed Conceptual Model

While in-house business analytics solutions are common in Western and Northern Europe, the situation is totally different, if one takes a look at the South East end of Europe. Economies in that area of Europe are largely SME-based but a sizeable percentage of these SMEs have a limited capability of exploiting business analytics in

order to improve their business performance. Part of the explanation for this phenomenon is attributed to cultural and historical reasons. Another important reason is the lack of effective funding; SMEs either decide to invest elsewhere or operate on very low margins, thus being unable to afford an ICT back-office in general (despite the generous funding from European Union sources). This situation is evidently a very serious obstacle for growth with respect to SMEs and SME-based economies in this region. The proposed model addresses this problem by moving the business analytics infrastructure out of the SME environment and into the cloud. This removes the requirement for dedicated hardware / software and combined with an affordable pricing scheme, significantly lowers the barrier of entry for medium-sized/small players. The model will provide SMEs with business analytics (sales, marketing and customer analytics) and reporting/knowledge extraction components for medium and long term business strategies. The model will be compatible with several business data formats and sources (e.g. MS Office files, XML-formatted files, ERP systems etc.) and will be equipped with the appropriate software tools in order for SME-derived data to be seamlessly loaded onto the platform for further processing.

The architectural design of the model is based on a complete data life cycle that starts with the acquisition of raw data and ends with business analytics reporting as depicted in Figure 1. In terms of system architecture, the model will implement, to the furthest extent possible, a modular architecture consisting of loosely coupled, self-contained software components. This approach guarantees portability, ease of integration and modification of the platform. The main components of the model are presented as the following:

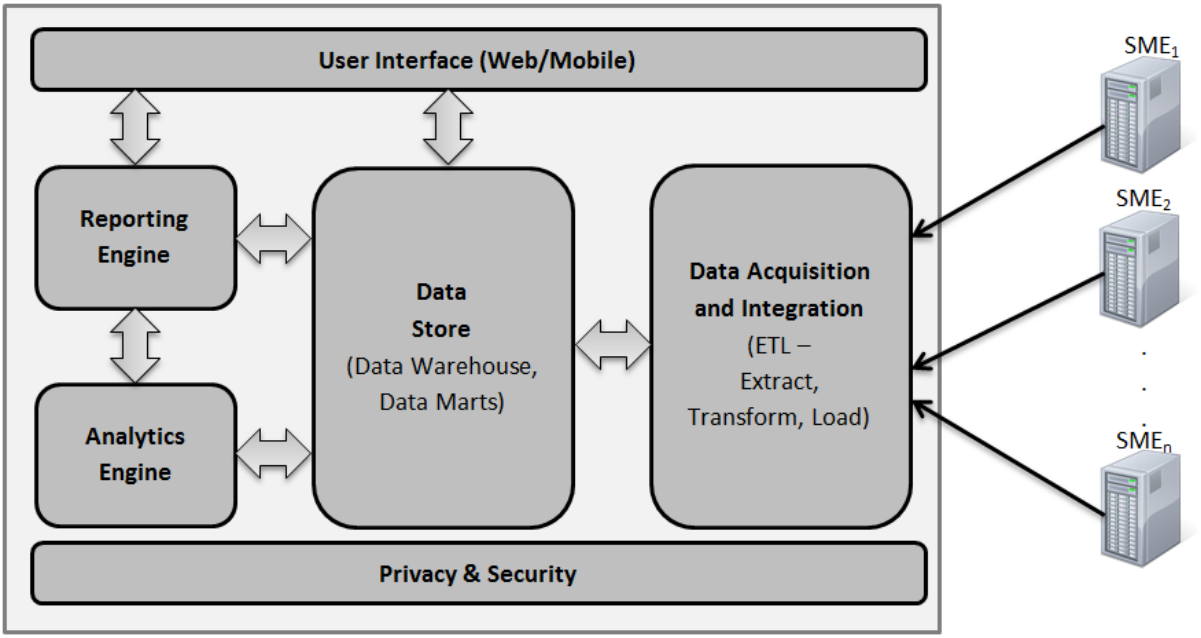


Figure 1. Architectural Design of the Proposed Model

Data Acquisition and Integration: These are services that would be used for the acquisition of input data from the retail SMEs business systems. As such, they are tailored to the interfaces of each external business system e.g. XML, ERPs, Office files etc. Evidently, original input data exists in several formats, depending on the originating business system, and its structure resides in several languages, depending on the country/countries the retail SME resides in. This component is supplied with input data from the various business connectors and transforms it, incorporating also semantic mediation techniques, to a common data schema that can be used for further processing.

Data Store: This component is responsible for the persistence of input, intermediate and output data. The exact nature of the data store (SQL, NoSQL, OWL, RDFS, etc.) must be decided based on requirements analysis.

Analytics Engine: This is the component that is responsible for processing the input data and producing analytics information and knowledge such as e.g. sales and market analysis, segmentation. The analytical component of the model includes the following specific modules.

Sales Performance Analytics Module: To be able to construct this module the followings must be identified: the common sales performance metrics, the target data set which includes the necessary data fields required for these metrics, some pre-processing activities (validation, aggregation, reduction, transformation, consolidation, and etc.) and their logical principles, the mathematical calculations needed to come up these metrics, determination of programming needs and procedures.

Some data pre-processing technologies, tools and principles regarding data preparation as well as potential data exploration methods based on mathematical and statistical models (eg., univariate and multi-variate analyses, linear and logistics regression) must also be utilized within this module.

Market/Sales Intelligence Module: This module includes several forecasting models for the aim of extracting marketing and sales intelligence information/knowledge. Therefore, after determining the types of information to be extracted, the necessary variables associated with each type of knowledge-domain must be specified. Through the identification of dependent and independent variables for each knowledge-domain several forecasting and econometric models can be employed.

Market Basket Analysis Module: In order to create this module some pre-processing activities on the consolidated data mart to make the data ready for market basket analysis must be carried out. The module also utilizes a robust and commonly used data mining based association rules technique (the Apriori algorithm) in order to carry out market basket analysis. Therefore, a user-defined flexible rule extraction design must be accomplished via taking into account different parameters (eg., confidence and support levels of the rules) of the algorithm.

Customer Analytics Module: The activities regarding this module have two folds. The first group of activities would ensure the most flexible and comprehensive target data selection process for conducting the analyses particularly customer segmentation, customer lifetime value calculation and customer churn modeling. Therefore, potential data preparation and exploration technologies, tools and principles within the context of data mining and knowledge discovery can be utilized. The second category of activities include the followings: (1) identification of segmentation, lifetime value and churn models and the variables (or the bases) required to perform each type of segmentation, lifetime value, and churn model; (2) determination and design/programming procedures of data mining based clustering, classification and prediction algorithms (including k-means, fuzzy c-means, self-organizing maps, back-propagation, logistic regression) needed to carry out the selected segmentation and churn models; (3) design/programming procedures of logical principles and the mathematical calculations of each selected lifetime value model; (4) determination of the criteria required for the evaluation and validity assessment of the selected segmentation, lifetime value and churn models and design their mathematical calculations.

Reporting Engine: This component properly combines the derived knowledge from the SME originated business data in order to produce custom reports that concern long term strategic decisions for SMEs (i.e. What will be the sales projections? Which sales channels have better performance? Which customers are more profitable or valuable? What are the main customer segments? What are the profiles of each segments? Should I tailor a pricing model for each segment? etc.).

User Interface (UI): The model will feature a web-based front-end based on responsive design principles in order to be compatible with desktop and mobile terminals. The main role of UI is to allow SMEs to benefit the business analytics reports.

Privacy and Security Component: One of the major issues concerning the platform is how to convince SMEs that it would be safe to send their business data to an external platform for further processing. There are open issues concerning data theft, data corruption, as well as trade secrets. A minimum set of requirements would most certainly have to offer secure authentication / authorization, secure connections, encrypted databases, as well as, an anonymization process that would make it very difficult for a 3rd party to associate a set of data with a specific SME.

5. Conclusions and Discussions

The proposed model aims to provide retail SMEs with a cloud-based open platform for customer data analytics by contributing them in acquiring the access to competences and resources that they need to develop innovative content and data analytics services. The model connects with numerous apps already employed at SMEs, acquire customer data and then perform analysis processes in order to produce a rich set of customer analytics reports. Considering the fact that most cloud computing-based BI tools may not be suitable for every organization, the proposed model provides a customized solution for retail SMEs within context of analytical processing and reporting regarding customer management.

The suggested technological architecture and business model will enable the sustainability and wider use of the model outcomes. The cloud model, for instance can provide a pay per use option for SMEs who can become register and as a member benefit from the provided services for a low and affordable price. The related bodies such as retailers, associations and development agencies can be contacted for ensuring the outreach and sustainability of the model outcomes. The sustainability of the model depends on the assumption that either a financially affordable pay per use business model will be facilitated to SMEs or an external fund is to be obtained from the related governmental institutions, retail/SME associations and unions.

References

[1] Schroeck, M. J. (2001). Customer analytics making the difference in CRM: Customer analytics amplify the value of integrated CRM solutions. DM Review, September.

[2] Bose, R. (2009). Advanced analytics: opportunities and challenges. Industrial Management & Data Systems, 109(2), 155-172.

[3] Srivastava, J., Wang, J. H., Lim, E. P., & Hwang, S. Y. (2002). A case for analytical customer relationship management. In Advances in Knowledge Discovery and Data Mining (pp. 14-27). Springer Berlin Heidelberg.

[4] Ngai, E. W., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. Expert Systems with Applications, 36(2), 2592-2602.

[5] Suleymanov, O., "Business Intelligence Strategy: Business Intelligence for Small and Middle Size Enterprises", http://www.slideshare.net/olimjons/business-intelligence-strategy-for-smes [Accessed on 30/11/2012].

[6] Turban, E., Sharda, R., Delen, D., King, D. (2011), "Business Intelligence: A Managerial Approach", Prentice Hall, Boston.

[7] Vercellis, C. (2009), "Business Intelligence: Data Mining and Optimization for Decision Making", Wiley, UK.

[8] Ranjan, J. (2009). Business intelligence: Concepts, components, techniques and benefits. Journal of Theoretical and Applied Information Technology, 9(1), 60-70.

[9] Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. Communications of the ACM, 54(8), 88-98.

[10] Olszak, C. M., & Ziemba, E. (2012). Critical Success Factors for Implementing Business Intelligence Systems in Small and Medium Enterprises on the Example of Upper Silesia, Poland. Interdisciplinary Journal of Information, Knowledge, and Management, 7.

[11] Fotache, M. V., & Fotache, G., (2005). The Economic Recovery of the SME's by Implementing BI Technologies.

[12] EU SMEs in 2012: at the crossroads - Annual report on small and medium-sized enterprises in the EU, (2012). http://ec.europa.eu/enterprise /policies/sme/facts-figures-analysis/performance-review/files/supporting documents/2012/annual-report_en.pdf [Accessed on 30/11/2012]

[13] Turkey SME Strategy and Action plan 2011-2013, (2011). http://www.sanayi.gov.tr /Files/Documents/KOSGEB_Katalog.pdf [Accessed on 30/11/2012]

[14] Deloitte: Retail Sector Update, (2012). http://www.deloitte.com/assets/Dcom-Turkey/Local%20Assets/Documents/turkey_tr_cb_retailsectorupdate_040612.pdf, [Accessed on 30/11/2012]

[15] SBA Fact Sheet Turkey, (2011). http://ec.europa.eu/enterprise/policies/sme/facts-figures-analysis/performance-review/files/countries-sheets/2010-2011/turkey_en.pdf, [Accessed on 30/11/2012]

[16] Grabova, O., Darmont, J., Chauchat, J. H., & Zolotaryova, I. (2010). Business intelligence for small and middle-sized entreprises. ACM SIGMOD Record, 39(2), 39-50.

[17] Mircea, M., Ghilic-Micu, B., & Stoica, M. (2011). Combining business intelligence with cloud computing to delivery agility in actual economy. Journal of Economic Computation and Economic Cybernetics Studies, 45(1), 39-54.

[18] Xu, M., Gao, D., Deng, C., Luo, Z., & Sun, S. (2009). Cloud computing boosts business intelligence of telecommunication industry. In Cloud Computing (pp. 224-231). Springer Berlin Heidelberg.

[19] Mell, P., & Grance, T. (2011). The NIST definition of cloud computing (draft). NIST special publication, 800, 145.

[20] Gurjar, Y. S., & Rathore, V. S. (2013) Cloud Business Intelligence–Is What Business Need Today. International Journal Of Recent Technology and Engineering, Vol 1 (6),

[21] Horakova, M., & Skalska, H. (2013). Business Intelligence and Implementation in a Small Enterprise. Journal of Systems Integration, 4(2), 50-61.

[22] Liyang, T., Zhiwei, N., Zhangjun, W., & Li, W. (2011). A conceptual framework for business intelligence as a service (saas bi). In Intelligent Computation Technology and Automation (ICICTA), 2011 International Conference on (Vol. 2, pp. 1025-1028). IEEE.

[23] Wu, L., Barash, G., & Bartolini, C. (2007). A service-oriented architecture for business intelligence. In Service-Oriented Computing and Applications, 2007. SOCA'07. IEEE International Conference on (pp. 279-285). IEEE.

[24] Vesset, D., Wardley, M., McDonough, B., Schubmehl, D. (2012), "Worldwide Business Analytics Software 2012-2016 Forecast and 2011 Vendor Shares", IDC Market Analysis.

[25] Gartner Press Releases on the Future of Business Intelligence: http://www.gartner.com/it/page.jsp?id=1971516 [Accessed on 30/11/2012].

[26] Blue Link Associates Ltd. White Paper on Business Intelligence for the SME http://www.bluelinkerp.com/newsroom/whitepapers/BusinessIntelligenceforSME_whitepaper.asp [Accessed on 30/11/2012].

[27] Predicting Future Outcomes With Business Analytics: http://busification.com/business-analytics-for-smes/ [Accessed on 30/11/2012].

[28] 3 barriers for wide adoption of business analytics in SMEs: http://www.simafore.com /blog/bid/62628/3-barriers-for-wide-adoption-of-business-analytics-in-SMEs [Accessed on 30/11/2012].

[29] SAP Report on Business Analytics for SMEs: www.sap.com [Accessed on 30/11/2012].

[30] Thompson, W. J., & Van Der Walt, J. S. (2010). Business intelligence in the cloud. SA Journal of Information Management, 12(1), 5-pages.

ISSN 2303-4521
PERIODICALS OF ENGINEERING AND NATURAL SCIENCES
Vol. 1 No. 2 (2013)
Available online at: http://pen.ius.edu.ba

Science-Technology Park Ilidža as a Generator of Innovation Potential and SME's Development in Bosnia and Herzegovina

Darko Petković
Faculty of Mechanical Engineering
University of Zenica
centar_mkr@mf.unze.ba

Hazim Bašić
Faculty of Mechanical Engineering
University of Sarajevo
basic@mef.unsa.ba

Benjamin Duraković
International University Sarajevo
Faculty of Engineering
durben2@yahoo.com

Sanja Prodanović
Faculty of Economics East Sarajevo
University of East Sarajevo
sanja.prodanovicc@gmail.com

Abstract
Many nations are currently adopting a variety of directed strategies to launch and support the development of research parks. Science and technology research parks are seen increasingly as a means to create dynamic clusters that accelerate economic growth and international competitiveness. Technology parks represent the place of SME's development and the place of the most direct technological transfer. Research parks encourage greater collaboration among universities, research laboratories, and SMEs, providing a means to help convert new ideas into the innovative technologies for the market. They are recognized to be a tool to create successful new companies, sustain them and attract new ones, especially in the science, technology, and innovation sector. Universities, in turn, benefit by exposure to the business world. What all these parks have in common is that they are, at heart, knowledge partnerships that foster innovation. University research and science parks provide the launch pad that startup companies need when they are "spun out" from a university or company. The current state of development of technology parks in Bosnia and Herzegovina and possibilities of establishment of a new technology park in Sarajevo (Ilidža) region are analyzed in this paper.

Key words: *Science-technology parks, technology transfer, entrepreneurial, innovations.*

1. Introduction

The Western Balkan region faces complex structural and macroeconomic challenges. On the structural front, a legacy of unfinished reforms still prevents most economies from reaping the benefits of investments in research and innovation — despite the progress they have achieved in recent years. At the macroeconomic level, adverse circumstances required governments to adopt tighter fiscal policies and cautious financial markets reduced public and private resources available for research and innovation. The countries share a common set of problems, such as low business expenditures on R&D, outdated research infrastructure and brain drain.

Many development documents at the state level of Bosnia and Herzegovina (B&H), and the Sarajevo Canton emphasize the need to improve development and innovative activity of SME's, [1, 4, 12, 13]. For example, at the state level, the strategy for the development of small and medium enterprises in B&H for the 2009-2011 year and the Strategy for Development of Science in B&H (2009) were adopted. The Government of the Federation of Bosnia and Herzegovina (F B&H), has adopted the Strategy for Development entrepreneurship and craftsmanship in F B&H 2010-2020, and the Strategy of development of science and research work, which was adopted 2012th year. At the level of the Sarajevo Canton, The study of economic activity of the Sarajevo Canton of 2009th year should be emphasized [12].

In this development documents the importance of SMEs in the economic development and provide for measures to encourage their development are emphasized. Among other activities, on the significant position are the measures of improving technical and technological level of production and improve the innovation capacity of SME's. It is anticipated that the research relies partly on the resources of the Fund for Science and Research, and the other part of the company assets.

The general assessment is that all these documents and studies declaratory support to SMEs, but measures for their

DOI: 10.21533/pen.v1i2.25

implementation absent. The main reasons for this are the lack of financial resources, but also the absence of a strategic approach to the management of small and medium-sized enterprises at different levels.

A majority of EU Member States are in the process of improving their governance structures and strategic guidance for research and innovation, which is often accompanied by a gradual integration of the two policy fields and increased targeting of public funding on selected areas. Business environment and economic recovery are in direct correlation, Figure 1, [2].

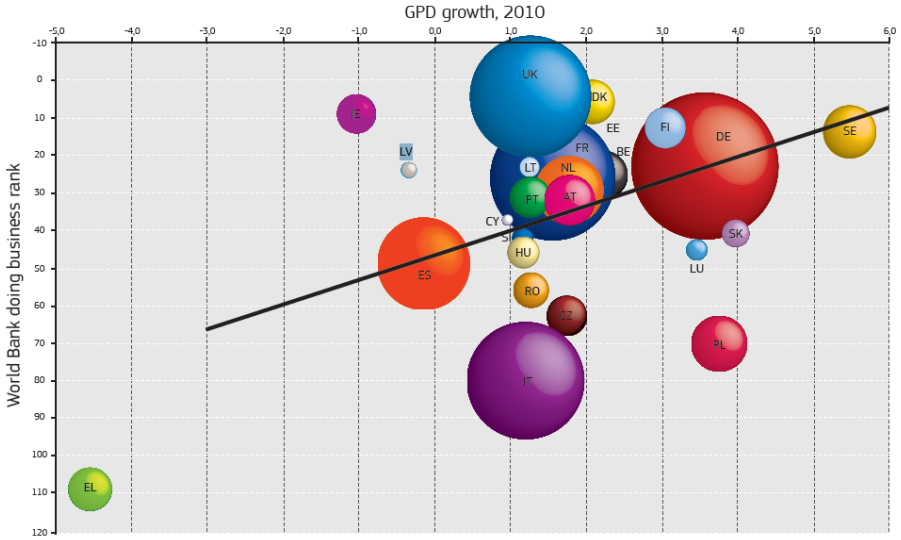


Fig. 1: Business environment and economic recovery, [2].

Establishing of technology /scientific parks (TP/SP) is a faster way to overcome the weakness and inertia of state structures, [7, 10].

According to IASP, "A TP/SP is an organization, managed by specialized professionals, whose main objective is to increase the wealth of its community by promoting the culture of innovation and the competitiveness of its associated businesses and knowledge-based institutions." To reach these goals a TP/SP stimulates and manages the flow of knowledge and technology amongst universities, R&D institutions, companies and markets; facilities the creation and growth of innovation-based companies through incubation and spin-off processes, and provides other value added services as well as high quality space and facilities.

2. Methodology

The EU noted that SMEs are dynamic source of employment, growth and competitiveness, but if care is taken that the authorities must develop a comprehensive strategy to support these businesses, which includes strategy of support technology and innovation, [3, 6, 11].

Research and development activities can play a very important role in regional development by providing a knowledge base that supports innovation. This is the most important way in which universities can contribute to the development of a particular region by the "translation" of their research in the form in which it can be downloaded and implemented by the private and public sectors.

It is important to recognize that university research can be involved in the process of regional development, not only through the support of innovation that begins with research, but also as a result of the reported demand by the regional public and private sector.

An overview of high-tech exports as % of total exports to the countries of the Western Balkans is given on Figure 2. It is apparent lagging behind B&H, which is the result of almost no connection between higher education and industry. The above suggests that the contribution of university research activities to the regional innovation must be treated as a multidi-sciplinary and iterative process.

Consideration of the present state of the industry in the Sarajevo Canton and the F B&H is not possible without taking into account the situation in the area before the war. It is clear that the region of Sarajevo was the most developed area in B&H, and the city at the time was the seat of great economic systems (ENERGOINVEST, FAMOS, UNIONINVEST, UNIS, ZRAK, etc.), which were generators and holders of economic activities and development in the most industries. The current state of the industry in the Sarajevo Canton is a result of general negative climate for economic development and incomplete or incorrect privatization process.

The privatization of large systems is not brought to an end. These processes are regularly accompanied by politicization instead of planned actions to restructure the company and change the character of the ownership.

The importance of ideas that are analyzed in this paper should be placed in the context of the industry in the Sarajevo Canton. After a difficult transition period it is now restored and trying to develop on the basis of inherited production and new companies that are appear in this sector. Sometimes extremely developed industrial sector in Canton Sarajevo, the war is degraded in two ways:

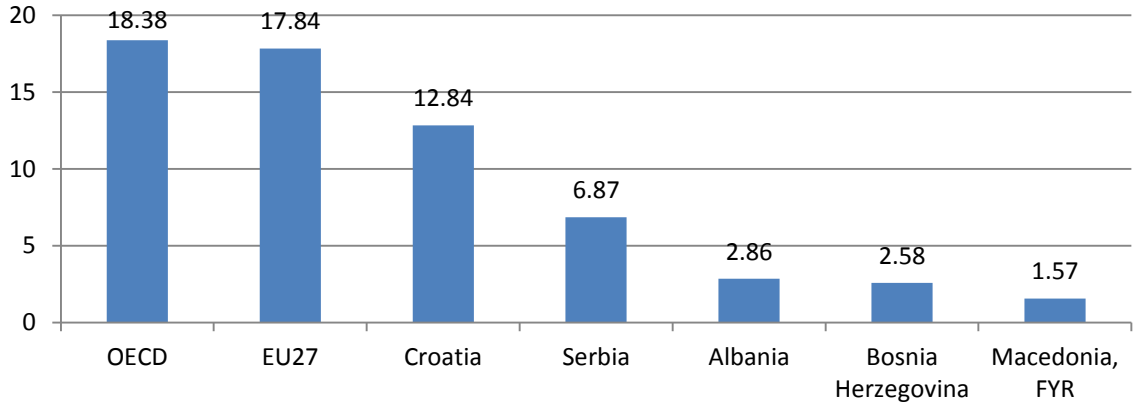


Fig. 2. High-tech exports (% of total commodity exports) 2009, (Source: UN Comtrade).

1. Through war destruction of research and development institutions that existed in large enterprises. These institutions were separate legal entities and arrange the research and development projects with legal entities within large companies.
2. Through lack of updates in research and development institutions and their customers' research and development services.

Policy in Bosnia and Herzegovina was to continue thinking only to small and medium enterprises (SMEs), which are formed mainly in the retail sector. Later, when SMEs began some kind of production, they are quite low-tech aspirations. Within themselves are not formed any research and development groups, or research and development centers, they ordered research and development projects to develop new products, technologies and services, either in the form of external R&D organizations.

Chambers of commerce that are usually main founders of technology parks (as ambient for cheap R&D activities and technology transfer, [8, 9, 10], have not formed any technological park in Bosnia and Herzegovina. There are only a few TP/SP established in B&H at this time: STP INTERA, [16], in Mostar (the Herzegovina region), TP of University of Zenica (the central B&H region) and TP Tuzla (the north-east B&H), [12]. Therefore, the formation of a technology park in Sarajevo Economic Region, in the municipality of Ilidza, should fill the existing gap.

3. Results

Almost complete destruction of the once great systems through improperly conducted privatization process has destroyed almost 90% of the manufacturing base. So they collapsed logically complex forms of clusters only for arrival in possession of a property. That's what today is best witnessed by the large industrial sites mentioned and other companies.

By introducing the principle of entity privatization and corporate bond rupture within the once great systems has led to further fragmentation of clusters, a division once a single economic space, and the consequence of the impossibility of creating a national strategy for reconstruction and development.

Most of the actions carried out last 10 years, nor result in the development of the industrial sector, but a change of ownership of the property and not on business.

This process was accompanied by all the phenomena of transition that are social, ethnic, ethical and psychological character. This is further deformed and destroyed the human potential of the former base of industrial workers. If you add the factor of elapsed time (over 20 years), it can be concluded that the greatest benefit of the last remaining time tradition, yet vital references and little remaining capacity and experienced personnel. At the same time, these are the main characteristics of Sarajevo region and Ilidza municipality.

The Sarajevo economic region properties

Sarajevo economic region is a unique natural, geographical, cultural, historical, economic and communications area. Covers an area of 32 municipalities, of which 13 are in the Republika Srpska (RS) and 19 in the Federation of Bosnia and Herzegovina. The total area of the region is 8699,9 km^2, of which the municipality of FB&H refers 3531,9 km^2, and the community / municipality from RS 5,168 km^2, [4].

According to the last census (1991st year), in the region's 838,216 inhabitants lived in 254,669 households, with an average number of household members of 3,3 members.

The average population density is 97,3 inhabitants per km^2. In urban areas lived 64.9% of total population.

Total in the region during the 2010th year lived 738,840 inhabitants, of which 79% of the municipalities in the F B&H and 21% in the municipalities of RS, [4].

The population of the region accounts for about 19,2% of the total population of B&H. Sarajevo Canton has the largest share in the total population of the region (59,1%), followed by the municipality of RS (20,8%).

In the region exists two of public higher education institutions (University of Sarajevo and University of East Sarajevo), and three private universities (Sarajevo School of Science and

Technology - SSST, International University in Sarajevo, The Burch University) are placed.

Forests and timber represent one of the most important natural resources of the region. Around 4477,44 km^2 or 51,7% of the region is covered by forests. There are rich mineral deposits, and the region also has a huge capacity of drinking, thermal and mineral waters.

In the region there are 46 large and small rivers with a total length of flow through the region of around 979 km. At the region existed 18 major or minor natural or artificial lakes where fishing is possible and / or sports tourism, [12].

Important hydropower potential of the area, especially the Drina river and its tributaries where conditions for the construction of mini, small and medium-sized hydro-power plants exists.

Demographic growth of Bosnian capital and its surroundings creates the need for new investment and jobs. It is certain that the sectors of tourism, hospitality, telecommunications, information techno-logy, banking and continues to be the focus of development policies.

When viewing the Sarajevo Canton, reveals that companies in the industrial sector, SMEs, some of them established after the war and some were made from the composition of the former large systems.

In 2011 GDP estimates FB&H per capita is 6,821 KM, Figure 3, GDP per capita compared to the average of the FB&H is the largest in the Sarajevo Canton, [12].

If cantonal government wants to develop the Sarajevo Canton economy, then it must develop industry capacities and put the development of SMEs on the key place. Highly developed countries have a very large share of industry in total domestic economy. It is very typical when looking through exports. Thus, for example, in the structure of Japan's total exports, 60-66% are industrial products, USA 40%, while the EU countries around 40%.

However, examples of countries with a high level of real economy (China, Turkey, Germany, Poland, USA) show how the production as a generator of added value unsurpassed growth and development factor in times of crisis and normal economic situation, [11,14,15]. Sarajevo region with its professional potential, openness, communication, confidence in supplying energy, sites and traditions will be attractive for investment both domestic and foreign investors.

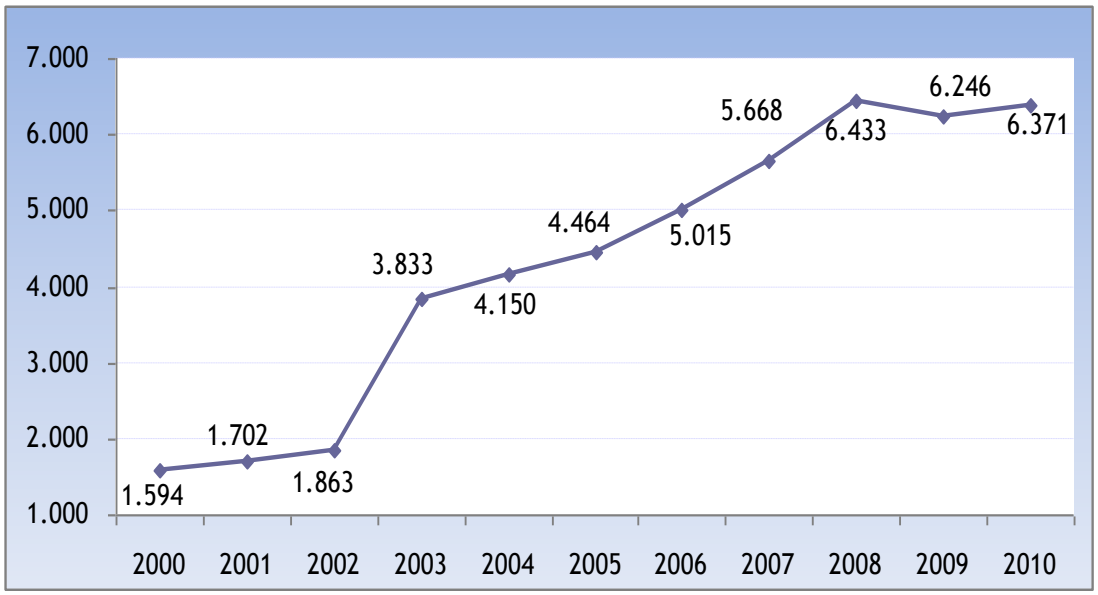


Figure 3. BDP per capita, 2000-2010 in KM, (Source: Agencija za statistiku B&H, 2012).

4. Discussion

Bosnia and Herzegovina does not have the capacity (organizational, political, institutional and financial) which is necessary for the process of restructuring of the system of scientific and technological development. It is, therefore, necessary to define the research and development as a priority of future interventions of the international community in B&H. SMEs require systematic support to services and technology transfer, technical cooperation and development, innovation support, etc.

The situation with regard to the development of technology or science-technology parks in Sarajevo region and B&H requires strategic definition of their development, government assistance in the establishment and profiling, equipment, registration and other activities, on the model of technology parks in the world .

Scientific parks and research and technology centers require significant capital investment, but the activities that are carried out within them and that have the potential for realizing the transformational impact, [5,7]. They provide help universities to make them closer to market commercialization and bridging the gap between research and its applications.

The founders of the technology park can be universities, associations of municipalities, regional developmental agencies, individual companies or local levels of government. The primary task of a technology park is to generate new small

and medium enterprises through the involvement of university staff to solve the problems of these companies.

Final goal is to create sustainable economic space in which to create jobs and promote investment opportunities.

The basic preconditions for the establishment of a technology park Ilidža already exist: the existence of several higher education institutions in Sarajevo region and especially in the Ilidža municipality, the strong tradition of industrial production in region and municipality Ilidža, land, human resources. After the initial phase of the existence of the technology park, the expected benefit would soon overcome regional frameworks.

References

[1] Ekonomski institut Sarajevo (2006), *Studija izvodljivosti u cilju formiranja industrijskih zona SMR*, Sarajevo.

[2] *EU Member States and Associated countries Innovation Union progress at country level 2013*, European Commission, Directorate-General for Research and Innovation, Brussels, 2013. (http://ec.europa.eu/research/innovation-union/pdf/state-of-the-union/2012/innovation_union_progress_at_country_level_2013.pdf#view=fit&pagemode=none)

[3] EU (2011): "*Connecting Universities to Regional Growth: A Practical Guide*", http://ec.europa.eu/regional_policy/sources/docgener/presenta/universities2011/universities2011_en.pdf.

[4] Federalni zavod za planiranje razvoja (2011), *Makroekonomski pokazatelji po kantonima u Federaciji Bosne i Hercegovine u 2011*., Sarajevo.

[5] Mrkonjic Z., Kostic M., Krezic D., Petković D. (2008), *''Technology park Mostar: from idea to realisation*'', 12th International Research/Expert Conference "Trends in the Development of Machinery and Associated Technology", TMT 2008, Istanbul, Turkey.

[6] OECD (2009), "*SME Policy Index 2009 - Progress in the Implementation of the European Charter for Small Enterprises in the Western Balkans''*, Paris.

[7] Petković D. (2004), "*Technology Parks in the Function of Entrepeneur Development and Growth of SME's''*, Int. Conference "Economic Recovery of Serbia through Development of Science Parks", Belgrade.

[8] Petković D. (2006), *''Tehnološki parkovi – Više od mjesta za tehnološki transfer i razvoj poduzetništva''*, Univerzitet u Zenici, Zenica.

[9] Petković D., et. al. (2003), "*Projekt naučno-tehnološkog parka u ZE-DO Kantonu''*, Mašinski fakultet u Zenici, Centar za menadžment, kvalitet i razvoj, Zenica.

[10] Porter, E. Michael, (1990): *The Competitive Advantage of Nations*, The Free Press, A Division of Macmillan, Inc., New York.

[11] Roxburgh, C. & Mischke, J., (2011), "*European growth and renewal: The path from crisis to recovery''*, London: McKnsey Global Institute.

[12] SERDA (2004), *Socio-ekonomska analiza Makroregije Sarajevo*, Sarajevo.

[13] Vijeće ministara BiH (2009), Strategija razvoja nauke u BiH 2010-2015, Sarajevo.

[14] Ziegenbalg C., et. al. (2011), "*Comprehensive analysis of programmes and initiatives in Bosnia and Herzegovina that assist the Collaboration between science and SME''*, The Report for the European Commission, Research Directorate-General, under FP7-SME- 2009-1, SP4-Capacities.

[15] World Bank (2011), *''Innovation Policy - A Guide for Developing Countries''*, New York.

[16] http://www.intera.ba/

[17]

ISSN 2303-4521
PERIODICALS OF ENGINEERING AND NATURAL SCIENCES
Vol. 1 No. 2 (2013)
Available online at: http://pen.ius.edu.ba

Hot Tear-Melt Quality Relationship in 3xxx Aluminium Alloys

Murat ÇOLAK
Department of Metal Education, Faculty of Technical Education, Sakarya University,
mcolak@sakarya.edu.tr

Engin TAN
Department of Material Science and Engineering, Faculty of Technology
Pamukkale University
etan@pau.edu.tr

Ramazan KAYIKCI
Department of Metallurgy and Materials Science Engineering
Sakarya University
rkayikci@sakarya.edu.tr

Derya DIŞPINAR
Department of Metallurgy and Materials Science Engineering
İstanbul University
deryad@sakarya.edu

Abstract:

Corrosion resistance, formability and high strength is few of the major properties that makes aluminium alloys preferred choice of material for extrusion or deep drawing processes. Particularly, buildings, windows, doors are main application areas and it is growing. Twin roll casting method is used to produce such materials. One of the common problems faced in this method is called hot tearing. The effect of alloying elements is the reason for this defect. As a result, segregation causes tear and decrease process efficiency. The economic impact is immanent. Thus, it is aimed to study the hot tear tendency of 3xxx series aluminium alloys and the effect of alloying elements will be investigated in this work.

Keywords: Hot-tear, 3xxx series, Aluminium alloys, melt quality, casting simulation.

1.Introduction

It is quite common to have defects in any sort of production methods. Hot tearing is one of the defects that occur during casting processes. The size and distribution of hot tears may be up to 30 vol % of the cast part. During solidification, certain locations on a cast part may freeze later than the surroundings. Thus, these areas will be under negative pressure due to the shrinkage. Depending on the tension generated, this remaining liquid may "tear" to form the casting defect called hot tear. Hot tears may even lead up to the distortion of the cast part. The variables that affect hot tearing can be mainly listed as follows: composition, mould design, filling conditions, process parameters etc.

Campbell [1] pointed out that these defects were associated with hydrostatic tensions in the dendritic network. Therefore many research have been carried out by using Darcy's Law to predict the formation of hot tear in castings [2-20]. However, hot tearing still remains a complex phenomenon. Morever, there is no single type of a test that can measure or standardize these defects. Campbell's model for hot tearing is given in Fig. 2.

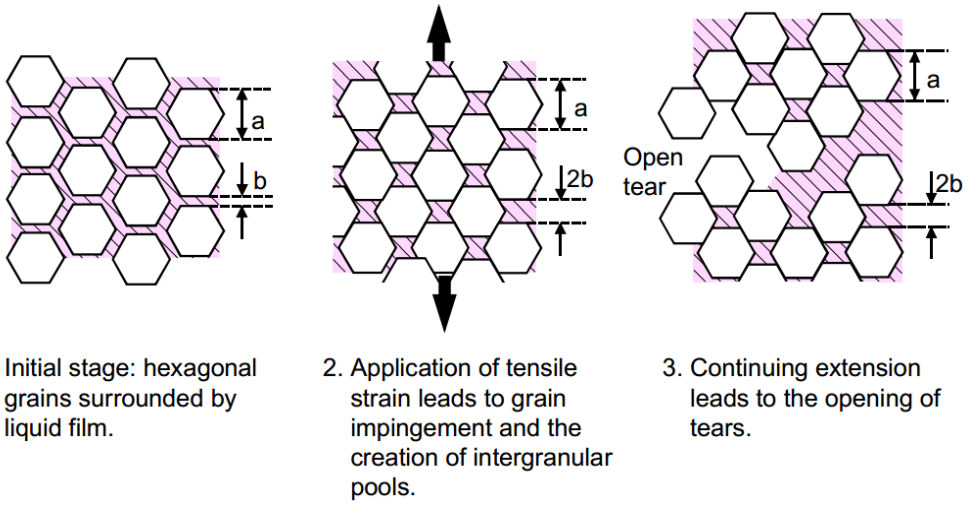


Fig. 1: A simple model of hexagonal grains of diameter "a" separated by a liquid film which initially has a thickness of "b" [1]

DOI: 10.21533/pen.v1i2.26

Therefore, in this study, two different mould designs were used to investigate the hot tearing susceptibility of 3xxx series aluminum alloys by using SolidCast simulation. The contribution of melt quality issue was also introduced into the hot tearing phenomena.

2. Experimental Work

Commercially available 3105 alloy was used to model the hot tearing susceptibility. The composition of the alloy is given in Table 1.

Table 1. Chemical Analysis of A3105 alloy

Al	Cu	Fe	Si	Zn	Mg	Mn	Ti	Cr
Balance	0,30	0,70	0,60	0,40	0,60	0,50	0,10	0,10

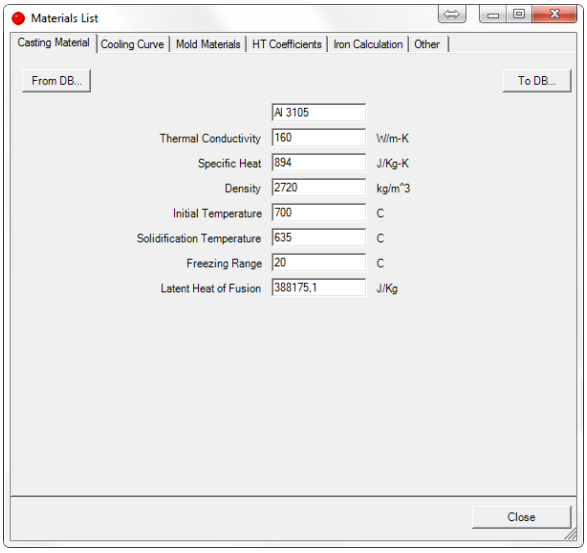


Figure 2. Parameters used in the modelling work.

For the modelling studies, two different melt temperature was selected: 700ºC and 750ºC; and three different mould temperatures were investigated: 20ºC, 120ºC and 250ºC. The parameters used in the simulations are given in Fig. 2.

SolidCast software was used to simulate the casting trials. The dimension of the mould used in the simulation is given in Fig. 3. The moulds were selected to be a sand mould and a permanent mould.

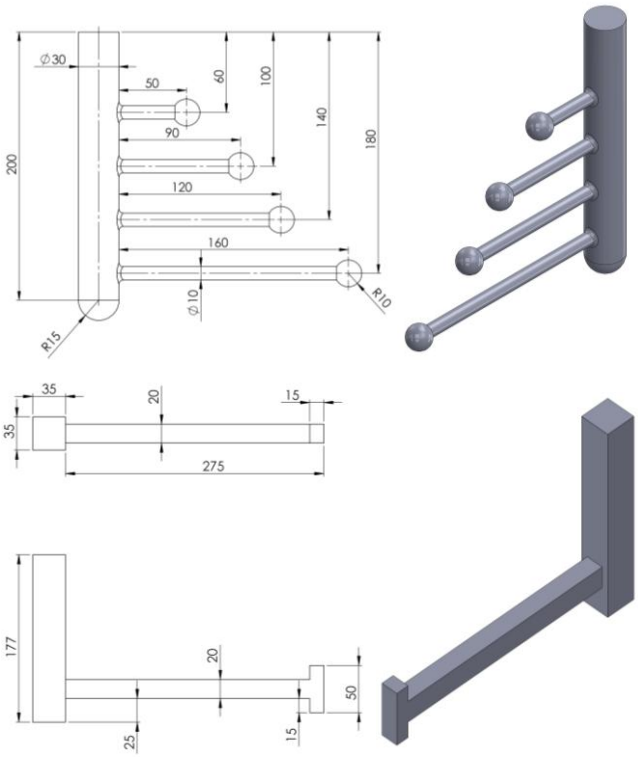


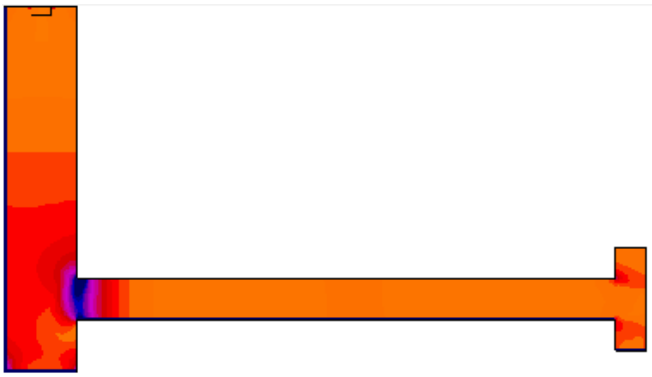
Figure 3. Dimension of the cast parts used in the modelling work.

3. Results

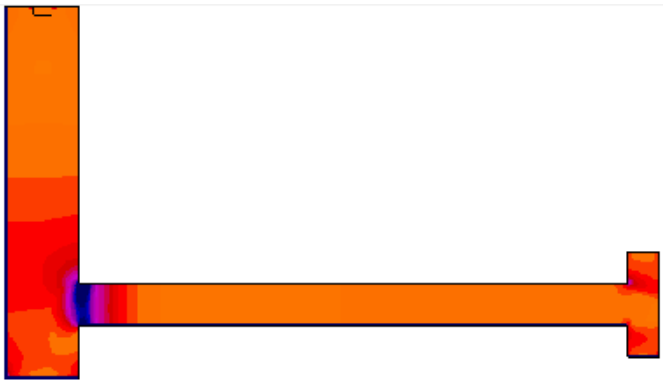
The simulation results that show the olidification time in the moulds are given in Table 2. The simulation results that show the hot spots in the moulds are given in Fig. 4-6.

Table 2. Solidification time results

T bone design	Solidification time(min.)	Hot tear susceptibility mould	Solidification time(min.)
Sand mould 700 °C	3.062	Sand mould 700 °C	1.895
Sand mould 750 °C	3.592	Sand mould 750 °C	2.242
Permanent mould (20ºC), cast at 700ºC	0.383	Permanent mould (20ºC), cast at 700ºC	0.287
Permanent mould (120ºC), cast at 700ºC	0.431	Permanent mould (120ºC), cast at 700ºC	0.324
Permanent mould (250ºC), cast at 700ºC	0.539	Permanent mould (250ºC), cast at 700ºC	0.405
Permanent mould (20ºC), cast at 750ºC	0.379	Permanent mould (20ºC), cast at 750ºC	0.292
Permanent mould (120ºC), cast at 750ºC	0.431	Permanent mould (120ºC), cast at 750ºC	0.337
Permanent mould (250ºC), cast at 750ºC	0.591	Permanent mould (250ºC), cast at 750ºC	0.443



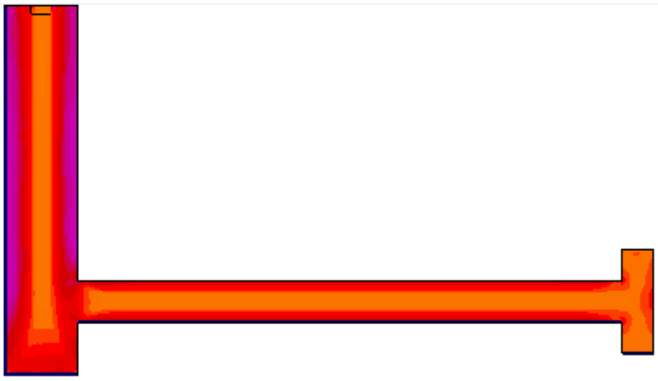
(a) Sand mould, cast at 700ºC



(b) Sand mould, cast at 750ºC

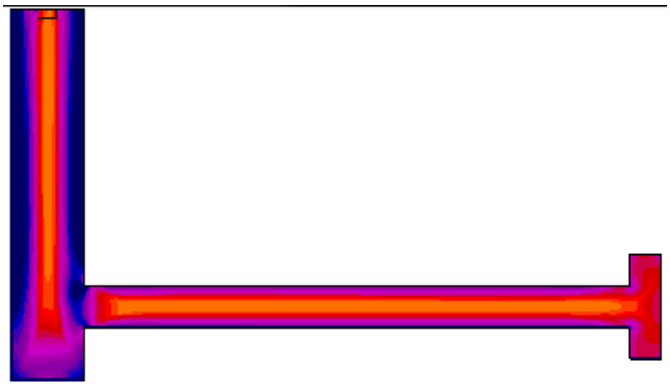
(c) Permanent mould (20ºC), cast at 700ºC

(d) Permanent mould (20ºC), cast at 750ºC



(e) Permanent mould (120ºC), cast at 700ºC

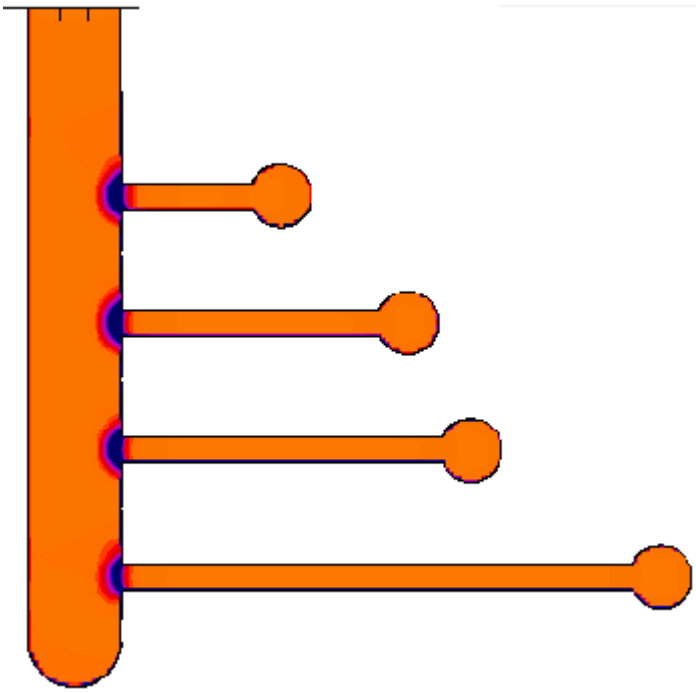
(f) Permanent mould (120ºC), cast at 750ºC



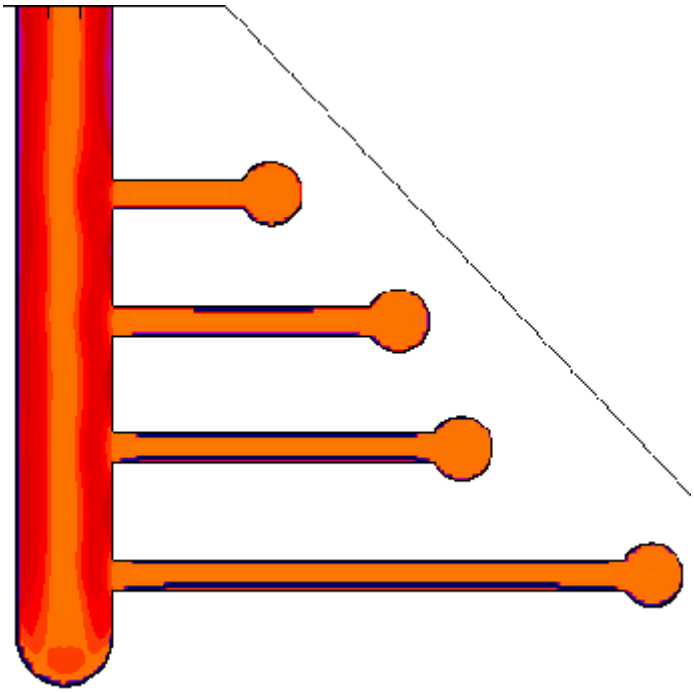
(g) Permanent mould (250ºC), cast at 700ºC

(h) Permanent mould (250ºC), cast at 750ºC

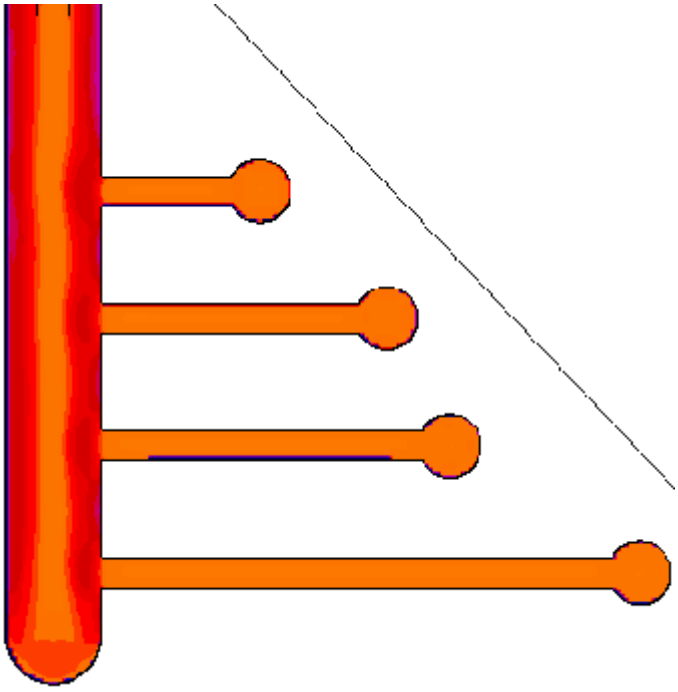
Figure 4: T-bone design simulation results



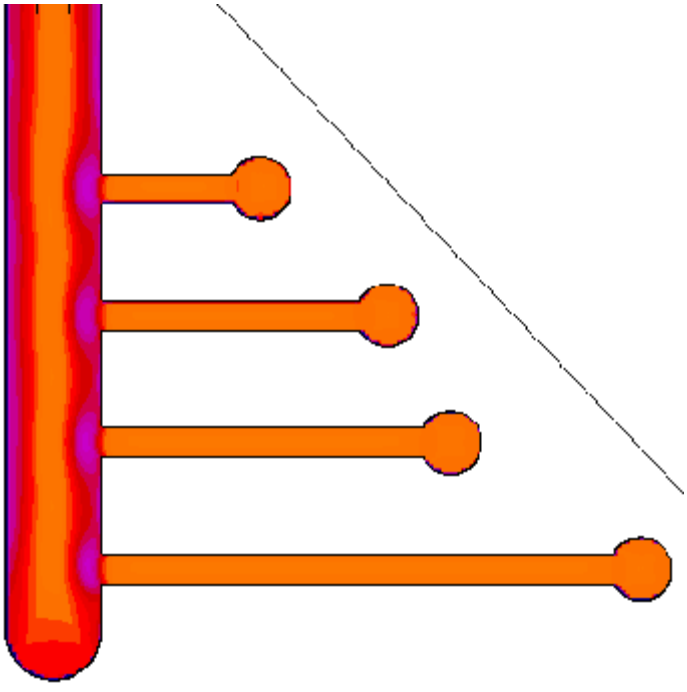
(a) Sand mould



(b) Permanent mould (20ºC)

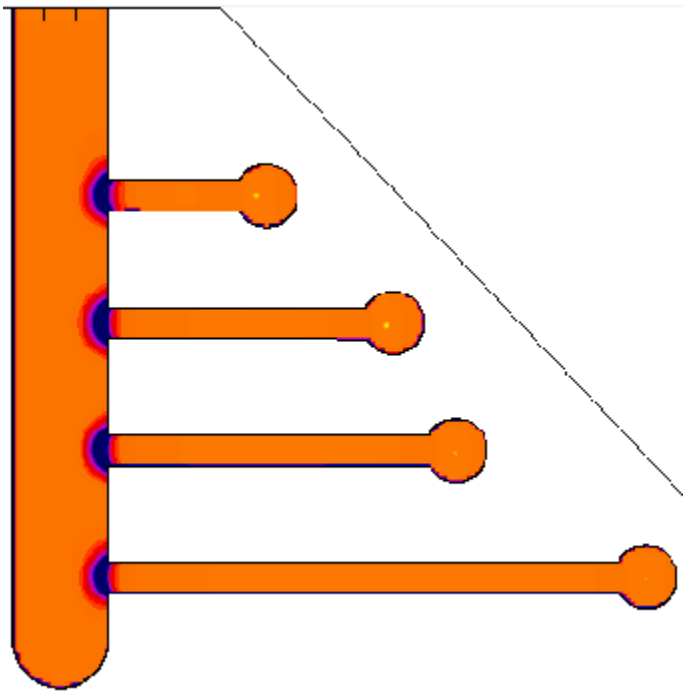


© Permanent mould (120ºC)

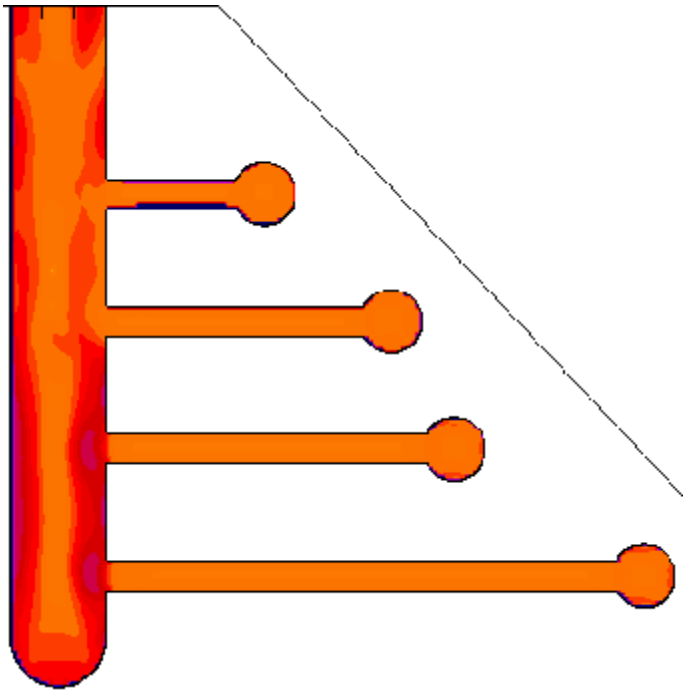


(d) Permanent mould (250ºC)

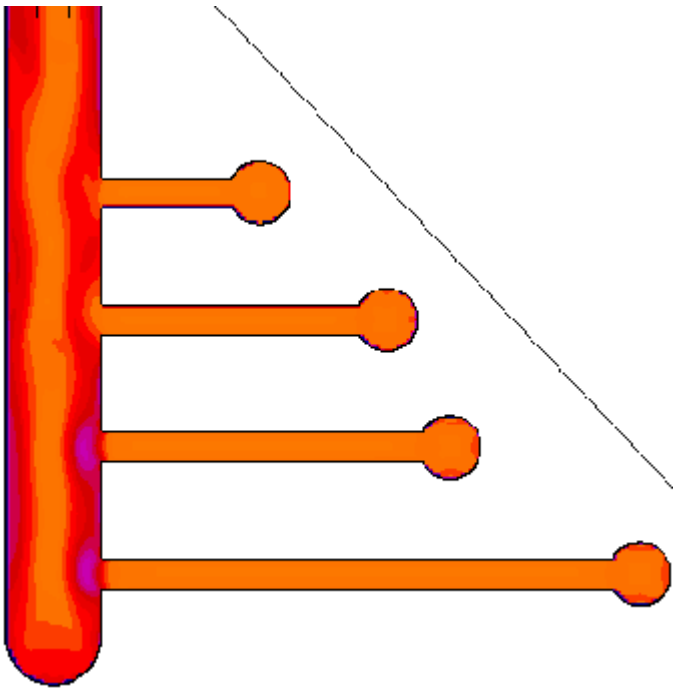
Figure 5: Hot tear susceptibility mould simulation results, cast at 700ºC



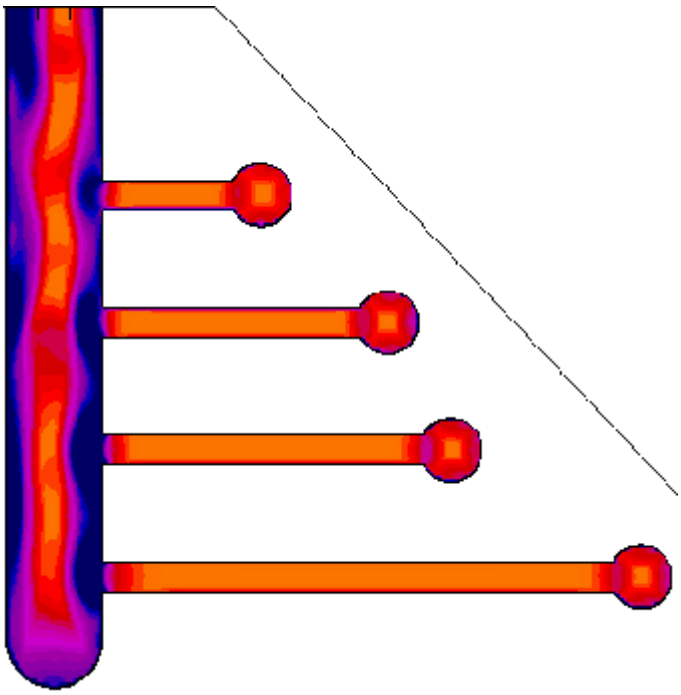
(a) Sand mould



(b) Permanent mould (20ºC)



© Permanent mould (120ºC)



(d) Permanent mould (250ºC)

Figure 6: Hot tear susceptibility mould simulation results, cast at 750ºC

4. Discussion

Two of the moulds were designed such that the long and thin cross-sectioned bars would solidify earlier than the thick pouring basin. In this way, there would be a hot spot at the junction points of the bars with the runner. These predictions were verified by the simulation results shown in Figs 4-6.

Different moulds (namely: sand and permanent) were used with different casting temperatures and various mould temperature in order to investigate the hot tearing susceptibility of 3105 alloy.

Thermal strain caused by the contraction during solidification was highest in the sand mould and at the permanent mould heated to 250 ^{o}C. It has been reported that 3105 type alloys exhibit a ductile fracture depending on the dendritic network (equiaxed or columnar).

The difference between the bar mould (Fig. 2a) and T-bone (Fig. 2b) is that with the bar mould, various effects of contraction rates can be examined. However, there are more studies with the T-junction type moulds in the literature.

The simulation results show the areas of hot spots. And the proportion and the distribution of these areas gives an indication

The existing hot tearing criteria have limited applicability. It has always been assumed that for a gap between grains to open into a tear, one of two conditions is required. Either a pressure drop within the mushy zone, or grains are pulled apart and the liquid film thickness increases. In either case, the pressure difference between the local liquid and the atmosphere exceeds the capillary pressure. However, one important aspect of hot tearing is that it needs to be nucleated. Aluminium alloys are known for their protective oxides that forms on the surface of the melt. When this oxide is disturbed during filling (i.e. uncontrolled pouring), a detrimental defect known as bifilms may form and these may be carried into the cast part. Bifilms are in the form of folded oxide skins. Therefore there exists an unbonded gap between these defect which may be readily opened to aid hot tearing.

Hot Tearing Susceptibility (HTC) in the mould given in Fig. 2a is calculated as follows:

$$HTS = \sum[L_i \times C_i] \quad (1)$$

Li is the length of the rod where the hot tear occurs and Ci is the severity of each tear. Severity ranges between 0-4 where 0 means no tear, 4 is the maximum tear where complete separation occur between the rod and the runner. An experimental work will be carried out to correlate melt quality with the HTS and simulation results.

4. Conclusions

Sand mould castings require chilling to eliminate hot tearing.

When permanent mould is used, the temperature of the mould play a significant role for hot tearing susceptibility. At room temperature, the feeding becomes difficult; at 250^{o}C, sharp edges increased hot tearing susceptibility; however, 120^{o}C appears to be an optimum temperature to eliminate hot tearing in 3105 alloy.

Simulation results are useful tool to identify the hot spots in a mould.

Further work is required to investigate the effect of melt quality and hot tearing susceptibility of aluminum castings.

References

[1] Campbell, J., *Castings*. 2003: Butterworth Heinemann.

[2] Eskin, D.G, L. Katgerman, 2007, *A Quest for a New Hot Tearing Criterion,* Mat Trans 38A, July, 1511

[3] Bichler, L. and C. Ravindran, *New Developments in Assessing Hot Tearing in Magnesium Alloy Castings.* Materials & Design, 2010. 31, Supplement 1(0): p. S17-S23.

[4] Dinnis, C.M., J.A. Taylor, and A.K. Dahle, *As-cast Morphology of Iron-intermetallics in Al–Si Foundry Alloys.* Scripta Materialia, 2005. 53(8): p. 955-958.

[5] Du, Z.-m., et al., *Tensile Properties of As-deformed 2A50 Aluminum Alloy in Semi-solid State.* Transactions of Nonferrous Metals Society of China, 2010. 20(9): p. 1597-1602.

[6] Esfahani, M.R.N. and B. Niroumand, *Study of Hot Tearing of A206 Aluminum Alloy Using Instrumented Constrained T-shaped Casting Method.* Materials Characterization, 2010. 61(3): p. 318-324.

[7] Fabrègue, D., et al., *Two- and Three-dimensional Characterizations of Hot Tears in a Al–Mg–Si Alloy Laser Weld.* Scripta Materialia, 2008. 59(3): p. 324-327.

[8] Farup, I., J.M. Drezet, and M. Rappaz, *In Situ Observation of Hot Tearing Formation in Succinonitrile-acetone.* Acta Materialia, 2001. 49(7): p. 1261-1269.

[9] Flemings, M.C., *Solidification Processing*. 1974: McGraw-Hill.

[10] Han, Q., et al., *Dross Formation During Remelting of Aluminum 5182 Remelt Secondary Ingot (RSI).* Materials Science and Engineering: A, 2003. 363(1–2): p. 9-14.

[11] Kamguo Kamga, H., et al., *Hot Tearing of Aluminum–copper B206 Alloys With Iron and Silicon Additions.* Materials Science and Engineering: A, 2010. 527(27–28): p. 7413-7423.

[12]Kwon, Y.-D. and Z.-H. Lee, *The Effect of Grain Refining and Oxide Inclusion on the Fluidity of Al–4.5Cu–0.6Mn and A356 Alloys.* Materials Science and Engineering: A, 2003. 360(1–2): p. 372-376.

[13]Li, M., et al., *The Effect of Y on the Hot-tearing Resistance of Al–5wt.% Cu Based Alloy.* Materials & Design, 2010. 31(5): p. 2483-2487.

[14]M'Hamdi, M. and A. Mo, *On Modelling the Interplay Between Microporosity Formation and Hot Tearing in Aluminium Direct-chill Casting.* Materials Science and Engineering: A, 2005. 413–414(0): p. 105-108.

[15]Phillion, A.B., S.L. Cockcroft, and P.D. Lee, *X-ray Micro-tomographic Observations of Hot Tear Damage in an Al–Mg Commercial Alloy.* Scripta Materialia, 2006. 55(5): p. 489-492.

[16]Phillion, A.B., S.L. Cockcroft, and P.D. Lee, *A New Methodology for Measurement of Semi-solid Constitutive Behavior and its Application to Examination of As-cast Porosity and Hot Tearing in Aluminum Alloys.* Materials Science and Engineering: A, 2008. 491(1–2): p. 237-247.

[17]Robinson, J.S., et al., *Influence of Processing on the Properties of the Aluminium Alloy 2025 with a Zirconium Addition.* Journal of Materials Processing Technology, 2009. 209(6): p. 3069-3078.

[18]Shabestari, S.G. and S. Ghodrat, *Assessment of Modification and Formation of Intermetallic Compounds in Aluminum Alloy Using Thermal Analysis.* Materials Science and Engineering: A, 2007. 467(1–2): p. 150-158.

[19]Suyitno, D.G. Eskin, and L. Katgerman, *Structure Observations Related to Hot Tearing of Al–Cu Billets Produced by Direct-chill Casting.* Materials Science and Engineering: A, 2006. 420(1–2): p. 1-7.

[20]Wang, J., et al., *The Apparent Viscosity of Fine Particle Reinforced Composite Melt.* Journal of Materials Processing Technology, 2003. 136(1–3): p. 60-63.

[21]Wang, Z., et al., *Hot Tearing Susceptibility of Binary Mg–Y Alloy Castings.* Materials & Design, 2013. 47(0): p. 90.