

IN THIS ISSUE

Ureas Impact on Curing
Dynamics

Alumina Nano and Micro
Particle Size

Characterization of Synthetic
Paraffin

Overview of Outlier Detection
Methods



Great Britain
Journals Press



IMAGE: OBSERVATORY WITH STAR
TRAILS ON MOUNTAINS FOR
CLEAR SKY

www.journalspress.com

LONDON JOURNAL OF ENGINEERING RESEARCH

Volume 24 | Issue 2 | Compilation 1.0

Print ISSN: 2631-8474
Online ISSN: 2631-8482
DOI: 10.17472/LJER





London Journal of Engineering Research

Volume 24 | Issue 2 | Compilation 1.0

PUBLISHER

Great Britain Journals Press
1210th, Waterside Dr, Opposite Arlington Building, Theale, Reading
Phone: +444 0118 965 4033 Pin: RG7-4TY United Kingdom

SUBSCRIPTION

Frequency: Quarterly

Print subscription
\$280USD for 1 year
\$500USD for 2 year

(color copies including taxes and international shipping with TSA approved)
Find more details at <https://journalspress.com/journals/subscription>

ENVIRONMENT

Great Britain Journals Press is intended about Protecting the environment. This journal is printed using led free environmental friendly ink and acid-free papers that are 100% recyclable.

Copyright ©2024 by Great Britain Journals Press

All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the publisher, except in the case of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law. For permission requests, write to the publisher, addressed “Attention: Permissions Coordinator,” at the address below. Great Britain Journals Press holds all the content copyright of this issue. Great Britain Journals Press does not hold any responsibility for any thought or content published in this journal; they belong to author's research solely. Visit <https://journalspress.com/journals/privacy-policy> to know more about our policies.

Great Britain Journals Press Headquaters

1210th, Waterside Dr,
Opposite Arlington
Building, Theale, Reading
Phone: +444 0118 965 4033
Pin: RG7-4TY
United Kingdom

Reselling this copy is prohibited.

Available for purchase at www.journalspress.com for \$50USD / £40GBP (tax and shipping included)

Featured Blog Posts

blog.journalspress.com

They were leaders in building the early foundation of modern programming and unveiled the structure of DNA. Their work inspired environmental movements and led to the discovery of new genes. They've gone to space and back, taught us about the natural world, dug up the earth and discovered the origins of our species. They broke the sound barrier and gender barriers along the way. The world of research wouldn't be the same without the pioneering efforts of famous research works made by these women. Be inspired by these explorers and early adopters - the women in research who helped to shape our society. We invite you to sit with their stories and enter new areas of understanding. This list is by no means a complete record of women to whom we are indebted for their research work, but here are of history's greatest research contributions made by...

Read complete here:
<https://goo.gl/1vQ3lS>



Women In Research



Computing in the cloud!

Cloud Computing is computing as a Service and not just as a Product. Under Cloud Computing...

Read complete here:
<https://goo.gl/VvHC72>



Writing great research...

Prepare yourself before you start. Before you start writing your paper or you start reading other...

Read complete here:
<https://goo.gl/np73jP>

Journal Content

In this Issue



Great Britain
Journals Press

- i.** Journal introduction and copyrights
- ii.** Featured blogs and online content
- iii.** Journal content
- iv.** Editorial Board Members

1. Electrodeposition of Copper Composites (Alumina Nano and Micro Particle Size) from Ionic Liquid (Ethaline) using Acoustic Impedance Electrochemical Quartz Crystal Microbalance (EQCM) at Cut Working Electrode in “Vertical” Position. **1-7**
2. Urea's Impact on Curing Dynamics, Resin Structure, and Properties of Urea-Formaldehyde Adhesives. **9-24**
3. Development and Characterization of Synthetic Paraffin Wax Scale for Oil Field Scaling Studies. **25-35**
4. An Overview of Outlier Detection Methods. **37-79**
5. Frequency Compensation Design of Three-Stage CMOS OTA Amplifier for RTD Applications. **81-87**

- V.** Great Britain Journals Press Membership

Editorial Board

Curated board members



Dr. Sharif H. Zein

School of Engineering,
Faculty of Science and Engineering,
University of Hull, UK Ph.D.,
Chemical Engineering Universiti Sains Malaysia,
Malaysia

Prof. Hamdaoui Oualid

University of Annaba, Algeria Ph.D.,
Environmental Engineering,
University of Annaba,
University of Savoie, France

Prof. Wen Qin

Department of Mechanical Engineering,
Research Associate, University of Saskatchewan,
Canada Ph.D., Materials Science,
Central South University, China

Dr. Luisa Molari

Professor of Structural Mechanics Architecture,
University of Bologna,
Department of Civil Engineering, Chemical,
Environmental and Materials, PhD in Structural
Mechanics, University of Bologna.

Prof. Chi-Min Shu

National Yunlin University of Science
and Technology, Chinese Taipei Ph.D.,
Department of Chemical Engineering University of
Missouri-Rolla (UMR) USA

Dr. Fawad Inam

Faculty of Engineering and Environment,
Director of Mechanical Engineering,
Northumbria University, Newcastle upon Tyne,
UK, Ph.D., Queen Mary, University of London,
London, UK

Dr. Zoran Gajic

Department of Electrical Engineering,
Rutgers University, New Jersey, USA
Ph.D. Degrees Control Systems,
Rutgers University, United States

Prof. Te-Hua Fang

Department of Mechanical Engineering,
National Kaohsiung University of Applied Sciences,
Chinese Taipei Ph.D., Department of Mechanical
Engineering, National Cheng Kung University,
Chinese Taipei

Dr. Rocío Maceiras

Associate Professor for Integrated Science, Defense University Center, Spain Ph.D., Chemical Engineering, University of Vigo, SPAIN

Dr. Rolando Salgado Estrada

Assistant Professor, Faculty of Engineering, Campus of Veracruz, Civil Engineering Department, Ph D., Degree, University of Minho, Portugal

Dr. Abbas Moustafa

Department of Civil Engineering, Associate Professor, Minia University, Egypt, Ph.D Earthquake Engineering and Structural Safety, Indian Institute of Science

Dr. Wael Salah

Faculty of Engineering, Multimedia University Jalan Multimedia, Cyberjaya, Selangor, Malaysia, Ph.D, Electrical and Electronic Engineering, Power Electronics and Devices, University Sians Malaysia

Prof. Baoping Cai

Associate Professor, China University of Petroleum, Ph.D Mechanical and Electronic Engineering, China

Dr. Kao-Shing Hwang

Electrical Engineering Dept., Nationalsun-Yat-sen University Ph.D., Electrical Engineering and Computer Science, Taiwan

Dr. Mu-Chun Su

Electronics Engineering, National Chiao Tung University, Taiwan, Ph.D. Degrees in Electrical Engineering, University of Maryland, College Park

Nagy I. Elkalashy

Electrical Engineering Department, Faculty of Engineering, Minoufiya University, Egypt

Dr. Vitoantonio Bevilacqua

Department of Electrical and Information Engineering Ph.D., Electrical Engineering Polytechnic of Bari, Italy

Prof. Qingjun Liu

Professor, Zhejiang University, Ph.D., Biomedical Engineering, Zhejiang University, China

Research papers and articles

Volume 24 | Issue 2 | Compilation 1.0



Scan to know paper details and
author's profile

Electrodeposition of Copper Composites (Alumina Nano and Micro Particle Size) from Ionic Liquid (Ethaline) using Acoustic Impedance Electrochemical Quartz Crystal Microbalance (EQCM) at Cut Working Electrode in "Vertical" Position

K. El Ttaib & A. Benhmid

ABSTRACT

Here we trace the electrolytic deposition of copper and copper composites from a solution of the metal chloride salt in ethylene glycol / choline chloride based eutectics. This work also explores the use of the electrochemical quartz crystal microbalance (EQCM) to monitor both the deposition and the inclusion of inert particulates into the copper coatings. This technique allows the first in-situ the determination the thickness of coated copper beside particulate inclusion. However the composition of composite material was strongly dependant on the amount of species suspended in solution and the size of incorporated alumina particles. It was also revealed that the majority of material was dragged onto the surface that has been shown in the difference on thickness between copper coated with 50 nm and 1 μ alumina.

Keywords: copper composites, ionic liquid, alumina, EQCM and SEM.

Classification: LCC Code: QD501-505

Language: English



**Great Britain
Journals Press**

LJP Copyright ID: 392931
Print ISSN: 2631-8474
Online ISSN: 2631-8482

London Journal of Engineering Research

Volume 24 | Issue 2 | Compilation 1.0



Electrodeposition of Copper Composites (Alumina Nano and Micro Particle Size) from Ionic Liquid (Ethaline) using Acoustic Impedance Electrochemical Quartz Crystal Microbalance (EQCM) at Cut Working Electrode in "Vertical" Position

K. El Ttaib^a & A. Benhmid^c

ABSTRACT

Here we trace the electrolytic deposition of copper and copper composites from a solution of the metal chloride salt in ethylene glycol / choline chloride based eutectics. This work also explores the use of the electrochemical quartz crystal microbalance (EQCM) to monitor both the deposition and the inclusion of inert particulates into the copper coatings. This technique allows the first in-situ the determination the thickness of coated copper beside particulate inclusion. However the composition of composite material was strongly dependant on the amount of species suspended in solution and the size of incorporated alumina particles. It was also revealed that the majority of material was dragged onto the surface that has been shown in the difference on thickness between copper coated with 50 nm and 1 μ alumina.

Keywords: copper composites, ionic liquid, alumina, EQCM and SEM.

I. INTRODUCTION

Electroplating of copper is crucial for a variety of commercial and embellishing purposes including mega scale use in the electronics industry for production of printed circuit boards, selective case hardening of steel for engineering components and productions od electrotypes in printing

industries.^{1,2} Copper may be easily deposited and electroplated with other metals and it is therefore particularly useful as a pre-coating for soft soldered work or for zinc alloy die-castings used by the automotive industry. In these cases the copper deposit provides a protective layer to the metal to allow further coatings to be applied³. Also the incorporation of particulate material into metal coatings has become an area of technological interest; for example incorporation of poly(ethylene) can act as a corrosion barrier, incorporation of PTFE or mica can reduce surface friction whereas incorporation of alumina, silicon carbide, boron nitride or diamond can indeed enhance and improve the hardness property of the metal coating.

Commercial copper electroplating is based on aqueous solutions which have high solubility for electrolytes and metal salts resulting in highly conducting solutions. They have high throwing power and components with complex shape and internal surfaces can be plated with effortlessness. However, the cyanide and acid based copper plating solutions that are most generally employed are highly corrosive and suffer from several drawbacks including, toxic effluent, high energy consumption and air pollution. Recently,

¹ P.C. Andricacos, C. Uzoh, J.O. Dukovic, J. Horkans and H. Deligianni, *Electrochemical Microfabrication*, 42, (1998), 5.

² P. C. Andricacos, *Interface*, 7, (1998), 23.

³ "The Canning handbook Surface Finishing Technology", W. Canning plc, in association with E. & F. N. Spon Ltd., London – New York, (1989).

ionic liquids have been used as alternative plating solutions, and they have shown to be a less environmentally hazardous option. Copper has been electrodeposited from many ionic liquids in between the 1960s and 1970s, the study of these systems was dominated by the deposition of copper using chloroaluminate ionic liquids, some of which still remain popular.^{4,5,6,7} With the introduction of discrete anions such as $[\text{BF}_4]^-$, and $[\text{F}_3\text{CSO}_2]_2\text{N}^-$ ($[\text{Tf}_2\text{N}]^-$) in the 1990s copper could be more easily electrodeposited and the air and moisture stability of these systems made them a great technological improve.^{8,9} Endres and co-workers have shown that copper may be deposited from $[\text{BMP}][\text{Tf}_2\text{N}]$ at various temperatures (BMP = butylmethyl pyrrolidinium). However, this particular ionic liquid has limited solubility for copper compounds and copper cations had to be introduced into the liquid *via* anodic dissolution of a copper electrode¹⁰. This work highlights that a key issue with the design of ionic plating systems is the coordination chemistry and concentration of the metal complex. While ionic liquids with discrete anions show significant potential for the electrodeposition of electronegative metals such as aluminium¹¹, issues such as toxicity, availability and cost may limit their practical use. We have recently shown that simple eutectic-based ionic liquids can be produced using quaternary ammonium salts $\text{R}_1\text{R}_2\text{R}_3\text{R}_4\text{N}^+\text{X}^-$ complexed with hydrogen bond donors such as acids, amides and alcohols. These

liquids, also known as deep eutectic solvents (DES) have been used for electropolishing^{12,13,14}, polymer synthesis¹⁵, electroless (immersion) deposition^{16,17} and metal oxide processing.^{18,19} Most of our previous studies have concentrated on choline chloride as the quaternary ammonium salt as it is non-toxic, biodegradable and is used already as a common constituent of various household and industrial products. Hence it can be applied economically to large-scale processes. DESs formed with choline chloride and either urea or ethylene glycol have successfully been employed for the electrodeposition of zinc, tin, and zinc–tin alloys²⁰. It was shown that the choice of hydrogen bond donor affects the type of alloy and the electrochemistry of the components in solution, as well as the morphology of the coatings. The area of metal deposition using ionic

⁴ R.T. Carlin, H.C. De Long, J. Fuller and P.C. Trulove, *J. Electrochem. Soc.*, **145**, (1998), 1598.

⁵ B.J. Tierney, W.R. Pitner, J.A. Mitchell, C. L. Hussey and G.R. Stafford, *J. Electrochem. Soc.*, (1998), **145**, 3110.

⁶ C. Nanjundiah and R.A. Osteryoung, *J. Electrochem. Soc.*, **130**, (1983), 1312.

⁷ F. Endres and A. Schweizer, *Phys. Chem. Chem. Phys.*, **2**, (2000), 5455.

⁸ P.-Y. Chen and I.-W. Sun, *Electrochim. Acta*, **44**, (1999), 441.

⁹ K. Murase, K. Nitta, T. Hirato and Y. Awakura, *J. Appl. Electrochem.*, **31**, (2001), 1089.

¹⁰ S. Zein El Abedin, A. Y. Saad, H. K. Farag, N. Borisenko, Q. X. Liu and F. Endres, *Electrochim. Acta*, **52**, (2007), 2746.

¹¹ F. Endres, *Chem. Phys. Chem.*, **3**, (2002), 144.

¹² A. P. Abbott, G. Capper, B. G. Swain, D. A. Wheeler, *Trans. IMF*, **83**, (2005), 51.

¹³ A. P. Abbott, G. Capper, K. J. McKenzie and K. S. Ryder, *Electrochimica Acta*, **51**, (2006), 4420.

¹⁴ A. P. Abbott, G. Capper, K. J. McKenzie, A. Glidle and K. S. Ryder, *Phys. Chem. Chem. Phys.*, **8**, (2006), 4214.

¹⁵ J. H. Liao, P. C. Wu, Y. H. Bai, *Inorg. Chem. Commun.*, **8**, (2005), 390.

¹⁶ A. P. Abbott, S. Nandhra, S. Postlethwaite, E. L. Smith and K. S. Ryder, *Phys. Chem. Chem. Phys.*, **9**, (2007), 3735.

¹⁷ A. P. Abbott, J. Griffith, S. Nandhra, C. O'Connor, S. Postlethwaite, K. S. Ryder and E. L. Smith, *Surf. Coat. Tech.*, **202**, (2008), 2033.

¹⁸ A. P. Abbott, G. Capper, D. L. Davies, R. Rasheed, P. Shikotra, *Inorg. Chem.*, **44**, (2005), 6497.

¹⁹ A. P. Abbott, G. Capper, D. L. Davies, K. J. McKenzie and S. U. Obi, *J. Chem. Eng. Data*, **51**, (2006), 1280.

²⁰ A. P. Abbott, G. Capper, K. J. McKenzie and K. S. Ryder, *J. Electroanal. Chem.*, **599**, (2007), 288.

liquids has been thoroughly reviewed in a series of recent articles^{21,22,23,24} and a book.²⁵

II. EXPERIMENTAL

Choline chloride [HOC2H4N(CH3)3Cl] (ChCl) (Aldrich 99%) was, when necessary, recrystallised from absolute ethanol, filtered and dried under vacuum. Ethylene glycol (EG) (Aldrich 99+%), was used as received. The mixtures were formed by stirring the two components together, in the stated proportions, at 50 °C until a homogeneous, colourless liquid formed. The liquids, once formulated, were kept in a thermostatic oven at 50 °C prior to use. Particulate suspensions were formed by mixing either Al2O3 (0.05µm or 1.0 µm, Aldrich, Aldrich) with the appropriate ionic liquid (1 ChCl : 2 EG, in wt./wt. % ratio).

Acoustic impedance electrochemical quartz crystal microbalance (EQCM) experiments were carried out using electrodes consisting of thin Au films (no Ti or Cr binding layer was used) evaporated on 10 MHz AT cut quartz crystals (International Crystal Manufacturing Co., Oklahoma City, USA); the surface finish of these crystals was unpolished. The piezoelectric active electrode area was 0.23 cm². Crystal impedance spectra were recorded using either Hewlett Packard HP8751A network analyser, connected to a HP87512A transmission / reflection unit via 50 Ω coaxial cable (or an Agilent ENA E5061A network analyser) such that the centre of the spectrum was near the centre of resonance, f_o (10 MHz), with a typical sweep width of 20-200 kHz depending on the interface. Peak

frequency/mass, Q factor data were extracted from the acoustic impedance spectra by fitting to a Lorentzian line shape using methods described elsewhere.^{16,20,26}

²¹ A. P. Abbott and K.J. McKenzie, *Phys. Chem. Chem. Phys.*, 8, (2006), 4265.

²² S. Z. El Abedin and F. Endres, *Acc. Chem. Res.*, 40, (2007), 1106.

²³ A. P. Abbott, K. S. Ryder and U. Koenig, *Trans. Inst. Metal Finishing*, 86, (2008), 196.

²⁴ N. V. Plechkova and K. R. Seddon, *Chem.. Soc. Rev.*, 37, (2008), 123.

²⁵“*Electrodeposition of Metals (Chapter 4)*”, T.J.S. Schubert, S.Z. El Abedin, A.P. Abbott, K.J. McKenzie, K.S. Ryder and F. Endres, in “*Electrodeposition in Ionic Liquids*”, Edited by F. Endres, A.P. Abbott and D.R. Macfarlane, Wiley-VCH Weinheim, (2008), ISBN: 978-3-527-31565-9.

²⁶G. McHale, C. Hardacre, R. Ge, N. Doy, R. W. K. Allen, J. M. MacInnes, M. R. Bown and M. I. Newton, *Anal. Chem.*, 80, (2008), 5806.

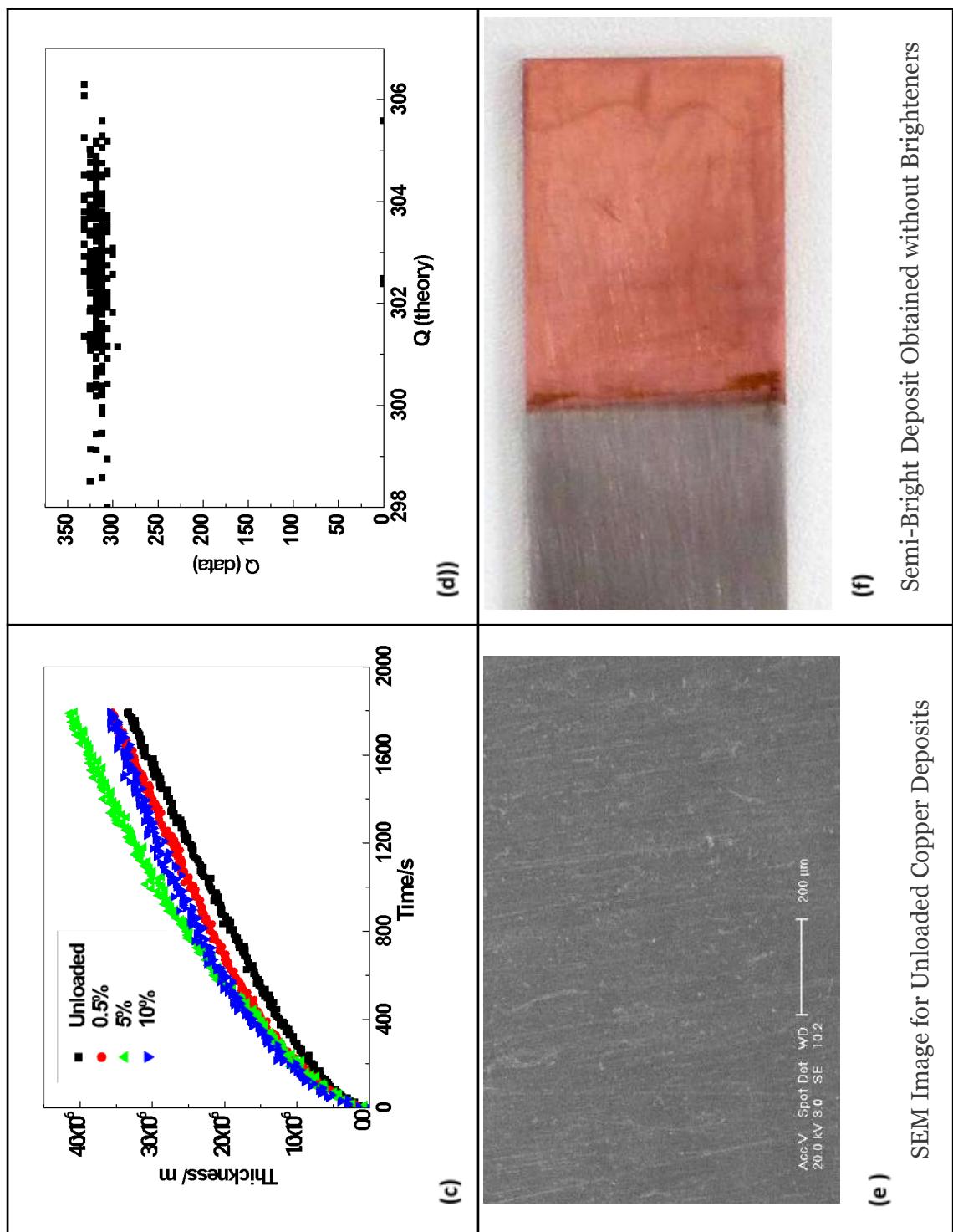


Figure 1: Mass-Charge Plots for Copper Composite Deposition at -0.8 v and 0.02 M [CuCl₂.2H₂O] in 1 ChCl: 2 EG Based Liquid for Different Solution-Phase Loadings of 0.05 μM Al₂O₃

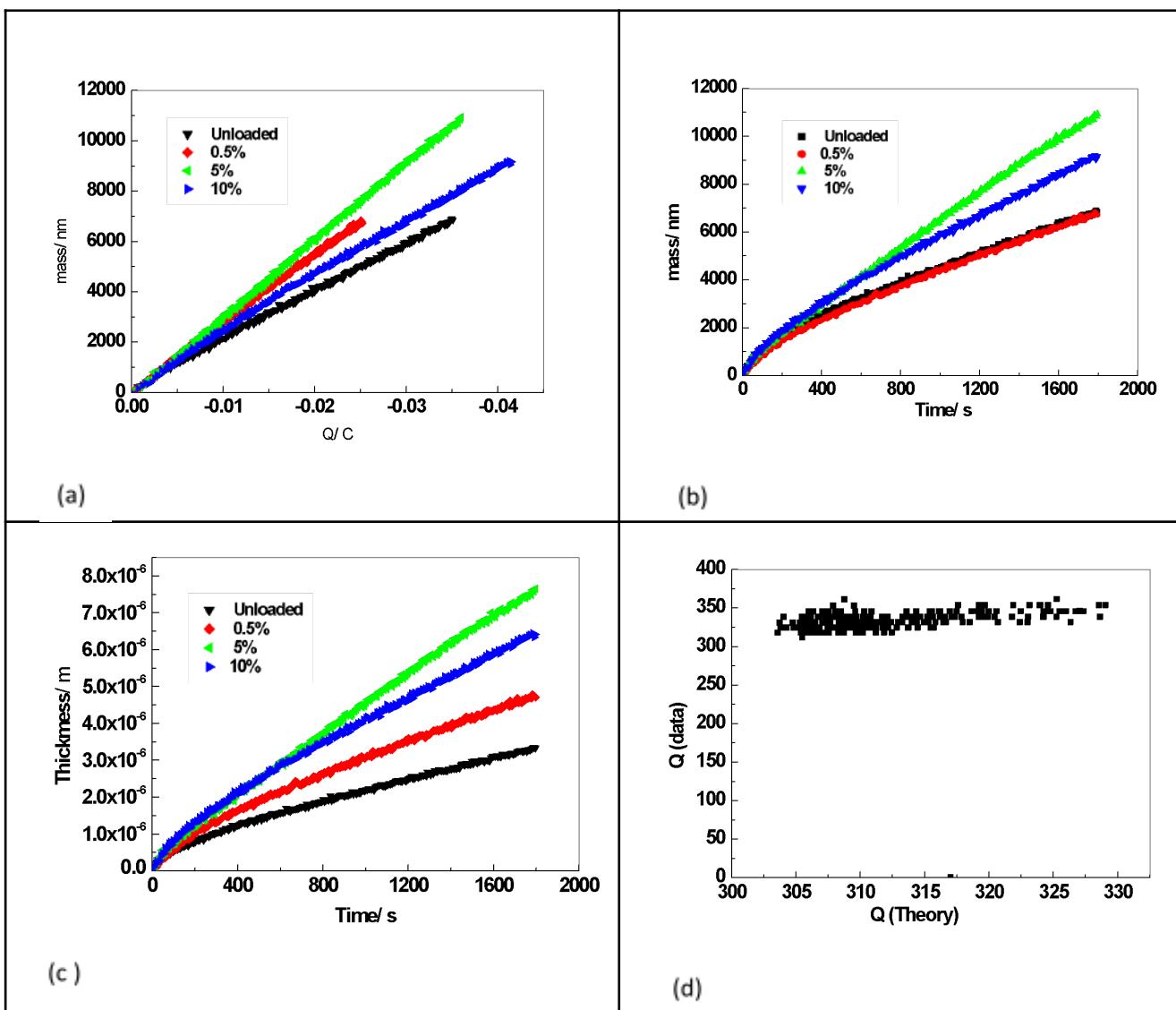


Figure 2: Mass-Charge Plots for Copper Composite Deposition at -0.8 V and $0.02\text{ M} [\text{CuCl}_2 \cdot 2\text{H}_2\text{O}]$ in $1\text{ M} \text{ChCl: 2 EG}$ Based Liquid for Different Solution-Phase Loadings of $1\text{ }\mu\text{M} \text{Al}_2\text{O}_3$

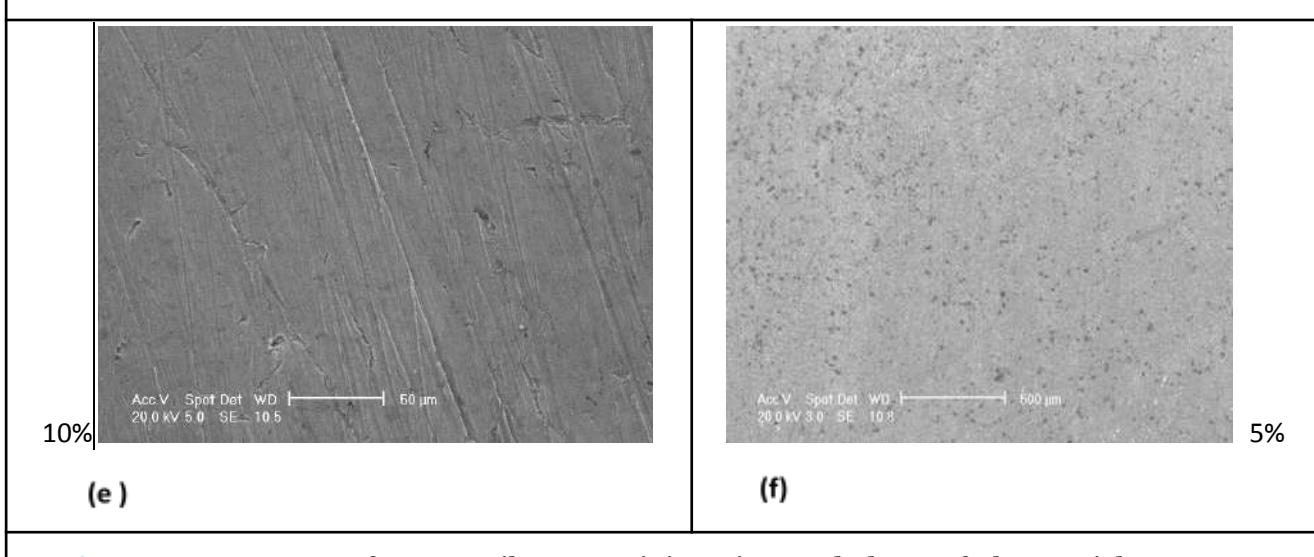


Figure 3: SEM Images of Copper Films Containing Dispersed Phase of Al_2O_3 : Right 5% $1\text{ }\mu\text{M} \text{Al}_2\text{O}_3$ - Left 10% $1\text{ }\mu\text{M} \text{Al}_2\text{O}_3$. Deposition at -0.8 V and $0.02\text{ M} [\text{CuCl}_2 \cdot 2\text{H}_2\text{O}]$ in $1\text{ M} \text{ChCl: 2 EG}$ Based liquid.(vertical Position)

III. RESULTS

Electrochemical quartz crystal microbalance experiments were used to obtain the mass / charge plots for copper deposition with two different sizes of alumina 50 nm and 1 μ m suspended in the ionic liquids. This technique utilises and permits real time *in-situ* in monitoring of both the copper deposition rate and the proportion of Al_2O_3 (or any particulate) incorporated into the composite. Since the particulate adds to the mass without consuming charge, where the place of AT cut working electrode was changed from horizontal to vertical position .This is important as it can address the issue of whether the composite material is settling on the surface or being dragged there by the copper complexes. The liquids were all kinetically stable on the time scale of the experiment i.e. the Al_2O_3 did not settle out at any time during the experiment, and this indicates that the sedimentation of the second phase is doubtful to be a crucial contributor to the overall rate of the material incorporation. *Figure 1a* depicts the mass-charge plot for copper at -0.8 V for different solution phase loadings of 0.05 μm Al_2O_3 in 1 ChCl: 2 EG based liquid (ethyline). *Figure 2a* shows the corresponding data for 1 μm Al_2O_3 . The comparison between the mentioned figures that they are not quite similar as in the deposition of the copper codeposits, where the gold substrate is in horizontal location.²⁷ By comparing between the alluded figures it has been observed that the mass of copper incorporated with the second phase sized 1 μm loaded is larger than the mass loaded with second phase 50 nm which is relatively interesting, moreover the deposition rates for similar experiments calculated from mass *versus* charge data is summarised in table1. The figures generally show that the codeposited copper with 1 μm particle alumina is larger than the codeposited copper with 0.05 μm particle alumina, it is a surprising results despite the 1 μm Al_2O_3 particles settle after about 24 hours whereas 0.05 μm particle suspensions are

²⁷A. P. Abbott, K. El Ttaib, G. Frisch, K. J. McKenzie and K. S. Ryder, *Phys. Chem. Chem. Phys.* 11(2009) 4269.

stable for over a week²⁸, also the rate of deposition by 5% micro particle is the highest, while the rate of deposition represented by 5% nano particles is lowest, this is needed to subject to further investigation. In the data shown in the table 1 the rate of copper deposition existence of 0.05 μm alumina do not have a remarkable change despite SEM-EDAX images perceive even distribution of nano particles in the deposits²⁹. The rate of copper deposition incorporated with 1 μm alumina is increased from 0% to 5% second phase. This can be attributed to that the mean mechanism of inclusion dispersed phase is dragging instead of settling on the substrate. The inclusion of 10% 1 μm alumina shows its rate of deposition than the expected value, this could due to high concentration of alumina particles in the ionic liquid. *Figure 1b* and *Figure 2 b* represent the mass loaded against time, for the inclusion of 0.5% nano particles found to be the most loaded material, this could be attributed to the small amount of 0.05 μm alumina is acting as micro stirrer²⁸. The thickness of both type of with codeposits nano and micro alumina viewed in *Figure 1c* and *Figure 2c*. The two Figures reveal that the thickness of copper codeposits is larger than the thickness of un loaded copper. The thickness of copper containing 1 μm Al_2O_3 particles almost the double the thickness of copper incorporated with 0.05 μm Al_2O_3 particles.

This is in a good agreement with the liquids were all kinetically stable on the time scale of the experiment i.e. the Al_2O_3 did not settle out at any time during the experiment and this suggests that the sedimentation of Al_2O_3 is unlikely to be a major contributor to the overall rate of material inclusion²⁹. The plot between the theoretical Q versus Q (data) is shown in *Figure 1d* and *Figure 2d*, where Q is the bandwidth and the resonant frequency determine the quality factor which can be seen in the mentioned Figures.

²⁸K. El Ttaib; The electrodeposition of composite materials using deep eutectic solvents (Thesis 2011).

²⁹Prof paper in press to Ttaib and Benhmid; *Int. Res. J. Pure Appl. Chem.*, vol. 25, no. 1, (2024) 22-27.

Table 1

| Cu deposits | Rate of deposits | Rate of deposits |
|-------------|-------------------------------------|-----------------------------------|
| | 50 nm alumina (g/cm ² C) | 1μm alumina (g/cm ² C) |
| Unloaded | 9.0 x10 ⁻⁹ | 9.0 x10 ⁻⁹ |
| 0.5% | 8.3 x10 ⁻⁹ | 12.2 x10 ⁻⁹ |
| 5% | 7.0 x10 ⁻⁹ | 13.8 x10 ⁻⁹ |
| 10% | 8.3 x10 ⁻⁹ | 9.8 x10 ⁻⁹ |

The SEM images give a substantial effect of incorporating aluminium particles; nano and micro particles on the morphology of the deposite, the size and the quantity of alumina, this is consistent with results by codeposition of hard second phase of silicon carbide (45-55nm / 1-3μ) and soft particles (PTFE) compared to deposition of pure copper, some mechanical properties such as hardness wear properties were studied^{29, 30}. One advantage of ionic liquids listed above is the ability to deposit metals onto surfaces that normally inactive or are water sensitive. *Figure 1 f* shows a photograph of nickel sample which was coated directly in a similar manner to the copper samples. The image reveal Semi-bright deposit obtained without brighteners.

IV. CONCLUSION

This work shows that ionic liquids, based on eutectic mixtures of choline chloride and hydrogen bond donors, such as ethylene glycol, can be used as electrochemical solvents for the electrodeposition of copper. The study revealed that in ionic liquid, the current efficiency for copper deposition is moderately efficient when AT cut working electrode is placed in a vertical position. Copper, incorporated with Al_2O_3 , has also been produced and shows that the loading of these species in the resulting electroplated films is not dependent on the concentration of particulate in a solution, when the working electrode is located horizontally. Gravitational settlement has been excluded as the main mechanism for particulate inclusion into the electrolytic deposits due to the position of the working electrode. It is surprising that the deposition copper incorporated nano particle alumina is fairly independent to the concentration of alumina, while the inclusion of micro alumina particles showed that 5% is the highest loading in the deposit, which is a very interesting observation.

The EQCM studies indicate that the copper codeposits thickness with 1 micro alumina particles is double the thickness with the 50 nm alumina particle, which supports that particle codeposition is gravitationally independent.

This page is intentionally left blank



Scan to know paper details and
author's profile

Urea's Impact on Curing Dynamics, Resin Structure, and Properties of Urea-Formaldehyde Adhesives

Daniil V. Ivanov, Aleksey S. Solovyov, Mark K. Loginov, Vadim A. Novakowski, Anton S. Mazur & Aleksey A. Kalashnikov

Saint Petersburg State Forest Technical University

ABSTRACT

The curing of urea-formaldehyde resin in the presence of urea used to reduce the formaldehyde content in particleboards was researched. Free urea was added into the adhesive with different amounts of latent curing agent (1 and 3% ammonium sulfate based on dry resin); a urea which is part of the modifiers-curing agents (products of the interaction of citric acid, urea and ammonia) was also studied separately. Based on the results of differential scanning calorimetry, solid-state NMR ^{13}C of the cured resin and tests of model particleboards, it was suggested that the effect of urea is significantly influenced by the pH value of the adhesive before and during curing. At a relatively high pH value (3.50-3.70 after curing at ammonium sulfate content of 1%), urea acts as a classical scavenger, that is, it chemically interacts with formaldehyde to form harmless products. At a reduced pH value (3.00-3.10 after curing at ammonium sulfate content of 3%), urea reacts with resin components (methylolureas and UF oligomers mainly) that is, it participates in the curing of the adhesive.

Keywords: particleboard, urea-formaldehyde resin, curing agents, modifier-curing agent, formaldehyde reduction, pH dependency, differential scanning calorimetry, solid-state NMR, toxicity, wood adhesives.

Classification: LCC Code: TP1150-1180

Language: English



Great Britain
Journals Press

LJP Copyright ID: 392932
Print ISSN: 2631-8474
Online ISSN: 2631-8482

London Journal of Engineering Research

Volume 24 | Issue 2 | Compilation 1.0

© 2024, Daniil V. Ivanov, Aleksey S. Solovyov, Mark K. Loginov, Vadim A. Novakowski, Anton S. Mazur & Aleksey A. Kalashnikov. This is a research/review paper, distributed under the terms of the Creative Commons Attribution- Noncom-mercial 4.0 Unported License <http://creativecommons.org/licenses/by-nc/4.0/>, permitting all noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.



Urea's Impact on Curing Dynamics, Resin Structure, and Properties of Urea-Formaldehyde Adhesives

Daniil V. Ivanov^a, Aleksey S. Solovyov^a, Mark K. Loginov^p, Vadim A. Novakowski^{CO},
Anton S. Mazur^Y & Aleksey A. Kalashnikov[§]

ABSTRACT

The curing of urea-formaldehyde resin in the presence of urea used to reduce the formaldehyde content in particleboards was researched. Free urea was added into the adhesive with different amounts of latent curing agent (1 and 3% ammonium sulfate based on dry resin); a urea which is part of the modifiers-curing agents (products of the interaction of citric acid, urea and ammonia) was also studied separately. Based on the results of differential scanning calorimetry, solid-state NMR ¹³C of the cured resin and tests of model particleboards, it was suggested that the effect of urea is significantly influenced by the pH value of the adhesive before and during curing. At a relatively high pH value (3.50-3.70 after curing at ammonium sulfate content of 1%), urea acts as a classical scavenger, that is, it chemically interacts with formaldehyde to form harmless products. At a reduced pH value (3.00-3.10 after curing at ammonium sulfate content of 3%), urea reacts with resin components (methylolureas and UF oligomers mainly) that is, it participates in the curing of the adhesive. Tests of model samples of particleboards have shown that the mechanism of including urea into the structures of the curing resin is less effective for reducing toxicity than its action as a classical scavenger; at the same time, the formaldehyde content in particleboards with 3% ammonium sulfate and urea is lower in absolute values due to the useful effect of the curing agent. Modifiers-curing agents provide a low pH value of the binder before curing (3.72-4.51 with a content of modifiers-curing agents of 5%), urea in their composition is also involved in curing (according to solid-state NMR data).

Keywords: particleboard, urea-formaldehyde resin, curing agents, modifier-curing agent, formaldehyde reduction, pH dependency, differential scanning calorimetry, solid-state NMR, toxicity, wood adhesives.

Author a: Department of Wood and Cellulose Composite Technology, Institute of Chemical Wood Biomass Processing and Technospheric Safety, Saint Petersburg State Forest Technical University, Institutskiy per. 5, 194021 St. Petersburg, Russia.

a: Department of Wood and Cellulose Composite Technology, Institute of Chemical Wood Biomass Processing and Technospheric Safety, Saint Petersburg State Forest Technical University, Institutskiy per. 5, 194021 St. Petersburg, Russia.

p: Department of Wood and Cellulose Composite Technology, Institute of Chemical Wood Biomass Processing and Technospheric Safety, Saint Petersburg State Forest Technical University, Institutskiy per. 5, 194021 St. Petersburg, Russia.

CO: Centre for Thermal Analysis and Calorimetry of Saint Petersburg State University Research Park, Saint Petersburg State University, Universitetskaya nab. 7/9, 199034 St. Petersburg, Russia.

Y: Magnetic Resonance Research Centre, Research Park, Saint-Petersburg State University, Universitetskaya nab. 7/9, 199034 St. Petersburg, Russia.

§: JSC «Slotex», Industrialny prospect, 64, 195279 St. Petersburg, Russia.

I. INTRODUCTION

Nowadays one of the most popular adhesive for wood boards remains of urea-formaldehyde resins (UF resin) [1, 2]. Its main disadvantage is high toxicity which determined by formaldehyde emission during resin curing. A common way to reduce toxicity is to use formaldehyde scavengers (acceptors) – modifiers that interact with formaldehyde to form harmless products. To date,

many different scavengers are known [3, 4], but due to the low cost and high availability, urea remains one of the most common modifiers.

Depending on the type of product (particleboard, medium density fiberboard, plywood, etc.), there are many ways to adding urea to the materials composition. The scavenger can be added to the composition or separately from adhesives, in solution or powder form, in one layer (only in core or only in surface layers) or in all volume of wood board, at that its efficiency can be different. It is easier to add urea to the adhesive composition, since it requires a minimum of special equipment and is suitable for both wood boards and plywood. In that case, the action of urea is to interact with resin components, the direction and depth of which define the properties of final product and effectiveness of reducing the release of formaldehyde emission.

It is known that urea reduces the content of free formaldehyde as a result of chemical interaction and the formation of urea methylol derivatives; in particular this principle is used in the synthesis of resin at the third stage in alkaline conditions aiming to capture previously unreached formaldehyde. In addition, urea can interact not only with free formaldehyde, but also with other resin components such as methylolureas and UF oligomers. Hui Wang at al. [5] showed that post-added urea (decrease of F/U molar ratio from 2.0 to 1.2 at pH 8.0–8.5, and further shutter speed at 60 °C for 20 min) may act to urea-formaldehyde oligomers (UF oligomers), that is, it not only neutralizes formaldehyde, but also changes the structure of the main resin components. According to the authors, urea caused the movement of the type II methylol groups from oligomers to modifier with methylolurea forming due to primary amino groups of urea is more reactive than the secondary amino groups of UF oligomers. Vyukov [6] explored the curing of UF resin with urea and 0.3 % of oxalic acid at 20 °C for 10–60; min the author claims that the use of scavenger at reduce of F/U molar ratio from 2.0 to 1.6 and 1.2 leads to distractions of methylene esters bridges of UF oligomers. In first 20 min observed increased of methylolureas amount, but next 40 minutes

contents decreased to zero, the methylol group's content stably decreased and free formaldehyde changed slightly. It is assumed that free urea causes of partial destruction of UF oligomers and further curing occurs due to the interaction between methylolureas. In some works it is reported that the curing of UF resin under traditional conditions (after latent curing agent adding and at the heating) it begins with the interaction of urea (unreached during synthesis) with the methylol groups of resin components [7–9].

Thus, it can be assumed that free urea participates to curing on a par with methylolureas and UF oligomers, but its effect on the cured resin is usually negative. The presence of urea leads to reduced of resin reactivity; this is typical both for low molar ratio resins [7, 10] and for resins in which urea is added as a scavenger [11]. As a rule, this leads to a decrease of wood boards properties, primarily water resistance [11–13]. An effective way to preserve the high resin reactivity and satisfactory wood board's properties while maintaining the ability of scavengers to neutralize formaldehyde is the use of urea derivatives. Park et al. [11] showed that the mono-methylolurea is more effective than urea in scavenging of the formaldehyde and in ensuring the adhesive ability of the UF resin. Yang at al. [14] disclosed the possibility of polyurea using for increase of particleboard properties and decreased of formaldehyde emission. Perminova at al. showed the efficiency of glycoluril – product of interaction urea with glyoxal, which reducing the formaldehyde emission of particleboard to 34 % without degradation of strength and water resistant [15].

As derivatives of urea, products of its interaction with organic acids, such as citric acid, can be used. The resulting urea citrate [16] introduced into the resin performs the function of a direct curing catalyst; the addition of ammonia to the formulation makes it possible to give the resulting product the properties of a latent curing agent. The product was named Modifier-Curing agent (MC), because it's able to replace the traditional curing agents (ammonium chloride, sulfate, nitrate etc.) in adhesive compositions and reduce

formaldehyde emissions from wood boards without negative effect on physical and mechanical properties. Action MC as a curing agent showed in article [17], action MC as a modifier required to research.

The aim of this work was to study of free urea and MC urea on the curing of UF resin, structure of UF polymer and properties of particleboard.

Table 1: Properties of Commercial UF Resin of the KF-MT-15 Brand

| The appellation of the parameter | The value of the parameter |
|---|----------------------------|
| Mass fraction of solid resin, % | 65.1 |
| Conditional viscosity according to the viscometer with a nozzle diameter of 4 mm, s | 69 |
| Concentration of hydrogen ions, pH | 8.3 |
| Curing time with the addition of 1% ammonium chloride at 100 °C, s | 43 |
| Free formaldehyde content, % | 0.15 |

Aqueous solutions of ammonium sulfate (Russian State Standards – GOST 9097-82 “Ammonium sulphate. Specifications”) with a concentration of 20% were used as curing agents. The urea's brand "A", highest grade (GOST 2081-2010 “Carbamide. Specifications”), were used as a formaldehyde scavenger.

II. MATERIALS AND METHODS

2.1 Materials

The UF resin of the KF-MT-15 brand (F/U molar ratio of 1.2/1) was produced by one of the Russian manufacturers. The resin characteristics are given in Table 1.

We also used modifier-curing agents first series which is a products of interaction of citric acid, urea and ammonia with different molar ratios; MC-1(0.5) – 1/2.5/0.5 respectively, MC-1(1.5) – 1/1.5/1.5 respectively. The properties of the modifier-curing agent are presented in Table 2.

Table 2: Properties of Modifier-Curing Agents

| The appellation of the parameter | The value of the parameter | |
|-------------------------------------|--|-----------|
| | MC-1(1.5) | MC-1(0.5) |
| Appearance | transparent liquid without mechanical impurities | |
| Mass fraction of the dry residue, % | 40 | 40 |
| Concentration of hydrogen ions, pH | 4.2 | 2.3 |
| Nitrogen content, % | 8.2 | 8.8 |
| The content of amino groups, % | 6.3 | 9.1 |

2.2 Methods

2.2.1 Curing Time of UF Resin (At 100 ± 1 °C)

Compositions (depending of the experiment's objectives, for example resin + water + urea + curing agent) were prepared in such a way that the mass fraction of solid resin in all cases was 55%. 12 g of commercial resin was weighed with an accuracy of 0.02 g, the calculated amount of distilled water and curing agent solutions were added. The tubes with the compositions were heated in a water bath, continuously stirring. The time from the immersion of the test tube in

boiling water to the compositions loss of fluidity is considered to be the duration of curing. The mass of adhesive components (curing agent, modifier-curing agent, urea) in all cases, based on solid resin was calculated.

2.2.2 pH-Value of the Cured Adhesive

Compositions (depending of the experiment's objectives, for example resin + water + urea + curing agent) were prepared in such a way that the mass fraction of solid resin in all cases was 57%. 20 g of commercial resin was weighed with an accuracy of 0.02 g, the calculated amount of

distilled water and curing agent solutions were added.

The adhesive were cured at 110 °C for 2 min 26 s (thus, the press-factor of 0.15 min/mm is reproduced for particleboard with a thickness of 16 mm). The cured compositions were conditioned at room temperature for 30 min. Then the compositions were crushed and sifted using a sieve with a diameter of holes 0.5 mm. The powders which passed through the sieve and remained on the pallet were taken.

The powder were placed in a glass with a capacity of 100 cm³ and filled with 50 cm³ of water room temperature. Extraction was carried out for 30 min, after which the pH value were determined.

2.2.3 DSC Measurement

The adhesive samples were prepared in the same way as in clause 2.2.2. DSC measurements were carried out using a Netzsch DSC 204F1 Phoenix. The sample resins were tested by placing about 12.7 mg of each sample into a hermetic pan. Heating rates of 10 °C/min and a temperature range of 30 to 200 °C.

2.2.4 Solid State ¹³C NMR Spectroscopy

The adhesive samples were prepared, cured and crushed in the same way as in clause 2.2.2. ¹³C solid-state NMR spectrometers were used for instrumental analysis of the structure of the cured compositions. NMR spectra were obtained using a BRUKER AVANCE III WB 400 spectrometer; the zirconium oxide 4 mm rotor of the device rotated at a frequency of 12.5 kHz; CP/MAS pulse sequence ¹³C {¹H} was used; relaxation delay – 2 s; contact time – 2 ms; number of pulses – 2048.

2.2.5 Preparation of Model Particleboards and Its Performance

Fourteen one-layer particleboards with dimensions of 200×200×4.0 mm and a target density of 700 kg/m³ were prepared in the laboratory under a specific pressure of 2.8 MPa, at 110 °C press temperature; the total press time was 2 min 26 s. The solid resin loading was 20% based

on dry wood particles; the adhesives composition depended on the objectives of the experiment.

Were tested such properties of particleboard as ultimate bending strength (BS) to GOST 10635–88 “Particle boards. Methods for determining ultimate strength and modulus of elasticity in bending”; water absorption (WA) and thickness swelling (TS) of 24 h to GOST 10634–88 “Wood particle boards. Methods for determination of physical properties”; formaldehyde emission (FE) to WKI method [18].

The aim of manufacturing and testing of the model particleboards was researched the behavior of adhesives under hot pressing temperature of the core layer of particleboards and plywood. For this reason, the thickness of the boards and the pressing temperature has been reduced, and the adhesive content has been increased.

2.2.6 Preparation of Three-Layer Particleboards and Its Performance

Four three-layer particleboards with dimensions of 200×200×16 mm and a target density of 650 kg/m³ were prepared in the laboratory under a specific pressure of 2.8 MPa, at 220 °C press temperature; the press-factor was 0.15 min/mm of thickness (total press time – 2 min 26 s). The mass fraction of surface layers was 40 %, core layer 60 %. The solid resin loading was 12% based on dry wood particles of surface layers and 10% on particles of core layer; the mass fraction of solid resin in surface layers adhesive was 55 %; in core layer 57 %. The core and surface layers compositions shown to table 3.

Properties of three-layer particleboards which were tested shown in clause 2.2.5; additionally, we determined internal bonding strength (IB) according to GOST 10636–2018 “Wood-shaving and wood-fiber plates. Strength definition method at stretching perpendicularly plate layer”.

Table 3: Adhesives Compositions for Core and Surface Layers of Particleboards

| Component | Component content, %, regarding to solid resin, for particleboard type | | | | | | | |
|-----------------------|--|----|------|------|----|----|----|----|
| | 1* | | 2 | | 3 | | 4 | |
| | SL | CL | SL | CL | SL | CL | SL | CL |
| Ammonium sulfate | 1 | 3 | 1 | 3 | — | — | — | — |
| Modifier-curing agent | | | | | | | | |
| • MC-1(0.5) | — | — | — | — | — | — | — | 5 |
| • MC-1(1.5) | — | — | — | — | 3 | 5 | 3 | — |
| Urea | — | — | 0.43 | 2.15 | — | — | — | — |

* Control samples were made according to composition 1

III. EXPERIMENTAL

Fig. 1 shows the typically action of ammonium sulfate as a curing agent. Increasing the amount of this catalyst gain of more than 3 % does not lead to decreased of curing time. The addition of 1-5 % urea to the adhesive doesn't leads to a significant change both with an ammonium sulfate content of 1% and with an ammonium sulfate content of 3%

(Fig. 2). MC-1(0.5) cures the resin longer than the ammonium sulfate at an amount less than 3 %, but with an increase in the content of this modifier-curing agent curing time decreased by 28-38 %. MC-1(1.5) is inferior to ammonium sulfate even at a high content (with 3-5 % of MC-1(1.5) the adhesive cures longer to 10-22 % than with 3 % of ammonium sulfate).

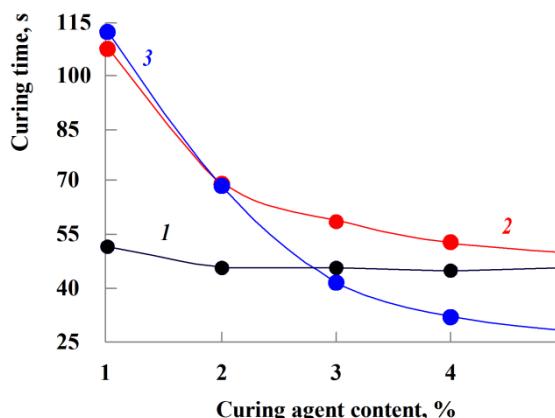


Fig. 1: Curing time of UF adhesive with various curing agents: 1 – adhesive with ammonium sulfate; 2 – adhesive with MC-1(1.5); 3 – adhesive with MC-1(0.5)

Since the increase in the content of ammonium sulfate is over than 3 % doesn't provide a significant effect, in further work the effect of this curing agent was researched only at two levels – 1 and 3 %. The action of modifier curing agents was researched at three levels – 1, 3 and 5 %, as they are influence to adhesive in a wider range of amount. The urea was added only to resin without modifier-curing agents (Table 4); its content was

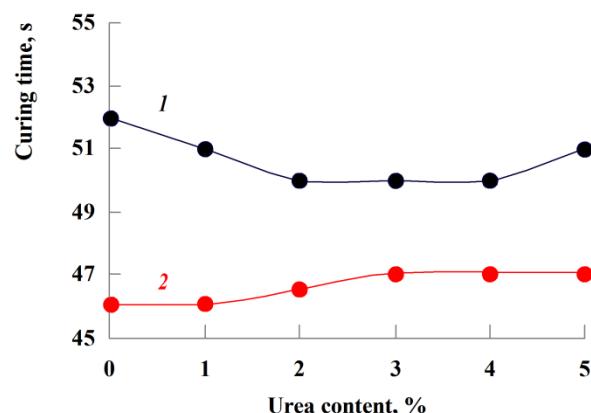


Fig. 2: Curing time of UF adhesive with ammonium sulfate and various urea content: 1 – adhesive with 1 % ammonium sulfate; 2 – adhesive with 3 % ammonium sulfate

varied in such a way as to correspond to the amount of scavenger introduced with MC-1(0.5) in parallel experiment.

The pH value of the initial and cured adhesive is shown in Table 4. Ammonium sulfate has a slight effect on the resin's pH immediately after addition, but the acidity of the cured adhesive decreases sharply under its action. Similar results

were obtained by Kantieva and Ponomarenko [19] with ammonium chloride. As expected, a larger amount of the curing agent provides a lower pH value during curing, therefore urea in combination with 3 % ammonium sulfate acts in a more acidic environment. Modifier-curing agents, on the contrary, sharply reduce the pH immediately after combining with the resin, so the difference in acidity of the adhesive before and after curing is not so significant. MC-1(0.5) provides the lowest pH value of the cured resin compared to other options; however, the pH value of the cured resin is comparable to the pH of the resin cured of 1 % ammonium sulfate.

Thermal curing behaviors of researched adhesives are shown in Fig. 3 and 4. Almost all of the observed peaks are endothermic, since the meeting was carried out without pressure. Such effects as water evaporation, emission of formaldehyde and reactions water overlap the exothermic effects of resin curing. In area of temperature peaks 67-92 °C endothermic effects may be caused by water evaporation [20, 21], in area 111-112 °C by emission of polycondensation's by-products [22]. Narrow and high peak in high temperatures area may be related with transformation of methylene ether bridges into methylene with formaldehyde's emission.

Table 4: pH Value of Curved UF Resin With Different Adhesive Compositions

| Adhesive composition | | | pH value | |
|----------------------|----------------------|---------------|---------------|--------------|
| Curing agent type | Curing agent content | Urea content* | Before curing | After curing |
| Ammonium sulfate | 1 | 0.00 | 6.72 | 3.51 |
| | | 0.43 | 6.83 | 3.61 |
| | | 1.29 | 6.87 | 3.72 |
| | | 2.15 | 6.91 | 3.70 |
| | 3 | 0.00 | 6.61 | 3.05 |
| | | 0.43 | 6.63 | 3.05 |
| | | 1.29 | 6.62 | 3.02 |
| | | 2.15 | 6.59 | 3.00 |
| MC-1(0.5) | 1 | 0.43 | 5.18 | 4.06 |
| | 3 | 1.29 | 4.04 | 3.64 |
| | 5 | 2.15 | 3.72 | 3.62 |
| MC-1(1.5) | 1 | 0.29 | 5.56 | 4.45 |
| | 3 | 0.88 | 4.69 | 3.79 |
| | 5 | 1.46 | 4.51 | 3.70 |

* Urea wasn't added when the modifier-curing agents using; in table shown urea which is part of the MC-1(0.5) and MC-1(1.5)

The adding of urea in adhesive has a significant effect on the DSC curve (not change only curve's form of sample with 3 % ammonium sulfate and 0.43 % urea). For samples with ammonium sulfate (1 %) and urea (Fig. 3, curves 2-4) top of the peaks in area 67-92 and 111-112 °C merge into one big peak in that is, the evaporation of solvent water and the release of reaction water occur simultaneously. With an increase of the scavenger content in the adhesive, the enthalpy of the first endothermic area decreases (Table 4), most likely due to increased exothermic effects [22]. DSC curve of the sample with 3 % ammonium sulfate and 0.43 % urea (Fig. 3, curve 6) not changes

compared to the curve of the sample without scavenger. At ammonium sulfate content 3 % and urea amount 1.29 % and 2.15 % in area 84-95 °C appears the exothermic peak (top of the peak 88 °C); at ammonium sulfate content 1 % exothermic peaks not observed. An increase in heat generation may be related to reaction urea with methylol groups of methylolureas and UF oligomers [7-9] or reaction between urea and free formaldehyde [22]. The higher curing agent content contribute to intensification of the exothermic reactions, wherein with the appearance of a peak in the within the temperature range of 84-95 °C suggests that urea

interact mainly with methylol groups, since it's in this area the effects from curing reactions usually observed [6-9, 24, 25].

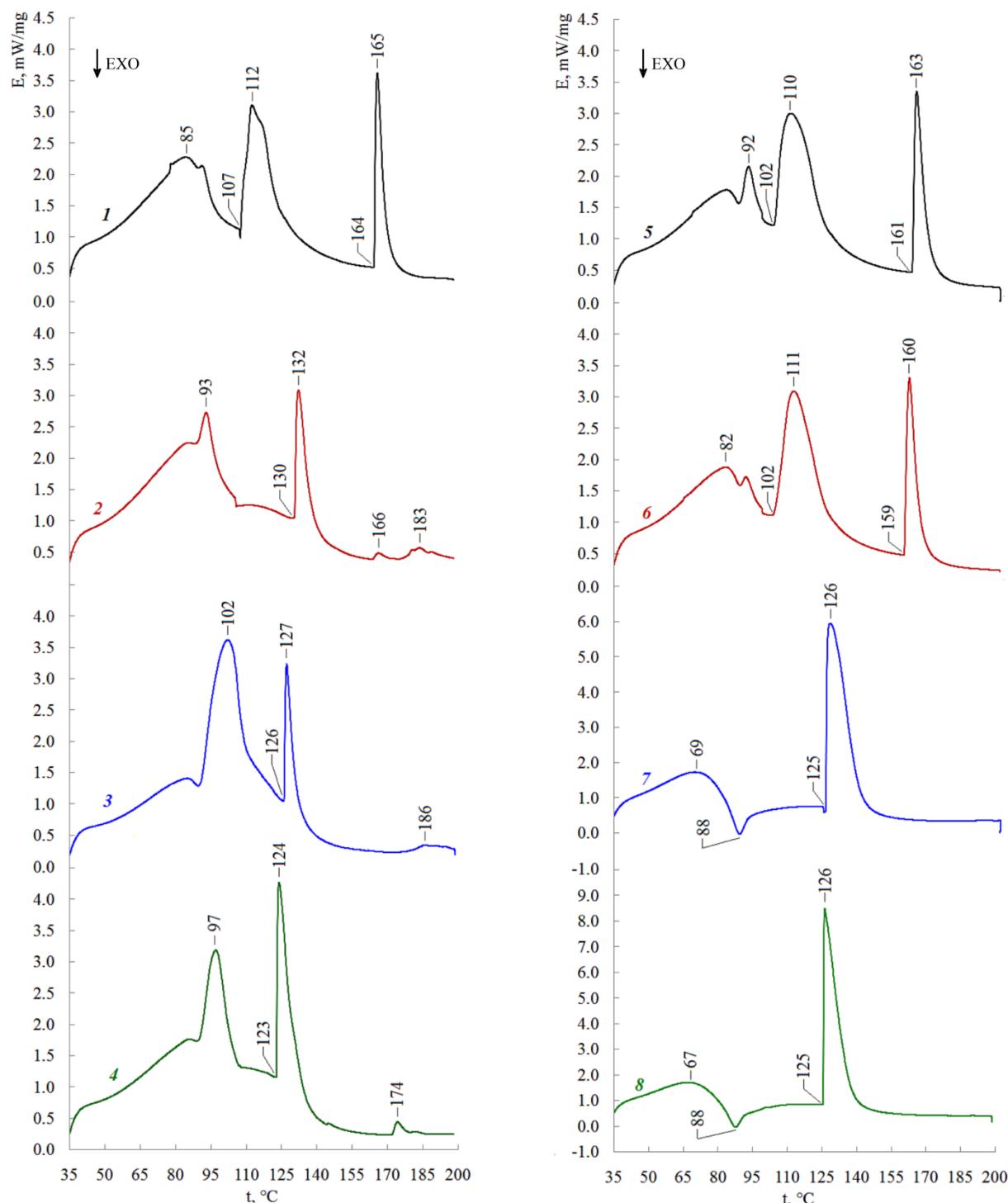


Fig. 3: DSC curves of UF adhesive with different composition: 1 – 1 % ammonium sulfate without urea; 2 – 1 % ammonium sulfate and 0.43 % urea; 3 – 1 % ammonium sulfate and 1.29 % urea; 4 – 1 % ammonium sulfate and 2.15 % urea; 5 – 3 % ammonium sulfate without urea; 6 – 3 % ammonium sulfate and 0.43 % urea; 7 – 3 % ammonium sulfate and 1.29 % urea; 8 – 1 % ammonium sulfate and 2.15 % urea

Thermal curing behavior of resin with MC-1(0.5) and MC-1(1.5) has a more complex dependency (Fig. 4), because at an increasing the content of modifier-curing agents the amount of both curing catalyst and urea in the adhesive is increasing. Deepening of resin curing is accompanied by an increase of the urea's action. The modifier-curing agents at the low content (1%) doesn't provide sufficient intensity of polycondensation; this is evidenced by both long curing time of adhesive (Fig. 1, curves 2, 3) and DSC data (Fig. 4, curves 1, 2, 3).

4); perhaps due to low yield of polycondensation's by-products the endothermic effects of the curing are poorly. The increasing in the content of MC-1(0.5) and MC-1(1.5) leads to curing deepening; weak endothermic effects (Fig. 4, curves 3, 5) can be explained by the intensification of the exothermic reactions. The DSC curve of adhesive with 5% MC-1(0.5) is similar in shape to the curves of samples with 3% ammonium sulfate, 1.29 and 2.15% urea, but there is no exothermic peak in the 88 °C region.

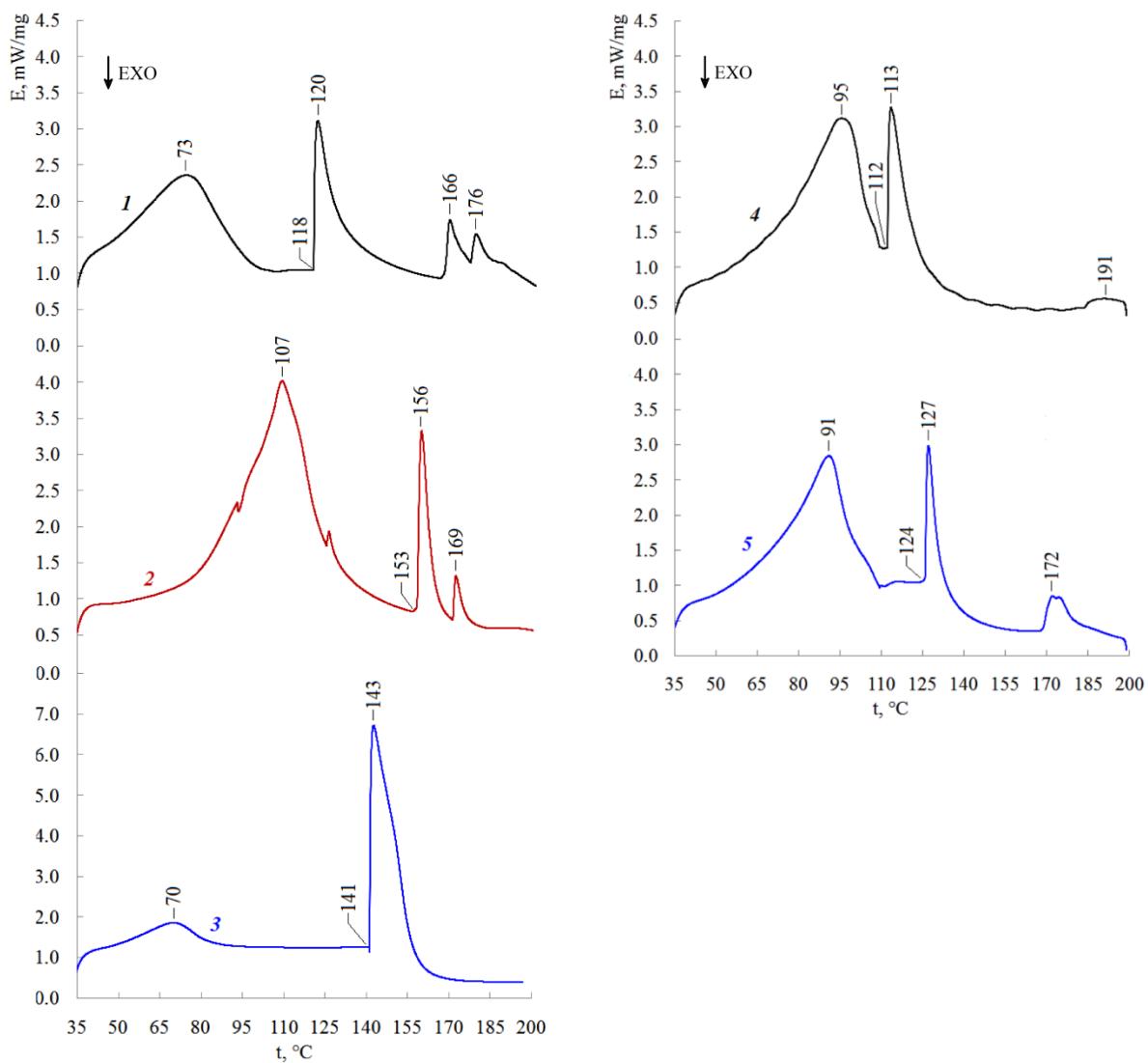


Fig. 4: DSC curves of UF adhesive with modifier-curing agent: 1 – 1 % MC-1(0.5); 2 – 3 % MC-1(0.5); 3 – 5 % MC-1(0.5); 4 – 1 % MC-1(1.5); 5 – 5 % MC-1(0.5)

Solid-state ¹³C NMR spectra showed significant differences to urea's influence to the UF polymer structural at different ammonium sulfate content (Fig. 5). At catalyst amount 1 % the adding of 1.29 % scavenger leads to decreases the amount of type

I methylene bridges, increases amount of methylol groups and methylene ether bridges (Table 5). With a further increase of the urea content (to 2.15 %), there is a sharp increase in the amount of type I methylene bridges (but not

to initial level) and decreases amount of methylene ether bridges. Probably, in combination with 1 % ammonium sulfate urea reacts mainly to free formaldehyde; but at a high content of scavenger (2.15 %), its interaction with resulting methylolureas is possible with the formation of type I methylene bridges.

At the higher curing agent content (3 %) the effects from the ureas action changes; the UF

polymer contains a high amount of not only methylol groups and methylene ether bridges, but also type I methylene bridges. Perhaps in this case urea reacts not only with free formaldehyde but also with methylol groups of methylolureas and UF oligomers. Thus scavenger participates in curing along with other resin components, which is consistent with the DSC data.

Table 4: Peak Temperatures of the Studied Adhesives by DSC Measurements

| Samples | | Peak's area, °C | Top of the peak, °C | ΔH, J/g |
|---------------------------------------|------|-----------------|---------------------|---------|
| 1 % ammonium sulfate, at urea content | 0.00 | 35-164 | 85, 112* | 719 |
| | 0.43 | 35-130 | 93 | 436 |
| | 1.29 | 35-125 | 102 | 439 |
| | 2.15 | 35-122 | 97 | 363 |
| 3 % ammonium sulfate, at urea content | 0.00 | 36-157 | 92, 110* | 717 |
| | 0.43 | 35-158 | 82, 111* | 690 |
| | 1.29 | 35-85 | 69 | 215 |
| | 2.15 | 35-84 | 67 | 216 |
| MC-1(0.5) content | 1 | 35-118 | 73 | 309 |
| | 3 | 35-153 | 107 | 705 |
| | 5 | 35-92 | 70 | 262 |
| MC-1(1.5) content | 1 | 35-110 | 95 | 389 |
| | 5 | 35-124 | 91 | 365 |

* Two peaks were observed in the considered area

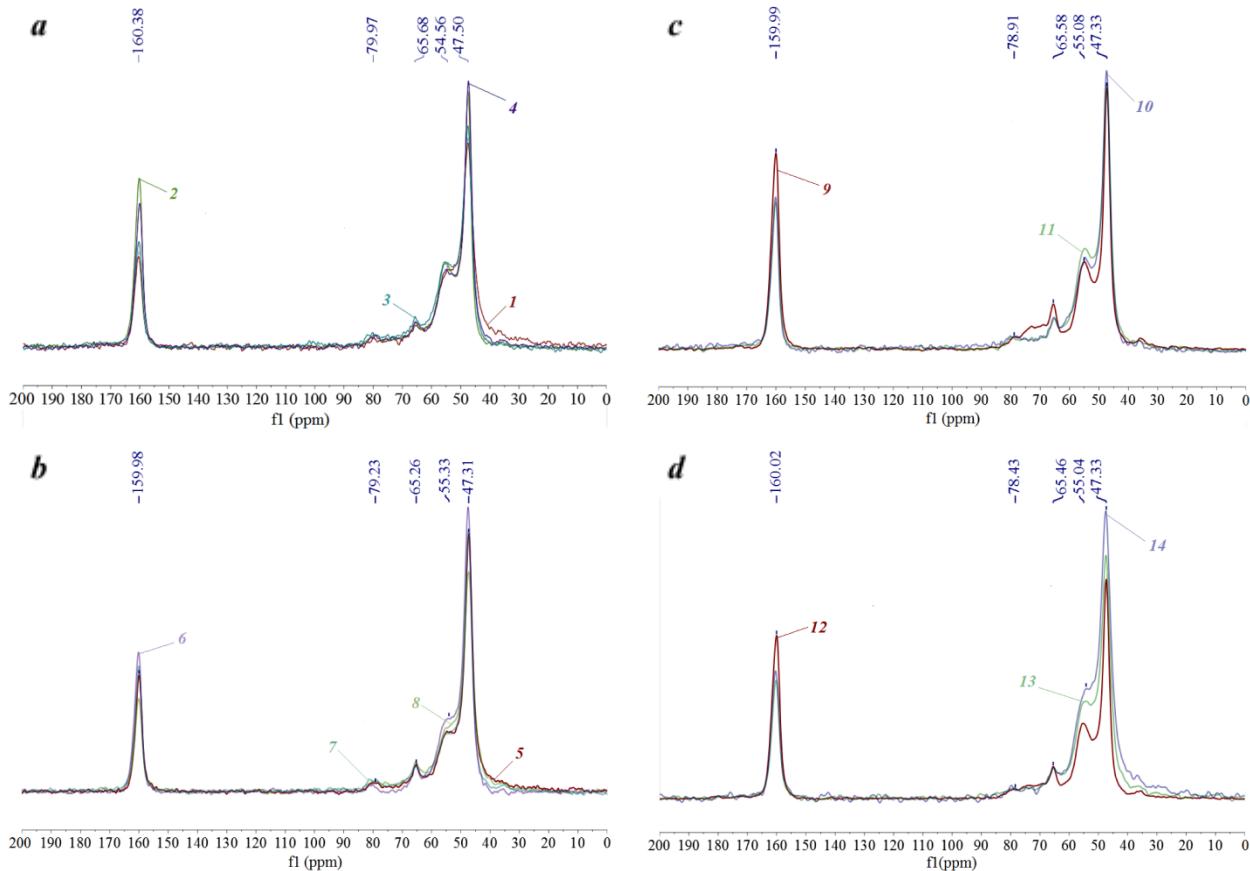


Fig. 5: Solid-state ^{13}C NMR spectra of UF resin cured of (a) 1 % ammonium sulfate, (b) 3 % ammonium sulfate, (c) MC-1(0.5), (d) MC-1(1.5): 1, 5 – UF resin without urea; 2, 6 – UF resin with 0.43 % of urea; 3, 7 – UF resin with 1.29 % of urea; 4, 8 – UF resin with 2.18 % of urea; 9, 12 – UF resin with 1 % of modifier-curing agents; 10, 13 – UF resin with 3 % of modifier-curing agents; 11, 14 – UF resin with 5 % of modifier-curing agents

Perhaps at low curing agent content urea act as a classic formaldehyde scavenger, that is react with formaldehyde with harmless products forming. At high curing agent content the activity of urea increased and its starts interacting with other resin components (methyloloureas mainly); in that case urea act as a modifier of UF resin by getting involved to the curing process. The UF polymer is characterized by a relatively low methylene bridge type II amount, thus it has an insufficiently dense cross-linking network, which can cause reduced water resistance of particleboard.

Similar conclusions were made based on the results of works [5, 26]; the authors suggested that as a result of the interaction of urea with methylolureas, linear polymers are formed that are weakly involved in cross-linking reactions. In this case, the structure of the cured resin is largely

formed due to physical bonds that are vulnerable to the action of water in the case of amorphous polymers. Vyukov and Vasiliev [27] argue that in case of an excess of acid, free urea is hydrolyzed to form ammonia, which in turn neutralizes formaldehyde. However, due to the destruction of urea and an increase in pH (due to the formation of ammonia), the resin lacks internal resources for curing.

Table 5: Structures and Chemical Shift of UF Resin Cured Using Ammonium Sulfate and Urea

| Grope | Chemical Shift (ppm) | Methylenic carbons content, %, at urea mass fraction | | | | | | | |
|---|----------------------|--|------|------|------|----------------------|------|------|------|
| | | 1 % ammonium sulfate | | | | 3 % ammonium sulfate | | | |
| | | 0.00 | 0.43 | 1.29 | 2.15 | 0.00 | 0.43 | 1.29 | 2.15 |
| -NH-CH ₂ -NH- | 40-52 | 62.5 | 52.9 | 53.9 | 59.4 | 59.3 | 67.9 | 69.0 | 67.6 |
| -NH-CH ₂ -NH< | 52-61 | 26.7 | 32.9 | 30.5 | 27.3 | 26.3 | 18.7 | 16.9 | 18.6 |
| -NH-CH ₂ OH | 62-70 | 6.4 | 8.6 | 7.0 | 7.2 | 7.5 | 9.1 | 9.0 | 8.2 |
| >N-CH ₂ -O-CH ₂ -N< | 72-80 | 4.4 | 5.7 | 8.5 | 6.1 | 7.0 | 4.3 | 5.2 | 5.6 |
| Total | | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Type I methylene bridges; ** Type II methylene bridges

The ¹³C NMR spectroscopy data (Table 6) confirm the deepening of the curing of the UF resin with an increase of the modifier-curing agent's content. With an increase of the MC-1(1.5) content, the amount of methylene bridges type I increased, the amount of methylol groups, methylene ether bridges and methylene bridges type II decreased.

Wherein, the urea participates in curing mainly through react with methylolurea and terminal methylol groups of UF oligomers, which leads to increased of methylen bridges type I amount (even compared to adhesive cured with ammonium sulfate without urea).

Table 6: Properties of Model Particleboards Made Using Urea

| Grope | Chemical Shift (ppm) | Methylenic Carbons Content, %, at MC Mass Fraction | | | | | |
|---|----------------------|--|------|------|-----------|------|------|
| | | MC-1(0.5) | | | MC-1(1.5) | | |
| | | 1 | 3 | 5 | 1 | 3 | 5 |
| -NH-CH ₂ -NH- | 40-52 | 48.4 | 58.7 | 57.1 | 51.6 | 59.5 | 64.6 |
| -NH-CH ₂ -NH< | 52-61 | 29.5 | 29.3 | 28.6 | 30.5 | 25.7 | 24.3 |
| -NH-CH ₂ OH | 62-70 | 11.6 | 6.7 | 5.9 | 9.2 | 8.1 | 6.9 |
| >N-CH ₂ -O-CH ₂ -N< | 72-80 | 10.5 | 5.3 | 8.4 | 8.8 | 6.8 | 4.2 |
| Total | | 100 | 100 | 100 | 100 | 100 | 100 |

Action of MC-1(0.5) at an amount of 1 and 3% is similar to the effect of MC-1(1.5), however, UF polymers with 5% of various modifier-curing agents are different. The adhesive cured with 5% MC-1(0.5) has more methylol groups, methylene ether bridges and methylene bridges of type II, as well as fewer methylene bridges of type I. Perhaps due to the very fast curing (Fig. 1, curve 3), the urea from the modifier-curing agent doesn't have time to react with the resin components, since the rate of their interaction with each other is higher. This explains the absence of an exothermic peak to the DSC curve of an adhesive with 5% MC-1(0.5) compared to samples cured of 3% ammonium sulfate with 1.29 and 2.15% urea.

According the results of model particleboards tests (Table 7) the increasing of the urea's amount

at low ammonium sulfate content leads to toxicity loss and increase of bending strength and water resistance. An increase of the curing agent content leads to a deterioration in the physical and mechanical properties of the boards mainly thickness swelling. Urea's influence to the toxicity of particleboard decreased at the high ammonium sulfate content; when 1 % catalyst used formaldehyde content reduced to 23 % (from 21.9 to 16.8 mg/100 g), when 3 % catalysts used to 9 % (from 16.4 to 14.9 mg/100 g). It is possible that the effectiveness of reducing formaldehyde emissions when free urea is included in the curing process is lower than when it acts as a classic scavenger.

Table 7: Properties of Model Particleboards Made Using Urea

| Properties | Urea content, % | | | | | | | |
|--------------------------------|----------------------|------|------|------|----------------------|------|------|------|
| | 1 % ammonium sulfate | | | | 3 % ammonium sulfate | | | |
| | 0.00 | 0.43 | 1.29 | 2.15 | 0.00 | 0.43 | 1.29 | 2.15 |
| Density, kg/m ³ | 723 | 751 | 758 | 786 | 709 | 731 | 703 | 724 |
| Bending strength, MPa | 34.9 | 38.5 | 39.4 | 41.5 | 30.0 | 36.9 | 36.0 | 36.4 |
| Water absorption, % | 72 | 62 | 61 | 62 | 68 | 59 | 51 | 70 |
| Thickness swelling, % | 28 | 25 | 25 | 26 | 27 | 26 | 31 | 31 |
| Formaldehyde content, mg/100 g | 21.9 | 21.4 | 18.5 | 16.8 | 16.4 | 15.7 | 15.4 | 14.9 |

It should be noted that increasing of the curing agent amount also has a beneficial effect on reducing toxicity; since ammonium sulfate acts through a reaction with free formaldehyde of UF resin, it can also act as a scavenger. Similarly, Nuno Costa et al [28, 29] explained the positive effect of ammonium sulfate to the formaldehyde emission compared with organic and mineral acids. Vyukov and Vasiliev [27] claim that the curing agent also reacts with methylol groups.

The physic and mechanical properties of particleboards manufactured with MC-1(0.5) and MC-1(1.5) improve with an increase of modifier-curing agents content (Table 8), which confirms the assumption of deepening the curing of the adhesive. In terms of strength and water resistance, they are comparable to boards based on an adhesive with 3% ammonium sulfate and urea (the thickness swelling of boards with modifier-curing agents is even less), but they are significantly inferior in toxicity. The formaldehyde content in particleboards with 5% MC-1(0.5) is 5% lower than in plates with 5% MC-1(1.5), but higher than in plates with 1% ammonium sulfate and 2.15% urea by 20%.

Apparently, the effectiveness of formaldehyde neutralization depends on the pH value provided

by the curing agents. Modifier-curing agents cannot provide the same low pH of the cured resin as 3% ammonium sulfate, the effect of their action is comparable to that of 1% ammonium sulfate. At the same time, urea belonging to the modifier-curing-agents is more likely including in the structure of the curing resin rather than acting as a classic scavenger, according to the data of solid-state NMR; perhaps because the pH of the initial binder (before curing) is significantly lower than in the case of ammonium sulfate. The pH value of 4.5-5.0 is sought at the second stage of the synthesis of UF resin of the KF-MT-15 brand, just at the polycondensation stage; that is, under such conditions, it is possible to expect a predominant interaction of urea with methylolureas and UF oligomers. In this case, the thesis of a less effective mechanism for reducing toxicity based on the inclusion of urea in the structure of the curing resin is confirmed. It should also be noted that modifier-curing agents have a complex composition, namely they contain citric acid. The effect of citric acid was not considered separately in this work; however, it can be assumed that its effect on both the action of urea and the structure of the UF polymer may be significant.

Table 8: Properties of Model Particleboards Made Using Modifier-Curing Agents

| Properties | Modifier-curing agent content, % | | | | | |
|--------------------------------|----------------------------------|------|------|-----------|------|------|
| | MC-1(0.5) | | | MC-1(1.5) | | |
| | 1 | 3 | 5 | 1 | 3 | 5 |
| Density, kg/m ³ | 690 | 755 | 721 | 678 | 713 | 712 |
| Bending strength, MPa | 28.2 | 34.3 | 29.6 | 28.6 | 30.9 | 33.5 |
| Water absorption, % | 104 | 69 | 81 | 93 | 79 | 77 |
| Thickness swelling, % | 48 | 29 | 28 | 41 | 28 | 24 |
| Formaldehyde content, mg/100 g | 33.6 | 22.1 | 17.9 | 33.1 | 24.4 | 18.9 |

The test results of the three-layer particleboards partially confirmed the previously obtained data (Table 9). Samples with urea have the lowest toxicity (30% lower than that of the control – particleboard type 1), but have a high thickness swelling. Samples with modifier-curing agents

have physical and mechanical parameters comparable to the control; when using MC-1(0.5) to cure the adhesive of the core layer, it was possible to reduce the formaldehyde content by 9%. Perhaps due to the better heating of the plate, the effectiveness of urea has increased.

Table 9: Three-Layer Particleboards Properties

| Properties | Particleboard type | | | |
|--------------------------------|--------------------|------|------|------|
| | 1* | 2 | 3 | 4 |
| Density, kg/m ³ | 629 | 649 | 635 | 655 |
| Bending strength, MPa | 22.3 | 22.1 | 21.5 | 25.1 |
| Internal bonding strength, MPa | 0.33 | 0.32 | 0.36 | 0.35 |
| Water absorption, % | 115 | 118 | 110 | 100 |
| Thickness swelling, % | 38 | 42 | 36 | 37 |
| Formaldehyde content, mg/100 g | 16.4 | 11.3 | 16.3 | 14.9 |

* Control samples

The Table 9 data shown that the combination of modifier-curing agents of the first series makes it possible to obtain three-layer particleboards with satisfactory physical and mechanical properties, as well as to achieve a reduction in the formaldehyde content compared with samples with ammonium sulfate. However, a 9% reduction in toxicity cannot be considered an acceptable result; the effectiveness of the modifier-curing agents needs to be enhanced.

IV. CONCLUSION

The conclusions of this article are as follows:

1. Free urea added to the UF resin-based adhesive acts differently depending on the content of the curing agent. With an ammonium sulfate content of 1% (the pH of the adhesive after curing is 3.50-3.70), urea acts as a classic scavenger, that is, it chemically interacts with formaldehyde to form harmless products. When the content of curing agents is 3% (the pH of the adhesive after curing is 3.00-3.10), urea reacts mainly with the methylol groups of resin components, participating in the curing process. This reduces the sources of free formaldehyde amount.
2. The action of urea through embedding in the structure of the curing resin is less effective for neutralizing formaldehyde than the action as a scavenger. With an amount of ammonium

sulfate and urea of 3 and 2.15% respectively, the formaldehyde content in model particleboards decreases by 9.1% compared with plates without a modifier; with an amount of ammonium sulfate and urea of 1 and 2.15% respectively, the formaldehyde content decreases by 23% compared with plates without a modifier. At the same time, the toxicity of plates with high curing agent content is lower in absolute values due to the positive effect of ammonium sulfate.

3. Urea in the composition of modifiers-curing agents, most likely, acts through interaction with the methylol groups of resin components and inclusion in the curing process. Probably due to the low pH value of the binder created before curing (3.72-4.51 with a content of modifiers-curing agents of 5%). Model particleboards made using modifiers-curing agents have relatively high formaldehyde content; by 14-18% less than samples without urea with 1% ammonium sulfate, but by 9-15% more than samples without urea with 3% ammonium sulfate.
4. Tests of three-layer particleboards have shown that, compared with samples without modifiers, samples with urea have comparable strength, reduced toxicity (formaldehyde content is 31% lower) and water resistance (thickness swelling is 10% higher). Samples with modifiers-curing agents have comparable

strength and water resistance, but the toxicity is only 9% lower.

ACKNOWLEDGMENTS

The DSC study was conducted using the equipment of the Centre for Thermal Analysis and Calorimetry of Saint Petersburg State University Research Park.

Disclosure statement

No potential conflict of interest was reported by the authors.

REFERENCE

- George I. Mantanis, Eleftheria Th. Athanassiadou, Marius C. Barbu & Kris Wijnendaele, Adhesive systems used in the European particleboard, MDF and OSB industries. *Wood Material Science & Engineering*, 2017. DOI: <http://dx.doi.org/10.1080/17480272.2017.1396622>.
- Pizzi, A.; Papadopoulos, A.N.; Pollicardi, F., Wood Composites and Their Polymer Binders. *Polymers*, 2020, 12(5), 1115. DOI: 10.3390/polym12051115.
- Lubos Kristak, Petar Antov, Pavlo Bekhta, Muhammad Adly Rahandi Lubis, Apri Heri Iswanto, Roman Reh, Jan Sedliacik, Viktor Savov, Hamid R. Taghiyari, Antonios N. Papadopoulos, Antonio Pizzi, Aleksander Hejna. Recent progress in ultra-low formaldehyde emitting adhesive systems and formaldehyde scavengers in wood-based panels: a review. *Wood Material Science & Engineering*, 2022. DOI: 10.1080/17480272.2022.2056080.
- Antov, P., Savov, V., Neykov, N. Reduction of formaldehyde emission from engineered wood panels by formaldehyde scavengers – a review. 13th International Scientific Conference WoodEMA2020 and 31st International Scientific Conference ICWST 2020, Sustainability of forest-based industries in the global economy. Vinkovci, Croatia, 2020.
- Wang, H.; Cao, M.; Li, T.; Yang, L.; Duan, Z.; Zhou, X.; Du, G. Characterization of the Low Molar Ratio Urea–Formaldehyde Resin with ¹³C NMR and ESI–MS: Negative Effects of the Post-Added Urea on the Urea–Formaldehyde Polymers. *Polymers*, 2018, 10(6), 602. DOI: 10.3390/polym10060602.
- Vyunkov S.N. The effect of urea on the curing of urea formaldehyde resins. *Izvestia Sankt-Peterburgskoj Lesotekhniceskoj Akademii*, 2019, 226, pp. 155–161 (in Russian with English summary). DOI: 10.21266/2079-4304.2019.226.155-16.1
- Cheng Xing, S. Y. Zhang, James Deng, Siquan Wang. Urea–formaldehyde-resin gel time as affected by the pH value, solid content, and catalyst. *Journal of Applied Polymer Science*, 2007, 103(3), pp. 1566–1569. DOI: 10.1002/app.25343.
- Mohsen Khonakdar Dazmiri, Mohammad Valizadeh Kiamahalleh, Ali Dorieh, Antonio Pizzi. Effect of the initial F/U molar ratio in urea-formaldehyde resins synthesis and its influence on the performance of medium density fiberboard bonded with them. *International Journal of Adhesion and Adhesives*, 2019, 95, 102440. DOI: 10.1016/j.ijadhadh.2019.102440.
- Ali Dorieh, Nosratollah Mahmoodi, Manouchehr Mamaghani, Antonio Pizzi & Massoud Mohammadi Zeydi. Comparison of the properties of urea-formaldehyde resins by the use of formalin or urea formaldehyde condensates. *Journal of Adhesion Science and Technology*, 2019, 32(23), pp. 1–15. DOI: 10.1080/01694243.2018.1492780.
- Arif Nuryawan , Byung-Dae Park , Adya P. Singh. Comparison of thermal curing behavior of liquid and solid urea–formaldehyde resins with different formaldehyde/urea mole ratios. *Journal of Thermal Analysis and Calorimetry*, 2014, 118, pp. 397–404. DOI: 10.1007/s10973-014-3946-5.
- Byung-Dae Park, Eun-Chang Kang, Jong-Young Park. Thermal Curing Behavior of Modified Urea-Formaldehyde Resin Adhesives with Two Formaldehyde Scavengers and Their Influence on Adhesion Performance. *Journal of Thermal Analysis and Calorimetry*, 2014, 118(1). DOI: 10.1007/s10973-014-3946-5.

12. S. Boran n, M. Usta, E. Gümüşkaya. Decreasing formaldehyde emission from medium density fiberboard panels produced by adding different amine compounds to urea formaldehyde resin. *International Journal of Adhesion & Adhesives*, 2011, 31, pp. 674–678. DOI: 10.1016/j.ijadhadh.2011.06.011.
13. İstek, A., Özluşoylu, İ., Bakar, S., Öz, E. Tutkal Çözeltisine Üre İlavesinin Formaldehit Emisyonu ve Levha Özelliklerine Etkisi. II. International Scientific and Vocational Studies Congress. Ürgüp, Turkey, 2018, pp. 824–830 (in Turkish).
14. Yang, H.; Wang, H.; Du, G.; Ni, K.; Wu, Y.; Su, H.; Gao, W.; Tan, X.; Yang, Z.; Yang, L.; et al. Ureido Hyperbranched Polymer Modified Urea-Formaldehyde Resin as High- Performance Particleboard Adhesive. *Materials*, 2023, 16(11), 4021. DOI: 10.3390/ma16114021.
15. Perminova Daria A., Malkov Victor S., Kurzina Irina A., Babushkina Tatiana B. Wood composite materials based on glycoluril-modified urea-formaldehyde resins. *Vestnik Tomskogo gosudarstvennogo universiteta*, 2015, 238-241 (in Russian with English summary). DOI: 10.17223/15617793/391/38.
16. R. Paleckiene, A. Sviklas, R. Slinksiene. Reaction of Urea with Citric Acid. *Russian Journal of Applied Chemistry*, 2005, 78(10), pp. 1651–1655. DOI: 10.1007/s11167-005-0579-2.
17. Ivanov D.V., Shevchenko S.V., Ekaterincheva M.A. The hardener of urea-formaldehyde resin, which reduce toxicity of wood board. *Izvestia Sankt-Peterburgskoj Lesotehniceskoj Akademii*, 2019, 229, pp. 215–230 (in Russian with English summary). DOI: 10.21266/2079-4304.2019.229.215-230.
18. Roffael E. Formaldehyde release from particleboard and other wood based panels. Kuala Lumpur, 1993, 281 p.
19. Kantieva E., Ponomarenko L. The effect of curing agents to urea-formaldehyde resins physico-chemical properties. *Wood Boards: Theory and practice. Proceedings of the 24th International Scientific and Practical Conference*. Saint Petersburg, Russia, pp. 94–98. (in Russian). EDN: XXIVXO.
20. K. Siimer, T. Kaljuvee, P. Christjanson. Thermal behavior of urea-formaldehyde resins during curing. *Journal of Thermal Analysis and Calorimetry*, 2003, 72, pp. 607–617. DOI: 10.1023/A:1024590019244.
21. T. Zorba, E. Papadopoulou, A. Hatjissaak, K. M. Paraskevopoulos, K. Chrissafis. Urea-formaldehyde resins characterized by thermal analysis and FTIR method. *Journal of Thermal Analysis and Calorimetry*, 2008, 92(1). DOI: 10.1007/s10973-007-8731-2.
22. Ding, Z., Ding, Z., Ma, T., Zhang, H. Condensation reaction and crystallization of urea-formaldehyde resin during the curing process. *BioResources*, 2020, 15(2), pp. 2924–2936. DOI: 10.15376/biores.15.2.2924-2936.
23. Šebenik, A. & Osredkar, U. & Zigon, Majda & Vizovišek, I. Study of the reaction between Urea and Formaldehyde by DSC and ¹³C NMR Spectroscopy. *Die Angewandte Makromolekulare Chemie*, 2003, 102(1), pp. 81–85. DOI: 10.1002/apmc.1982.051020109.
24. Popović, Mlađan & Miljković, Jovan & Budinski-Simendic, Jaroslava & Pavličević, J. & Ristić, Ivan. Curing characteristics of low emission urea-formaldehyde adhesive in the presence of wood. *Wood research*, 2011, 56(4), pp. 589–600.
25. Chang Liu, Jianlin Luo, Xiaona Li, Qiang Gao, Jianzhang Li. Effects of Compounded Curing Agents on Properties and Performance of Urea Formaldehyde Resin. *Journal of Polymers and the Environment*, 2018, 26(6), pp. 158–165. DOI: 10.1007/s10924-016-0913-1.
26. Taohong Li, Jiankun Liang, Ming Cao, Xiaoshen Guo, Xiaoguang Xie, Guanben Du. Re-elucidation of the acid-catalyzed urea-formaldehyde reactions: A theoretical and ¹³C-NMR study. *Journal of Applied Polymer Science*, 2016, 133(48), DOI: 10.1002/app.44339.
27. Viunkov S.N., Vasilyev V.V. Modification of the hardener of urea-formaldehyde resins with a low molar ratio of formaldehyde to urea. *Izvestia Sankt-Peterburgskoj Lesotehniceskoj Akademii*, 2023, 243, pp. 284–296 (in Russian with English summary). DOI: 10.21266/2079-4304.2023.243.284-296.

28. N. A. Costa, J. Pereira, J. Ferra, P. Cruz, J. Martins, F. D. Magalhães, A. Mendes, L. H. Carvalho. Formaldehyde emission in wood based panels: Effect of curing reactions. *International Wood Products Journal*, 2014, 5(3), pp. 146–160. DOI: 10.1179/2042645314Y.0000000070.
29. N.A. Costa, J. Pereira, J. Martins, J. Ferra c, Paulo Cruz c, F.D. Magalhães, A. Mendes, L. H. Carvalho. Alternative to latent catalysts for curing UF resins used in the production of low formaldehyde emission wood-based panels. *International Journal of Adhesion and Adhesives*, 2012, 33, pp. 56–60. DOI: 10.1016/j.ijadhadh.2011.11.003.



Scan to know paper details and
author's profile

Development and Characterization of Synthetic Paraffin Wax Scale for Oil Field Scaling Studies

K. H. Yaradua & A. J. Abbas

University of Salford, Manchester

ABSTRACT

Paraffin scale deposition poses a significant challenge in various industries, including oil and gas production, where it leads to reduced operational efficiency and increased maintenance costs. This study focuses on the utilization of off-the-shelf household candle wax to simulate typical oilfield paraffin scale deposits and, simultaneously, to characterize them in order to gain deeper insights into their chemical, and compositional properties in relation to typical oilfield paraffin scale deposits. Controlled laboratory experiments were conducted to mimic paraffin scale deposition under conditions resembling those encountered in production systems. The deposition process was monitored and chemically characterized using advanced analytical techniques, such as Fourier-transform infrared spectroscopy (FTIR) and nuclear magnetic resonance (NMR) spectroscopy. This analysis provided valuable information about the molecular composition of the deposits, including the types of paraffin involved and the presence of other organic and inorganic compounds. In conclusion, mimicking and characterizing the simulated paraffin scale deposits provide valuable insights into their behaviour and composition, shedding light on factors influencing paraffin deposition preservation.

Keywords: NA

Classification: LCC Code: TP690-692

Language: English



Great Britain
Journals Press

LJP Copyright ID: 392933
Print ISSN: 2631-8474
Online ISSN: 2631-8482

London Journal of Engineering Research

Volume 24 | Issue 2 | Compilation 1.0



Development and Characterization of Synthetic Paraffin Wax Scale for Oil Field Scaling Studies

K. H. Yaradua^a & A. J. Abbas^a

ABSTRACT

Paraffin scale deposition poses a significant challenge in various industries, including oil and gas production, where it leads to reduced operational efficiency and increased maintenance costs. This study focuses on the utilization of off-the-shelf household candle wax to simulate typical oilfield paraffin scale deposits and, simultaneously, to characterize them in order to gain deeper insights into their chemical, and compositional properties in relation to typical oilfield paraffin scale deposits. Controlled laboratory experiments were conducted to mimic paraffin scale deposition under conditions resembling those encountered in production systems. The deposition process was monitored and chemically characterized using advanced analytical techniques, such as Fourier-transform infrared spectroscopy (FTIR) and nuclear magnetic resonance (NMR) spectroscopy. This analysis provided valuable information about the molecular composition of the deposits, including the types of paraffin involved and the presence of other organic and inorganic compounds. In conclusion, mimicking and characterizing the simulated paraffin scale deposits provide valuable insights into their behaviour and composition, shedding light on factors influencing paraffin deposition preservation.

This knowledge is critical for the development of tailored solutions to mitigate or remove paraffin scale-related challenges, ultimately leading to improved operational efficiency and reduced economic impact across various industries.

Author a ^a: Spray and Petroleum Research Group, Salford Innovation Research Centre (SIRC), School of Science, Engineering and Environment (SEE), University of Salford, Manchester, UK.

I. INTRODUCTION

Paraffin scale deposition is a pervasive and persistent challenge across various industries, with a particularly pronounced impact on the oil and gas sector. The accumulation of paraffin wax deposits in production systems poses a significant threat to operational efficiency and cost-effectiveness. These deposits, often resembling the stubborn build-up of candle wax on a holder, can lead to reduced flow rates, increased energy consumption, and elevated maintenance [9].

Consequently, Scale deposition (paraffin scale deposits) remain the biggest stumbling block for both flow assurance and energy security till date ([1], [2], [3]). In spite of significant investments in terms of both financial resources and time dedicated to addressing the issue of scale deposits, no single solution has proven universally effective for all types of scale deposits [23]. Additionally, these solutions have not consistently met the criteria of being economically viable, time-efficient, user-friendly, and safe for both rig completion and personnel, as well as environmentally friendly [24]. Accordingly, treatment options have often been limited to aggressive chemical solutions, such as the use of acids as inhibitors and dissolver ([8], [13], [22]), as well as destructive mechanical techniques like explosives, cutters, and mills [4].

In some cases, complete rig workovers have been necessary, including tubing replacement, or production has been deferred [12]. These challenges are primarily attributed to inadequate planning and the failure to incorporate effective scale management strategies (prevention) into the asset life cycle management of a field during the capital expenditure (capex) phase [18]. This lack of proactive planning has led to increased removal

and inhibition costs during the operational expenditure (Opex) phase of field development [11]. Notwithstanding, scale deposition can occur either before the deployment of inhibition measures or at the end of the inhibition treatment life [10]. This often leaves operating companies with no choice but to react with emergency removal responses, resulting in confrontational approaches to scale management.

In an attempt to unravel this enigma, this paper embarks on a journey to mimic and characterize typical oilfield paraffin scale deposits using a novel approach - the utilization of off-the-shelf household candle wax. Through controlled laboratory experiments, we endeavour to simulate the conditions and processes leading to paraffin scale deposition, closely mirroring the challenges faced in real-world oil and gas production environments [25]. The choice of household candle wax as a surrogate material for mimicking oilfield paraffin scale deposits is an intriguing one. While seemingly mundane, [26] believes candle wax shares fundamental compositional similarities with the paraffin wax compounds found in crude oil. Both consist primarily of long-chain saturated hydrocarbons, with the former representing a simplified model system for the latter.

This innovative approach offers several research advantages like the accessibility of household candle wax been readily available and affordable, making it a cost-effective alternative to using authentic oilfield paraffin wax for experimental purposes. This accessibility as mentioned by [24] facilitates wider and more diverse research efforts. In addition to controlled experiments advantages of candle wax enables precise control over the experimental parameters, allowing researchers to isolate and manipulate specific variables to gain deeper insights into paraffin deposition mechanisms [27]. So also, safety in terms of candle wax been non-toxic and poses minimal environmental hazards, making it a safer option for laboratory studies compared to working with authentic oilfield paraffin wax.

In addition to the aforementioned advantages, the intricate interplay of factors influencing real oil

field paraffin deposition preservation (sample extraction, transportation, etc.) presents a complex puzzle [6] that has long confounded researchers and engineers alike: Thereby, significantly reducing impact and equity of the experimental outcomes and results as a result of factor such as : First and famous temperature fluctuations due to atmospheric conditions that can cause wax to change states, transitioning between solid and liquid phases as scholarly particularized by [21] that Paraffin deposits typically form at lower temperatures, and abrupt increases in temperature during an experiment can lead to the dissolution of wax, affecting the reproducibility and accuracy of results. Secondly, is oxidation as a result of exposure to oxygen in the atmosphere that can trigger the oxidation of both the oil and the wax. This chemical reaction can alter the properties of the wax and make it more resistant to deposition [28]. Thus researchers must carefully consider the potential effects of oxidation on the wax's behaviour and characteristics. Thirdly, is contaminants challenge as a result of atmospheric conditions that expose the wax to various contaminants such as dust, particulate matter, and other airborne impurities [17]. These contaminants can interact with the wax, affecting its properties and potentially accelerating or altering deposition processes.

Fourthly, is the challenges of moisture introduced by atmospheric conditions that can interact with both the wax and the oil [20]. Water can change the physical and chemical properties of both substances, leading to unintended outcomes in the experiment. Fifthly, is UV exposure If the experiments are conducted outdoors or under natural light, exposure to ultraviolet (UV) radiation can impact the wax's stability and properties [14]. UV exposure can cause photo degradation, which can result in alterations to the wax's molecular structure, further complicating the mimicking process. The sixth challenge as itemised in [15] is related to wind and airflow that can significantly influence the distribution of wax deposits during experiments. These factors may cause uneven deposition and affect the repeatability and consistency of results, necessitating careful consideration and control.

The seventh challenge includes collecting and analysing samples under atmospheric conditions that can introduce errors and inaccuracies [29]. Researchers must meticulously plan their sampling and analysis procedures to minimize such issues and ensure data reliability. Finally, is the duration of experiments knowing that in real oilfields, wax deposition occurs over extended periods, often years [5]. Hence experiments conducted in atmospheric conditions may not fully replicate the long-term deposition processes.

Researchers must consider the duration of their experiments and potential deviations from real-world scenarios.

Therefore, by acknowledging and addressing these challenges, we can better navigate the intricate path towards understanding paraffin scale deposits. The utilization of off-the-shelf household candle wax as a proxy material represents a creative and pragmatic approach to conducting controlled experiments that closely mirror the conditions encountered in the oil and gas industry [30]. Through this innovative research endeavour, we aim to shed light on the behaviour, composition, and preservation factors influencing paraffin deposition, ultimately paving the way for the development of tailored solutions to mitigate or remove paraffin scale-related challenges. These solutions hold the promise of improved operational efficiency and reduced economic impact across various industries, from oil and gas production to beyond.

II. METHODOLOGY

The methodology for this technical paper is structured into two phases: Phase one, which focuses on Paraffin Scale Deposit Mimicking, and Phase two, titled "Characterization of the produced paraffin scale replica. Each phase consists of specific stages to systematically replicate and characterize paraffin scale deposits for research and analysis. This comprehensive methodology ensures that the paraffin scale deposits produced and characterized using household candle wax are systematically replicated and thoroughly analysed, providing valuable insights into their properties and

behaviour for research and practical applications in the oil and gas industry.

2.1 Phase 1: Paraffin Scale Deposit Mimicking

This phase concentrate on mimicking to simulates the oil type paraffin deposit using house hold candles in stage two after designing and construction of the scale fabrication moulder in stage one for the replication.

2.1.1 Stage 1: Wax Moulder Design

Solid Works software was utilized to craft a multi-purpose holder capable of transforming between a hollow and a solid form, specifically designed for the preparation of wax scale samples. The design was first developed in separate components and subsequently assembled. The choice of material for the mould was made with the consideration of its ability to endure the wax's melting temperature. You can refer to [23]: Appendix A:15 and A:16 for comprehensive 2D and 3D design drawings created using SolidWorks, along with images depicting the individual components of the mould and the fully assembled mould for both solid and hollow shapes, as illustrated in Figure 2.1



Figure 2.1: Pictures of Wax Molder (a) Component and (b) Assembled[23]

2.1.2 Stage 2: Wax Scale Preparation

To prepare the wax scale sample, household candles, with a similar low API gravity to paraffin, were utilized. The candles were cut into pieces and melted, then carefully embedded into the appropriately designed moulder to achieve the desired shapes. The fully assembled moulder produces a 150mm outside diameter and a 110mm inner diameter (20mm thickness) hollow wax sample, simulating the early-stage growth of scale in production tubing before complete blockage. Removing the centre component of the moulder results in a 150mm diameter solid sample with a 40mm thickness, simulating complete tubing blockage. The production process is as follows as it schematically shown in Figure 2.2

1. Household candles were cut into smaller pieces and placed in a metal baking pan.
2. The pan was put in an oven and set to 120°C.
3. The melted wax was carefully removed from the oven and allowed to cool and solidify

slightly to reduce the risk of slippage when poured into the moulder.

4. The melted wax was poured from the metal baking pan into the desired scale moulder.
5. The wax was allowed to cool and solidify in the desired shape before removal from the mould.
6. Each sample was marked with a reference number on both sides and weighed.
7. Individual sample details were recorded in the scale identification register.

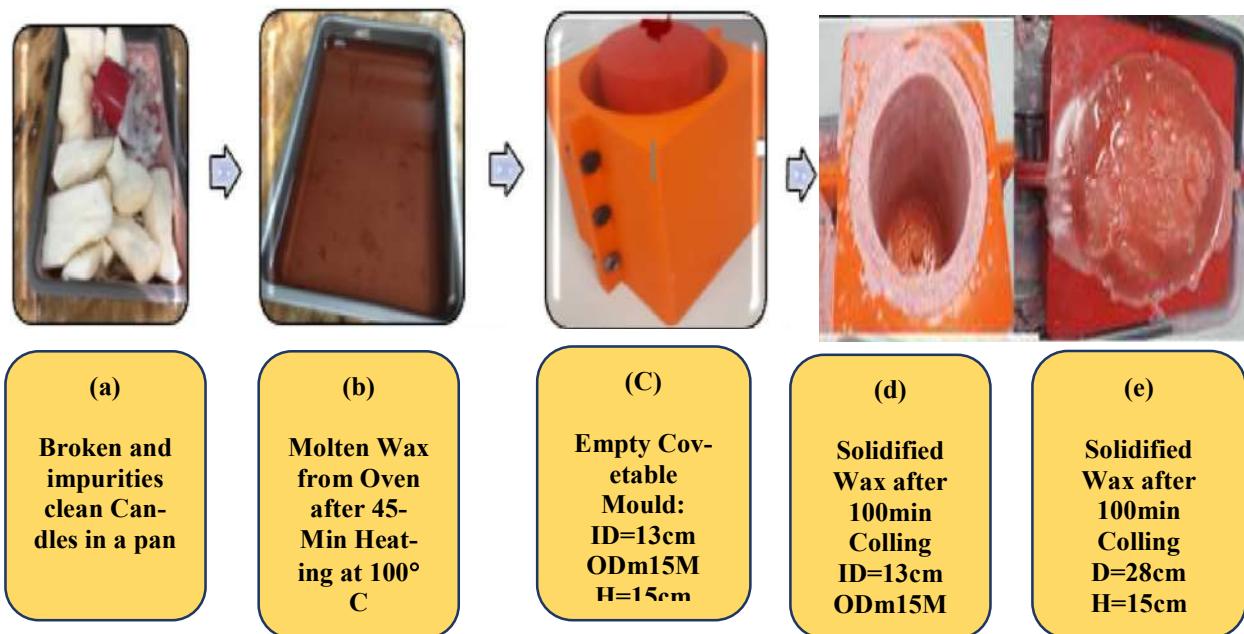


Figure 2.2: Constructed wax scale samples preparation Procedure

2.1.3 Stage: 3 Colling Curve Taste

To ensure that the fabricated scale samples were representative of typical oil field paraffin scales, a cooling curve matching test was performed on the molten waxes. The testing materials and apparatus were molten wax and long probe waterproof digital thermometer, respectively. The validation test involved obtaining the cooling curve for the wax samples and matching it with established cooling curves of typical oil field paraffin waxes. Consistency of the freezing points of the samples versus typical oilfield paraffin waxes was used as the quality indicator. Key steps of the validation test included the following.

- a) More than 75% of the thermometer probe was inserted into the molten wax, and the device was switched on.
- b) The thermometer was held steady in the above position until temperature readings stabilized.
- c) The temperature at that stage was taken and recorded.
- d) The above three steps were repeated at two minutes' intervals until the wax solidified.

2.2 Phase 2: Characterization of Produced Paraffin Scale Deposit

The manufactured soft scale samples of various dimensions and configurations, as described in

Section 2.1.1, underwent through NMR and FTIR analyses. These analyses aimed to ascertain whether they exhibited chemical properties consistent with the paraffin scale deposits typically encountered in petroleum production tubing. Figure 3.1 illustrates the fabricated soft scale samples, displaying (a) solid shapes and (b) hollow, respectively.

2.2.1 Stage 2: Chemical and Compositional Analysis

To determine the molecular composition of the deposits, the following techniques were employed:

2.2.2 Fourier-Transform Infrared Spectroscopy (FTIR)

A Thermo Scientific Nicolet iS10 Fourier-transform infrared spectroscopy (FTIR) was utilized in analysing and identifying the chemical substance, functional group in the compound of the constructed wax samples as per the working mechanism and experimental set-up in Figure 4.17 [23] using the following procedures.

1. Started by logging in to the OMNIC software windows in the FTIR analyzer dedicated computer system and select the smart FITR Diamond ATR accessory
2. The system drop window was properly clean with propanol and a lab tissue paper.

3. Then the ATR was screwed against the sample holder until a click sound is heard.
4. The 'Col BKg' was clicked on the software menu for sample background 'run' and generate background spectrum at the completion of 36 scans while monitoring the status bar
5. A small piece of the wax deposit (as in Figure 3.1) was please on the sample holder and the ATR is screwed against the sample until a click sound is heard.
6. The sample spectra were then obtained by clicking the 'Col Smp' button with 36 scan background collection as required.
7. The Selection of the peak of interest was done by putting ON the 'Analze' menu and changing the axis absorbance to transmittance and the general spectra were replaced with the new specific spectra.
8. The paraffin spectra were selected from the national institute of standard and technology (NIST) through the FTIR data base and superimposed against the wax scale spectra for comparison.
9. Step 6 to 7 were repeated for liquid paraffin analysis after pouring small quantity on the sample drop and properly screwing the ATR screw.
10. The liquid paraffin spectra generated results were superimposed with the generated wax sample spectra for comparison.

2.2.3 Nuclear Magnetic Resonance (NMR) Spectroscopy:

NMR spectroscopy was employed to provide insights into the molecular structure and composition (both aliphatic and aromatic) of the constructed wax deposits as per elaborated in the Figure 4.18 of [23] including set-up, working mechanism and procedure.

1. 10mg of starting material was used to clean the tube
2. The crushed wax sample was dissolved in 0.7ML of deuterated solvent ($CDCl_3$ in this case) as 4.5-5 cm is recommended as a suitable high of solvent for good spectrums
3. The NMR tube is carefully cap and the sample name is written on it.

4. The sample was gently shaken to ensure effective dissolving of the sample and precaution were taken to avoid contaminating the sample due to solvent cap contact.
5. The spinner and the NMR tube were clean with 2 propanol and lab tissue to wipe out all dirt and fingerprints
6. The NMR tube was gently inserted into the spinner and the spinner automatically rotate into the magnet to ensure the whole sample experience a homogenous magnetic field. The spinner is placed in a sample gauge to prevent the bottom of the NMR tube from sitting far into the NMR probe to reduce the risk of damaging the spectrometer as each probe has its own sample depth.
7. The sample was place into the NMR spectrometer and a Varian 400 MHz spectrometer equipped with an auto-sampler was utilized
8. The spectrums were processed, and peak were assigned in the spectrum.

III. RESULT AND DISCUSSIONS

Findings from the chemical and compositional analysis of the constructed wax scale samples using NMR and FTIR techniques as elaborated in Section 2.2.3 & 2.2.4 are detailed in the two coming Sections (3.3.1 and 3.3.2) respectively

3.1 Fabrication of Paraffin Wax Deposit

The constructed soft scale samples of different sizes and shapes from Section 4.3.2 of and 4.3.2 of [23] that were subjected to, NMR and FTIR analysis determine their true representation of typical oilfield paraffin deposit are showcase in Figure 3.1 Were (a) solid shape represent complete tubing blockage scenario and (b) hollow shapes represent early paraffin deposition stage respectively.



Figure 3: SEQ Figure: _3. * ARABIC 1 Constructed Soft Scale (a) Solid Shaped, (b) Hollow Shape Samples [23]

3.2 Cooling Curve Result

A temperature-versus-time graph was created from [24] to represent the cooling process. The cooling pattern and freezing point of the wax sample we made were assessed in comparison to

published paraffin cooling curves [7] using both qualitative and quantitative methods. Samples that did not meet these established criteria were not accepted.

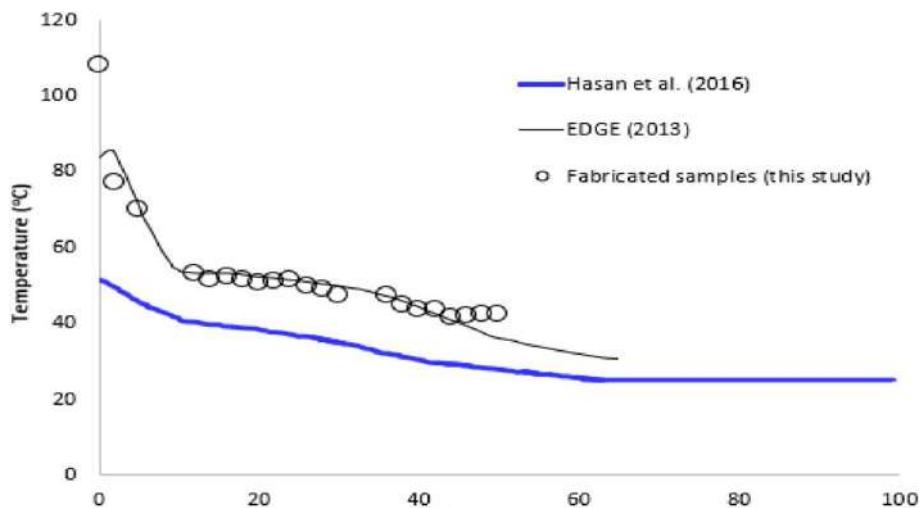


Figure 3.2: Comparison of Cooling Curves of the Scale Samples Against Those of Known Paraffin Waxes [24]

3.3 Chemical/compositional Analysis

The outcome and findings of the experiment from the chemical and compositional analysis of the constructed soft scale samples using NMR and FTIR techniques are as follows:

3.3.1 NMR Analysis of Soft Scale Sample

As in Section 2.2.3, we explain how nuclear magnetic resonance spectroscopy was employed to examine the chemical properties of typical oil field scale deposits (paraffin) within the created soft scale samples. The ^1H NMR spectra in Figure 3.3 demonstrate the existence of olefinic protons

within the range of $\delta = 0.5$ ppm to $\delta = 1.5$ ppm, which are characteristic of hydrogens found in CH , CH_2 , and CH_3 groups. This particular peak range aligns with what has been reported in the literature [19]. The singlet at $\delta = 0.0$ ppm is designated for TMS (tetramethylsilane) and primarily serves as a calibration point. The singlet peak at the far end ($\delta = 7.278$ ppm) is attributed to the deuterated chloroform (CDCl_3) solvent utilized for dissolving the sample. The distinct signals in the spectra validate the presence of saturated hydrocarbons (peaks in the upfield), and no peaks were observed in the aromatic

region of the spectra within the range of $\delta = 7.0$ ppm to $\delta = 8.0$ ppm.

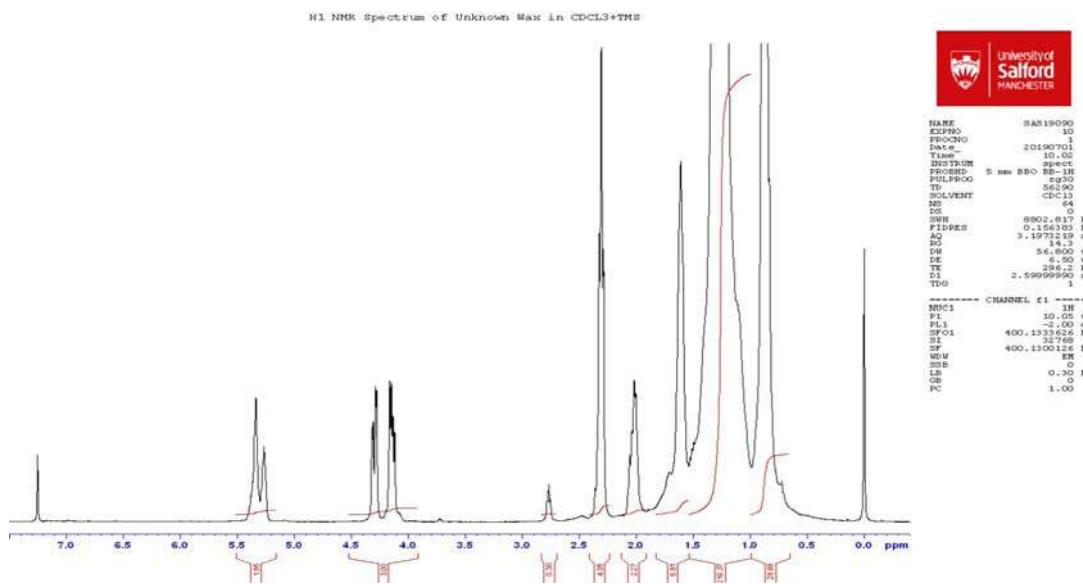


Figure 3.3: NMR Result of Constructed Wax Scale Sample [23]

3.3.2 Infrared Analysis of Soft Scale Sample

The synthesized wax scale deposit underwent additional analysis using Infrared spectroscopy to validate the NMR findings and ensure its chemical representation of the oil field scale deposit (paraffin). As described in Section 2.2.4 of the methodology chapter, the results obtained from the Thermo Scientific Nicolet iS10 for the prepared wax sample were confirmed. These results were compared to archived database in the system, which included paraffin flakes, and they appeared to share similar functional groups, as depicted in Figure 3.4. Both spectra aligned by exhibiting analogous fingerprint patterns and bands characteristic of paraffin's functional groups.

Moreover, in Figure 3.4 of the FT-IR analysis, the absorption peaks observed between 2900 cm^{-1} and 2800 cm^{-1} were attributed to the stretching and vibrations of CH_2 and CH_3 groups, affirming the presence of aliphatic paraffin in the sample, consistent with findings by [16]. These absorption peaks also corresponded with those found in the National Institute of Standards and Technology (NIST) database of FT-IR spectra.

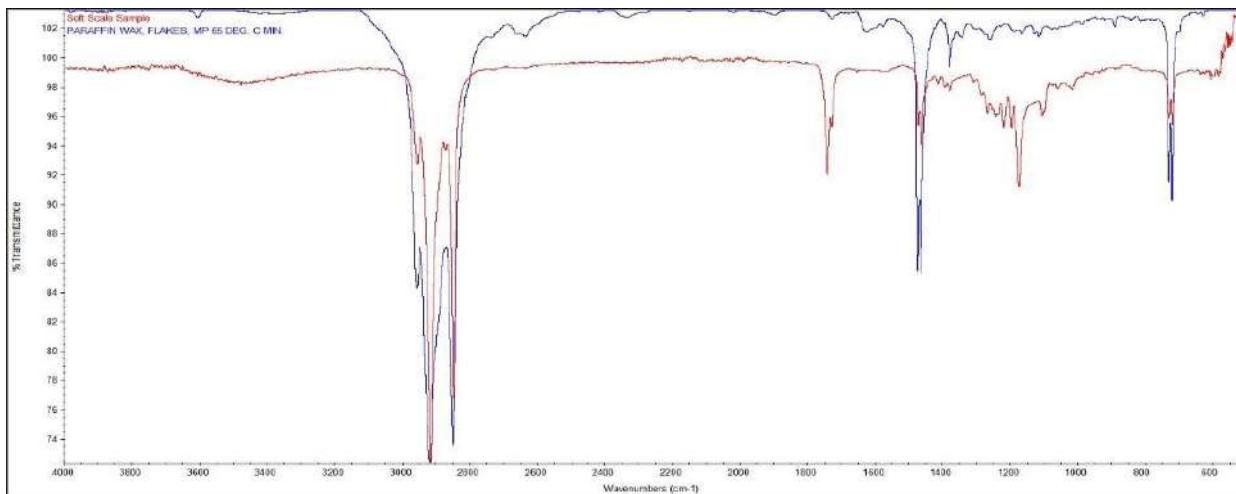


Figure 3.4: Infrared Analysis Result Compared to Paraffin Flakes From NIST Data Base [23]

Likewise, to ensure additional validation and confirmation, the spectra of the soft wax sample were overlaid and contrasted with the outcomes of the liquid paraffin confirmatory test, elaborated upon in Section 2.2.4 and displayed in Figure 3.5

In this comparison, it was evident that both the spectrum of the soft wax and the liquid paraffin exhibited identical peaks and bands characteristic of the same paraffin functional groups.

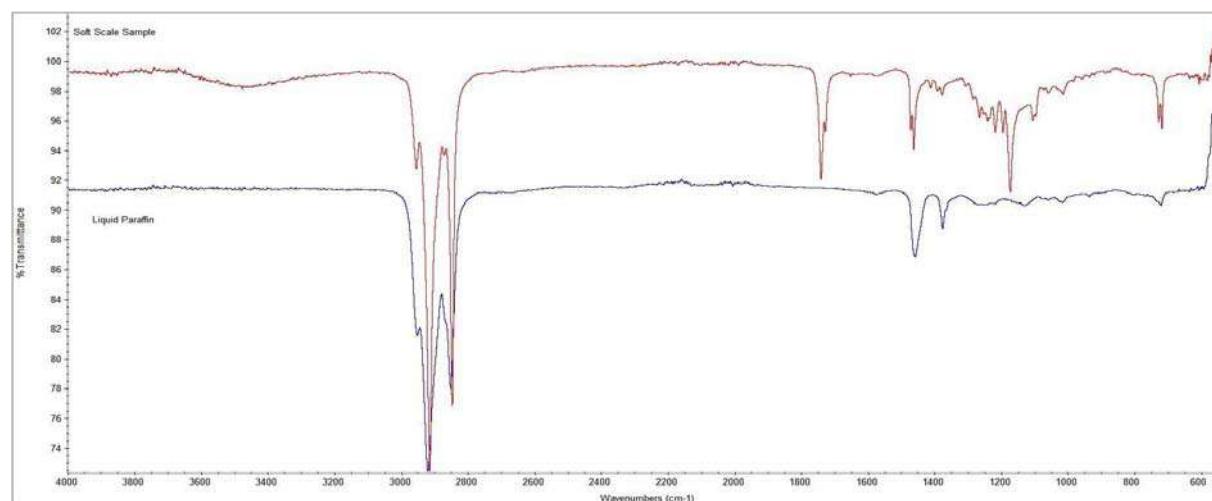


Figure 3.5: Infrared Analysis Result Compared to Liquid Paraffin [23]

VI. CONCLUSION

Taken in to consideration the aforementioned preservative and atmospheric challenges that might affect the accuracy and relevance of the experimental results. The constructed wax deposit as shown in Figure 3.1 proven to have striking some meaningful balance between the utilised controlled laboratory conditions and real- oilfield scenario from the series of compositional and chemical analysis outcomes established in Figure 3.2, 3.3, 3.4. and 3.5 respectively.

ACKNOWLEDGEMENTS

This research would not have been possible without the support of the Sparay Research group of the University of Salford. Also, more importantly Petroleum Trust Development Fund (PTDF) Nigeria for financing the research.

REFERENCES

1. Abbas AJ, Nasr GG, Burby ML, Nourian A . "Entrained Air around a High Pressure Flat Jet Water Spray". ILASS Americas, 25th

Annual Conference on Liquid Atomization and Spray Systems, May.

2. Abbas, A. J. "Descaling of Petroleum Production Tubing Utilising Aerated High-Pressure Flat Fan Water Sprays." *PhD Thesis, University of Salford, UK. 2014.*
3. Abbas A. J., Nasr G. G., Nourian A., Enyi G. C. "Impact pressure distribution in flat fan nozzles for descaling oil wells". *Paper ICLASS 2015, presented at 13th Triennial international conference on liquid atomization and spray systems, 23-27 Aug., Tainan.*
4. Alabdulmohsin, Yousif A., El-Zefzafy, Ibrahim Mohamed, Al-Malki, Mohammed H., Al-Mulhim, Abdullah A., Al-Ramadhan, Ali A., Al Hamawi, Kaisar , Arfin, Mohammad , and Danish Ahmad. "Underbalanced Coiled Tubing Mechanical De-Scaling Operations: A Case History for an Oil Well in Ghawar Field from Saudi Arabia." *PE Annual Technical Conference and Exhibition, Dubai, UAE, September 2016, pp. 1-14.*
5. Armacanqui, J. S., de Fatima Eyzaguirre, L., Flores, M. G., Zavaleta, D. E., Camacho, F. E., Grajeda, A. W, Alfaro, A. D., Viera, M. R. "Testing of Environmental Friendly Paraffin Removal Products". *SPE Latin America and Caribbean Heavy and Extra Heavy Oil Conference. 2016. pp. 1-9.*
6. Davidson, C.O., Michael, K. O., Stanley, E. N., Tobechukwu, C. O., Anaele, F. E., Chidiebere, C. O. & Victor, N. N. "Crude oil and the problem of wax deposition on pipeline systems during transportation": *A review. World Journal of Advanced Research and Reviews, 2022, https://doi.org/10.30574/wjarr.2022.15.1.0770.*
7. EDGE (2013). "100% wax cooling curve and melting point", *retrieved 21 May 2019, from http://edge.rit.edu/edge/P13411/public/Wax Test Data.*
8. Elmorse, S. A. "Challenge and Successful Application for Scale Removal Gemsa Oil Field, Egypt: Field Study." *SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain, March 2013. pp. 1-7.*
9. El Khamki MA, Nasr GG, Burby M, Nourian A. "High pressure overlapping flat sprays nozzles". *23rd Annual Conference on Liquid Atomization and Spray Systems, September 1-6. http://www.ilasseurope.org/ICLASS/ilas s2010/FILES/ABSTRACTS/0 46.pdf*
10. Farrokhrouz, M., & Asef, M. R. "Production Enhancement via Scale Removal in Nar Formation". *SPE Production and Operations Conference and Exhibition. 2010. pp. 1-9.*
11. Ghouri, S. A., Ali, I., Shah, N. M., Ejaz, M., & Haider, S. A. "Downhole Scales Management in Mature Gas Fields". *PAPG/SPE Pakistan Section Annual Technical Conference and Exhibition. 2018. pp 1-11.*
12. Guimaraes, Z., Almeida, V., Duque, L. H., Costa, G., Pinto, J. C., Chagas, J. V., Costa G. V., Franca, A. B., Alberto, C. "Efficient Offshore Scale Removal Without a Rig: Planning, Logistics, and Execution." *SPE/ICoTA Coiled Tubing and Well Intervention Conference and Exhibition. 2008. pp. 1-7.*
13. Hamdy E, Abu Bakar M, Abd El-Hay A, Sisostris S, Anwar M El Farouk O . "Challenges and Successful Application of Scale Removal In oil Field". *IPTC Paper 18139 Presented at Internation Petroleum technology Conference, 10-12-Dec., Kaula Lumpur.*
14. Janaina, I.A, Anthony, A. N, Amir, M. "Can Paraffin Wax Deposit Above Wax Appearance Temperature? A Detailed Experimental Study" .*SPE Annual Technical Conference and Exhibition, Virtual, October https://doi.org/10.2118/201297-MS.*
15. Jha, N. K., Jamal, M.S., Singh, D., & Uma S. Pr. "Characterization of Crude Oil of Upper Assam Field for Flow Assurance". *Paper presented at the SPE Saudi Arabia Section Technical Symposium and Exhibition, Al-Khobar, Saudi Arabia, https://doi.org/10.2118/172226-MS.*
16. Manoj, K., Kumar, B. and Padhy, P. K. "Characterization of Metals in Water and Sediments of Subarnarekha River along the Projects' Sites in Lower Basin, India." *Universal Journal of Environmental Research & Technology, 2(5). 2012. pp. 1-6.*
17. Mmata, B., Ajienka, J., Onyekonwu, M., & Godwin C. "Impact of Oil Base Mud Contamination OBM on Waxy Crude Properties". *SPE Nigeria Annual*

International Conference and Exhibition, Lagos, Nigeria. <https://doi.org/10.2118/184338-MS>

18. Nuhu Mohammed, Abbas .J, Godpower C E., Muhammad K. A, Onukak I. E, Bello S, Suleiman, S.M Yar'adua K.H. "Investigating the Flow Behaviour of CO₂ and N₂ in Porous Medium Using Core Flooding Experiment". *Journal of Petroleum Science and Engineering.* <https://doi.org/10.1016/j.petrol.2021.109753>.

19. Palou, R. M., Olivares-Xomelt, O. and Likhanova, N. V "Environmentally friendly corrosion inhibitors", *Developments in corrosion protection. InTech*, 2014. pp. 431–465.

20. Suleiman S.M, Abbas, A.J, Babaie. M, Akpan E.U, Yahaya A.A, Hassan K.Y."Effect of temperature and salt concentration on rheological behaviour of surfactin". *SPE NAICE Conference Lagos*.

21. Tao, Z., Wang, H., Liu, J., Zhao, W., Liu, Z., & Guo, Q. " Dual-level packaged phase change materials – thermal conductivity and mechanical properties". *Solar Energy Materials and Solar Cells*, 169(September), 222–225. <https://doi.org/10.1016/j.solmat.2017.05.030>.

22. Vazirian, M. M, Thibaut V. J. Charpentier, M, Oliveira P, Anne Ne. "Surface inorganic scale formation in oil and gas industry: As adhesion and deposition processes". *Journal of Petroleum Science and Engineering. Elsevier*. 137. 2016. pp. 22–32.

23. Yar' Adua K. H. "Descaling Petroleum Production tubing using multiple nozzles". *PhD Thesis. University of Salford, Manchester. UK*. 2020. pp. 1-25.

24. Yar'Adua KH, John IJ, Abbas AJ, Lawal KA, Kabir A. "An experimental study on high-pressure water jets for paraffin scale removal in partially blocked production tubings". *Journal of Petroleum Exploration and Production Technology*. <https://doi.org/10.1007/s13202-020-01009-w>.

25. Yar' Adua, K. H., Abbas, A. J., Salihu. S. M., John, I. J., & Aisha. K.Y. "Investigating the Effect of Varying Tubing Air Concentration during the Descaling of Petroleum Production Tubing using Multiple High-Pressure Nozzles". *American Scentific Research Journal for Engineering, Technology and Science*, 2021. Vol 76 No 1. pp. 1-9

26. Yar'Adua, K. H., Abbas, A. J., Salihu.S.M., A. Ahmadu.,& Aisha.K.Y. "An experimental approach to the removal of paraffin scales in production tubing using high pressure flat fan nozzles". *Journal for Petroleum and Gas Engineering* 2021. <https://doi.org/10.5897/JPGE2020.034>.

27. Yar'Adua, K. H., Abbas, A. J., John, I. J., Salihu.S.M., A. Ahmadu.,& Aisha.K.Y. "Laboratory-Scale Investigation of the Utilization Multiple Flat Fan Nozzles in Descaling Petroleum Production Tubing". *S.P. E Nigerian Annual International Conference and Exhibition NAICES 2021*. <https://doi.org/10.2118/207203-MS>.

28. Yar'Adua, K. H., Abbas, A. J., John, I. J., Salihu.S.M., A. Ahmadu., & Aisha.K.Y. "Investigating the impact of non-hydrodynamic connected descaling parameters in removing different stages of paraffin deposits from petroleum production tubing with multiple high-pressure spray". *American Scientific Research Journal for Engineering, technology and Science*, Vol .80, No.1, 34-51;

29. Yar'Adua, K. H., Abbas. "Hydraulic Techniques of Descaling Paraffin Inflicted Petroleum Production Tubing" <http://www.eliviapress.com/en/book/book-40300298875/>

30. Yar'Adua, K.H. "Determining the optimum descaling Requirement for effective descaling of different types of scale deposits using multiple jets". *American Scientific Research Journal for Engineering, technology and Science*, Vol .85, No.1, 319-347;

This page is intentionally left blank



Scan to know paper details and
author's profile

An Overview of Outlier Detection Methods

Dr. Maciej Celi Ski

ABSTRACT

One of the first important steps in achieving informed data analysis is detection of outliers. Even in cases where the final values are often considered to be incorrect calculations or noise, they can still provide very important information in some cases. Therefore, it is very important to detect them before modeling and analysis. In this paper, we present a structured and comprehensive review of residual detection research. There are many different methods, hence the purpose of this article is to help the novice researcher to formulate his ideas and gain an easier understanding of the various lines of research in which research has been conducted on this topic.

Keywords: machine learning, outlier, outlier detection.

Classification: LCC Code: QA76.9.D32

Language: English



Great Britain
Journals Press

LJP Copyright ID: 392934
Print ISSN: 2631-8474
Online ISSN: 2631-8482

London Journal of Engineering Research

Volume 24 | Issue 2 | Compilation 1.0



An Overview of Outlier Detection Methods

Dr. Maciej Celiński

ABSTRACT

One of the first important steps in achieving informed data analysis is detection of outliers. Even in cases where the final values are often considered to be incorrect calculations or noise, they can still provide very important information in some cases. Therefore, it is very important to detect them before modeling and analysis. In this paper, we present a structured and comprehensive review of residual detection research. There are many different methods, hence the purpose of this article is to help the novice researcher to formulate his ideas and gain an easier understanding of the various lines of research in which research has been conducted on this topic.

Keywords: machine learning, outlier, outlier detection.

I. INTRODUCTION

In the era of advanced technology and rapidly evolving digital landscapes, the need for robust anomaly detection techniques has become paramount across various domains. Among these techniques, Isolation Forests (IF) have emerged as a versatile and powerful tool with applications spanning cybersecurity, fault monitoring, big data analysis, fraud detection, and more. This collection of texts delves into the myriad ways Isolation Forests have been harnessed to address specific challenges and enhance various aspects of modern life.

As we delve into the texts that follow, we will uncover the multifaceted role that Isolation Forests play in ensuring the security of computer systems, identifying fraudulent activities, monitoring the health of mechanical systems, and analyzing vast datasets for hidden insights. The texts also highlight the adaptability and effectiveness of Isolation Forests in addressing the evolving demands of these diverse fields.

Our journey begins by exploring the significant role Isolation Forests play in safeguarding the digital realm. We delve into their application in Intrusion Detection Systems (IDS), where they excel in identifying unauthorized access and novel cyber threats, bolstering the security of sensitive data and computer systems.

Next, we venture into the domain of fault monitoring and system diagnostics, where Isolation Forests offer invaluable capabilities in detecting anomalies and irregularities across various mechanical and technical systems. Their rapid anomaly detection and response prove indispensable in industries where system failures can lead to substantial financial losses.

In the context of big data analysis, Isolation Forests provide a lens through which we can uncover hidden patterns and correlations within extensive datasets. This facilitates more informed decision-making in the business world, offering opportunities for comprehensive data exploration.

Our exploration then takes us into the realm of fraud and abuse detection, where Isolation Forests excel in identifying counterfeit reviews, misuse, and fraudulent activities, particularly in the context of customer service. Their role extends not only to consumer protection but also to upholding product and service integrity.

Lastly, we reflect on the broader significance of Isolation Forests, underlining their efficacy and potential in diverse fields. This collection of texts serves as a testament to the ever-evolving nature of Isolation Forests, opening new horizons for researchers and professionals alike.

Together, these texts paint a comprehensive picture of the significance and multifaceted applications of Isolation Forests, offering insights into how this powerful anomaly detection technique continues to shape and protect our digital world.

What is already Known about this Topic:

- Anomaly detection helps in data analysis.
- Anomaly detection saves time needed for subsequent analysis of this data, which is essential.
- Anomaly detection is widely applicable to various types of data.

What this Paper Adds:

- Allows us to notice the diversity of applications of anomaly detection methods.
- Constitutes a valuable collection of information about current methods of detecting anomalies.
- Is particularly important in the search for knowledge in the field of applications of fuzzy methods.

Implications for Practice and/or Policy:

- Is valuable for those wishing to conduct a comprehensive review of the current literature on this topic.
- Allows you to understand current trends in the use of fuzzy methods.
- It allows you to notice the potential of using fuzzy methods in various aspects of human life and its environment.

The purpose of this paper is to provide a comprehensive review of the literature on Outlier Detection and the current outlier detection techniques used in various areas. The methodology for this review presented in this paper is as follows:

- We select articles from 2018-2023 (94 works) from leading online scientific databases such as MDPI, Science Direct, IEEE Explorer, etc.).
- In addition, the work includes some required and particularly important and original articles from before 2018, in the context of the following keywords: outlier detection, anomaly detection.
- In our work we choose the most relevant and up-to-date articles that focus on the subject of our analyses.

The texts can be categorized into several thematic areas related primarily to the applications of Isolation Forests in various domains. These areas include:

1. *Cybersecurity*: Texts in this category focus on the role of Isolation Forests in Intrusion Detection Systems (IDS) and the identification of unauthorized activities and cyberattacks in the field of computer security.
2. *Fault Monitoring and State Diagnostics*: This category contains texts that describe the use of Isolation Forests for monitoring the state of devices and detecting faults in various mechanical and technical systems.
3. *Big Data Analysis*: Texts in this category concentrate on the applications of Isolation Forests in the analysis of large datasets, enabling the identification of unusual patterns and hidden dependencies in data.
4. *Fraud and Abuse Detection*: This category encompasses texts discussing the role of Isolation Forests in detecting fake product reviews, abuse, and fraud, particularly in the context of customer service.
5. *General Category*: In this category, there is a general discussion about the advantages and potential of Isolation Forests in various domains, emphasizing their effectiveness and significance. It's worth noting that the texts in each of these categories address different applications of Isolation Forests and their impact on specific fields.

II. SECTION TWO

2.1 Anomaly Detection Methodology

In the texts below, several important anomaly detection methods, widely utilized across various domains, are described. Here are the most significant of these methods.

Isolation Forests (IF): This is a pivotal method that appears in all the texts. Isolation Forests are an algorithm based on decision trees, exceptionally efficient in identifying outlier observations. This algorithm isolates anomalous points by employing a straightforward logic of partitioning data into smaller subgroups.

Local Outlier Factor (LOF): Another popular method for outlier detection. LOF assesses how isolated each data point is compared to its nearest neighbors. Points with low LOF scores are considered normal, while those with high LOF scores are deemed outliers.

One-Class Support Vector Machine (One-Class SVM): This approach utilizes Support Vector Machines (SVM) for data classification into two categories: normal and anomalies. The One-Class SVM establishes a boundary around the normal data and endeavors to ascertain whether data points situated beyond this boundary qualify as anomalies.

Histogram-Based Outlier Detection (HBOS): This method is based on creating histograms from data and utilizes differences in the distribution of normal and outlier data to identify anomalies.

Decision Trees (DT): Decision trees are used in some texts as part of more complex anomaly detection techniques, such as algorithm combinations.

Logistic Regression (LR): Logistic regression is also used as part of more advanced anomaly detection methods, often for classifying data as normal or outliers.

K-Means: In some texts, K-Means is employed as a data clustering algorithm, which can aid in identifying outlier groups or clusters.

Radial Basis Function Networks (RBF): RBF is used as a model to classify data as normal or outliers based on their distance from the cluster center.

Multi-Layer Perceptron (MLP): MLP is used in the context of data classification and is one of the neural network models employed for outlier detection.

These methods are applied in various contexts, from cybersecurity to big data analysis and fraud detection. Each of these techniques possesses unique advantages and applications, enabling researchers to choose the most suitable method depending on the specific problem and type of data they are currently working with.

Figure 1 below presents the percentage of pseudocode, source code or the lack of these two elements provided by researchers in the works discussed.

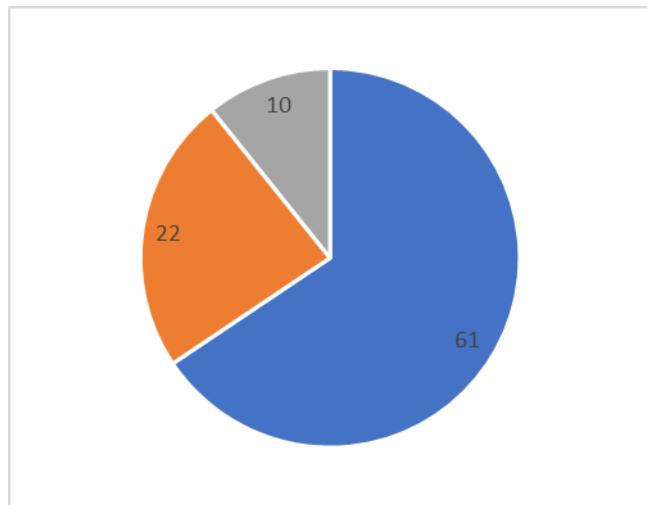


Fig. 1: Gray Color - Source Code Provided, Yellow Color - Pseudocode Included in the Work, Blue - No Source Code or Pseudocode

III. SECTION HEAD

3.1 Outlier Detection Literature Review

In [1] the paper, the authors introduce an approach to identify and diagnose irregularities through a reduction in their dimensionality, variable selection, and the application of an

isolation forest algorithm. The researchers focused their attention on incorporating isolated variables, which might go unnoticed when employing traditional techniques like Principal Component Analysis (PCA). Within the article, the authors delineate a technique for autonomously identifying and diagnosing anomalies using

historical OES data, with a notable emphasis on addressing both dimensionality and aspects associated with their interrelation.

In [2] this study, the researchers introduce ideas for a system designed to identify credit card fraud in real-time, with a strong focus on maximizing its accuracy by capturing as many new transactions as possible. Furthermore, in their research, the authors conduct a comparative analysis of various techniques for detecting credit card fraud, such as LOF and SVM methods. The utilization of Isolation Forest (IF) in detecting fraudulent credit card transactions, as demonstrated through a series of experiments, has proven to be highly efficient in anomaly detection. It's noteworthy that the researchers explicitly state their intention to implement an architecture capable of real-time detection of fraudulent credit card transactions.

In [3] to a varying extent, in the given study, researchers employ techniques such as Isolation

Forest (IF) and Long Short-Term Memory (LSTM) to identify anomalies within the dataset.

Subsequently, they utilize neural networks to forecast and rectify irregular data points. In this research, scientists conducted experiments using datasets derived from electromagnetic systems, with the primary objective of refining these datasets for future operational and maintenance purposes. This preparatory phase enables subsequent experiments utilizing the previously mentioned IF and LSTM methods. The experimental outcomes, as presented by the researchers, indicate that the proprietary software for data cleansing related to power supply operations and maintenance, which is based on CiF-AL and incorporates the use of IF and LSTM techniques, has undergone significant optimization, both in terms of anomaly detection and the accuracy of their correlation.

Table 1: Comparison of Corrected RMSE Index for Prediction

| Algorithm | RSME |
|--|-------|
| Model based on time series analysis in | 0.779 |
| Traditional LSTM neural network | 0.292 |
| CiF-AL | 0.263 |

In [4] this study, the researchers introduce an alternative approach to enhancing the security of avionics systems, specifically the ADS-B system, a critical component of most aircraft responsible for ensuring air navigation safety. The primary focus of their work lies in proposing an innovative security solution designed to detect abnormal ADS-B messages. This solution primarily aims to identify falsified or tampered ADS-B data sent either by malicious third parties (attackers) or compromised aircraft. The researchers present an algorithm rooted in Long Short-Term Memory (LSTM) to model an aircraft's flight path by analyzing sequences of crucial data and messages within the ADS-B system. Based on this analysis, the aircraft assesses and evaluates ADS-B messages to spot deviations from the correct flight path. To validate their approach, the researchers conduct a series of experiments using thirteen diverse datasets of flight paths, each containing various types of anomalies. The final result of

their research entails a performance comparison between the LSTM algorithm and other frequently employed methods for identifying unusual data patterns within datasets, such as GMM-HMM, DBSTREAM, SVM, LOF, and IF. The research results outlined in the study demonstrate that the proposed approach successfully identifies all injection attacks, with a minor increase in the average false-positive rate by 4.5%. This outcome provides evidence that the effectiveness of the presented algorithm surpasses that of the other methods under examination.

In [5] this work, the authors delve into the application of the Isolation Forest (IF) method, particularly focusing on the critical aspects of selecting the appropriate attribute and determining the optimal split point within that attribute when constructing an IF model. To address these challenges, the authors propose an innovative computational structure, which, in their view, excels at identifying the most

distinguishable attributes and pinpointing the optimal division points for them. The researchers highlight a key limitation of the IF method, namely, its reliance on a binary tree structure, which comes with constraints related to tree depth, leaf nodes, or external nodes. These aspects are pivotal in assessing whether a given instance is normal or anomalous. In this work, the authors present a novel solution aimed at streamlining the process of building a binary tree within the framework of the IF method.

Algorithm GradFindSpli

Input:

X^k : sorted values of the k th attribute;

Output:

bestX: the attribute value having the largest value of the separability index;

bestStep: the largest values of the separability index;

1. Initiate step as $\text{ceil}(X^k * 0.1)$
2. Let bestStep be step(X^k, X_1) and bestX be $(x_1 + x_2)/2$ and $i=0$
3. *while* $i < |X^k|$ *do*
4. Set $i = i + \text{step}$
5. Set currStep = step(X^k, x_i)
6. *If* currStep > bestStep *then*
7. Set bestStep = currStep
8. Set bestX = $(x_i + x_{i+1})/2$
9. *end if*
10. Update step by following the formulas (6) and (7)

Table 2: Results of Anomaly Detection on Actual Data. TPR Refers to True Positive Ratio, and Fpr Refers to False Positive Ratio

| Training Data | Time Range | Anomaly Ratio | TPR | FPR |
|---------------|-------------------|---------------|--------|-------|
| T1 | 20171118-20171127 | 4.4% | 97.61% | 1.48 |
| T2 | 20171112-20171121 | 1.0% | 94.42% | 0.22% |
| T3 | 20171108-20171117 | 0.0% | 98.21% | 0.66% |

In [7] the researchers primary focus lay in the identification of anomalies within network traffic. While there exist numerous studies in this domain, the novelty of their work lies in the parallel algorithm they introduce, which combines the Isolation Forest (IF) method with Spark. This amalgamation greatly expands the scope of anomaly detection within computer networks, notably by enabling the processing of vast volumes of data. During the computational

11. *end while*
12. *Return* best X and bestStep

In [6] the study discusses the authors' adaptation of the Isolation Forest (IF) method for the detection of network anomalies originating from the network management system. The authors emphasize the limited availability of techniques for identifying network anomalies and propose modifications to the IF method to address this issue. The adapted IF method has been tailored to eliminate recurring patterns, detect peaks and troughs, and fine-tune various metrics, thus increasing its applicability. By incorporating these IF algorithms, the research methodology presented enables the rapid identification of anomalies within extensive datasets containing multidimensional network management data.

Additionally, the paper presents experimental results based on six simulated datasets, showcasing the adaptability of the proposed method to various types of data derived from network performance measurements. Moreover, the authors conduct a comparative assessment of the suitability of five distinct feature extractors, evaluating them based on their characteristics and Area Under the Curve (AUC) scores. Finally, the researchers demonstrate the real-world application of the method in a computer network scenario, utilizing network traffic data for evaluation.

procedures, the authors distribute tasks across multiple IF and Spark computing nodes, thereby harnessing the potential of handling Big Data. The experimental findings, derived from real-world datasets sourced from network traffic, demonstrate the algorithm's efficiency and its effectiveness in detecting anomalies within network traffic data.

Table 3: AUC and Accuracy of iForest and SPIF

| Algorithm | AUC | Accuracy |
|-----------|--------|----------|
| iForest | 0.8831 | 8.872 |
| SPIF | 0.8927 | 87.144 |

In [8] their research, the scientists grapple with the challenge of applying the Isolation Forest (IF) method to multidimensional data, where it exhibits instability despite its low time complexity and effective anomaly detection capabilities. They stress that this method is vulnerable to noise in the data, prompting them to develop an enhanced anomaly detection approach rooted in the IF method, leveraging dimensional entropy. In this new method they present, the researchers incorporate dimensional entropy as a crucial component in both the selection of the isolation attribute and the determination of the isolation point. Within this method, the researchers outline a process for choosing isolation attributes and points based on the information entropy associated with each attribute within random samples. The outcomes of their research experiments demonstrate that this presented approach operates significantly faster and offers.

Algorithm 2: EiTTree($D_i, e, l, \text{bin}, \text{pathL}, \alpha$)

Inputs: $D_i, e, l, \text{bin}, \text{pathL}, \alpha$

Output: An EiTTree

1. If $e \geq l$ or $|D_i| \leq 1$ then
2. return exNode
3. Else
4. Calculate $\{\text{ent}_1, \text{ent}_2, \dots, \text{ent}_m\}$
5. End if;
6. If $0 < \text{ent}_1 < \alpha * \text{ent}_{\max}$ then record the attribute i into $\{\text{ch}_1, \text{ch}_2, \dots, \text{ch}_{\text{chnum}}\}$;
7. If $\{\text{ch}_1, \text{ch}_2, \dots, \text{ch}_{\text{chnum}}\}$ is not empty then
8. Isolation strategy 2 $\rightarrow D_{\text{left}}$ and D_{right} :
9. return inNode {
 $\text{pathL} = \text{pathL} * \text{pathL}_i$
 $\text{left} \rightarrow \text{EiTTree}(D_{\text{left}}):$
 $\text{right} \rightarrow \text{EiTTree}(D_{\text{right}}):$
} 10. end if.

In [9] the researchers address the limitations of current mobile device user authentication methods, particularly in terms of achieving the necessary security levels. They propose the utilization of an Isolation Forest (IF)-based

authentication model as a prototype for enhancing security. Their research centers on behavioral biometrics, specifically continuous user authentication based on the user's interaction with the device's touchscreen. Notably, the presented system's effectiveness in security operations is noteworthy. In the case of continuous authentication, the system was capable of detecting an unauthorized user after an average of seven actions. This achievement holds great unauthorized access.

In [10] to a certain extent, the researchers in this article highlight the connection between variations in vital signs as an indicator of potential health issues requiring attention. Consequently, they suggest that anomalies can be characterized by deviations from typical vital sign patterns, which may signal the necessity for immediate intervention. In their research, the authors introduce a hybrid method combining K-Means clustering and the Isolation Forest (IF) for detecting anomalies. They compare the outcomes of their research with those obtained from the standard IF algorithm and a hybrid technique that combines K-Means clustering with IF. The research findings presented by the authors indicate that the effectiveness of the proposed hybrid technique varies depending on the dataset, yielding either lower or higher error rates.

In [11] their study, the researchers introduce an Isolation Forest (IF) approach for low-fidelity event detection and monitoring, using real-time data from phasor measurement units (PMUs). Their method is capable of identifying early events even within data of subpar quality, and it boasts significantly reduced computational complexity compared to traditional machine learning techniques. Moreover, the suggested method exhibits enhanced reliability when compared to alternative techniques used for detecting and monitoring low-quality events, such as Principal Component Analysis (PCA) or Singular Value Decomposition (SVD).

In [12] the researchers focus their investigation on the Isolation Forest (IF) method, specifically delving into its IForestASD implementation. They employ this method within the scikit-multiflow framework, an open-source machine learning platform designed for data stream handling. The scientists then proceed to conduct experiments using three real datasets to evaluate the predictive performance and resource consumption of the IForestASD method. Additionally, they compare its performance with another data stream anomaly detection method referred to as half-space trees. Based on their research findings, the authors conclude that while IF demonstrates accuracy in traditional batch settings, it may not be well-suited for efficient data stream processing.

In [13] the researchers present an original approach to anomaly detection by incorporating the Isolation Forest (IF) method into streaming data analysis using a sliding window mechanism. They have termed this approach the IForestASD algorithm and have implemented it within Scikit-multiflow, a versatile machine learning platform tailored for handling data with multiple outputs and labels. This innovative solution holds promise for effectively detecting anomalies within streaming platforms, thus offering a substantial contribution to the field of anomaly detection in both industrial and research settings.

In [14] this study, the authors introduce a technique for identifying damage in hydroelectric generators using the Isolation Forest (IF) method. However, they highlight a significant challenge: failures in hydroelectric generators are infrequent, making it difficult to gather failure data. Therefore, the authors stress the importance of automatically constructing models exclusively with normal data and subsequently performing error detection. The researchers evaluate the effectiveness of IF for feature data by comparing it with an error detection method based on multivariate statistical process control (MSPC).

The experimental results validate that the damage detection approach proposed by the researchers, utilizing IF, outperforms the MSPC-based method in terms of accuracy. It's noteworthy that the researchers assess the correctness and

functionality of their approach using the Wilcoxon signed-rank test.

In [15] this presented research, the authors introduce a novel method termed "Mondrian Forest Isolation" or iMondrian Forest. The primary objective of this method is to detect anomalies both in batch processing and real-time streaming data. The researchers place particular emphasis on combining two established techniques: Isolation Forest (IF) and Mondrian Forest. They clarify that these methods are already recognized for their efficacy in batch anomaly detection and online random forest applications. The outcome of the researchers' efforts is a fresh data representation format capable of accommodating online data streaming while simultaneously serving as an anomaly detection tool. The conducted experiments reveal that iMondrian Forest outperforms the IF method significantly in batch processing scenarios and either surpasses or matches the performance of other anomaly detection methods designed for both batch and real-time processing. The method proposed by the researchers represents an innovative hybridization of Isolation Forests and Mondrian Forests, harnessing the strengths of both techniques. This combination leverages the isolation-based anomaly detection features and the advantages of online ensemble learning, as underscored by the research findings.

In [16] the researchers address the specific challenge of analyzing datasets that encompass various data types, including temporal, spatial, image, and categorical data. The presence of multiple data types often leads to the emergence of numerous outliers. To address this challenge, the authors introduce modifications to the Isolation Forest (IF) method, which they refer to as the Fuzzy Set-Based Isolation Forest (FSBIF).

The modification introduced by the researchers is based on the concept of measuring the distance from the center of the analyzed cluster during the construction of the binary search tree. Test results indicate that when utilizing the FSBIF approach, all of these data points achieve the highest degree of isolation compared to the conventional IF method, resulting in reduced degrees of isolation.

The research findings presented by the authors suggest that the method proposed in their work outperforms other existing methods in terms of both performance and effectiveness.

Table 4: Execution times of if and FSBIF

| Method | Artificial dataset | Taxi Database |
|--------------------|--------------------|---------------|
| IF | 11,023 pp | 55,997 pp |
| Fuzzy set-based IF | 1,152 pp | 7,298 pp |

In [17] the authors emphasize the challenges and expenses associated with obtaining damaged industrial samples, which are relatively scarce in practical applications. The researchers address the issue of dealing with imbalanced sample datasets, a common concern in industrial settings.

To tackle this problem, they employ a fusion anomaly detection method that combines Artificial Neural Networks (ANN) with the Isolation Forest (iForest) algorithm. The researchers' primary focus is on expanding the dataset through a technique known as upsampling, which involves increasing the

number of samples to maintain a balanced representation of different sample types. In the subsequent step, the researchers utilize Empirical Mode Decomposition (EMD) to extract specific energy features from each sample, aiming to capture the inherent signal characteristics more effectively. In the final stage of their research, the authors extract latent features, which are then input into the Isolation Forest algorithm to detect anomalies within the industrial samples. This comprehensive approach enhances the capacity to identify unusual patterns and anomalies in the data, particularly in scenarios where damaged samples are limited in availability.

Table 5: Details of the Experimental Method and Accuracy

| Method | Accuracy (%) |
|------------------|--------------|
| SVM | 83.86 |
| Light GBM | 93.54 |
| One ClassSVM | 80.06 |
| LOF | 54.56 |
| Isolation Forest | 92.35 |
| ANN | 95.6 |
| Proposed method | 97.46 |

The researchers conducted tests on the presented method along with various traditional methods using CWRU bearing datasets. The results demonstrated by the researchers indicate that the method they introduce outperforms the traditional approaches by a significant margin in terms of performance.

In [18] the authors highlight the urgent need for early prediction of diabetes, given that diabetes is currently one of the leading causes of global mortality. Late detection and untreated diabetes can result in severe health complications. To address this challenge, the researchers introduce a prognostic model designed to enable the early prediction of type 2 diabetes in individuals. In their model, the researchers employ the IFO (Isolation Forest-based outlier detection) method

and the synthetic minority pooling technique with sequential sampling. These methods are utilized to identify and remove outlier data points while also balancing the data distribution. For learning and predicting type 2 diabetes at an early stage, the researchers utilize ensemble classifiers. To assess the model's reliability, the authors utilize three publicly available datasets and compare the performance of their proposed model with other established models. They employ a 10-fold cross-validation method and evaluate four performance metrics: precision, recall, F1-score, and accuracy.

The research findings indicate that their model significantly outperformed other models, achieving accuracy rates of 93.18%, 98.87%, and 96.09% for datasets I, II, and III, respectively. This suggests the potential effectiveness of their model in early diabetes prediction.

In [19] the researchers propose a multimodal approach for detecting anomalies in embedded systems, especially in situations where there exists a correlation between energy consumption over time and memory access. They utilize time series data that encompass diverse power consumption patterns, L2 cache availability, and memory bus access patterns to train one-class Support Vector Machine (SVM) and Isolation Forest (IF) classifiers. This approach harnesses side channels with the capability to identify anomalies. The authors conducted a series of experiments employing a high-fidelity processor emulator. The experimental results demonstrate that the SVM-based detector, using the power consumption channel, achieved a high level of accuracy (ranging from 95% to 100%, depending on the type of attack). In subsequent attacks where code modifications altered the code's computational workload (e.g., adding more numerical calculations), the IF anomaly detector, which utilized the detected memory side-channel attacks, also achieved high accuracy (ranging from 95% to 100%). In these cases, code modifications led to changes in memory access patterns or the number of memory accesses. The researchers emphasize that the combined utilization of both side channels, in which the presented detectors were employed, leverages their respective strengths and proves highly effective in accurately detecting the considered attacks in their experiments.

In [20] this study, the authors underscore the significance of advancements in technologies for treating type 1 diabetes (T1D), with a particular focus on the field known as the "artificial pancreas" (AP). They highlight the use of mobile and transportable devices equipped with a closed-loop system, designed to administer insulin based on real-time data from a glucose sensor. In this presented work, the researchers shift their attention to the challenge of automatically detecting faults in these insulin pumps using multidimensional data-driven anomaly detection (AD) methods. They emphasize the use of unsupervised methods, which eliminate the need for labeled training data, often difficult to obtain in practical applications. The research

outcomes are highly promising, primarily due to the well-designed approach they propose. This approach paves the way for the utilization of these methodologies in AP systems and their potential integration into remote monitoring systems, offering significant benefits in the field of T1D treatment.

In [21] this study, the authors bring attention to the issue of traffic fluctuations in mobile networks, which can lead to inaccuracies in network management and, subsequently, a decrease in service quality. To address this problem, the researchers present an online data mining technique for detecting anomalous data within network operations. Since this data forms the foundation for network operator decisions, including network monitoring and real-world phenomena control, ensuring its accuracy is crucial for minimizing potential harm. The experiments conducted in this study are based on actual traces of cellular communication, employing an automated platform known as STAD. This platform aims to identify spatiotemporal anomalies using a combination of machine learning techniques, including one-class Support Vector Machine (OCSVM), support vector regression (SVR), and recurrent neural networks such as Long Short-Term Memory (LSTM). The researchers validate their findings against a real dataset from Call Detail Records (CDR). The results presented in the study demonstrate significantly higher accuracy compared to methods such as Isolation Forest (IF) and Auto-Regressive Integrated Moving Average (ARIMA). This suggests that the approach proposed by the researchers offers a more effective solution for anomaly detection in mobile network traffic data.

In [22] this study, the authors focus on the detection of very low-frequency electromagnetic anomalies associated with earthquakes. They introduce a multifunctional method for identifying anomalies in AETA ULF (Ultra Low-Frequency) electromagnetic disturbance signals. This method is rooted in the Isolation Forest algorithm and includes an additional step for feature extraction and selection. To assess the effectiveness of their approach, the researchers

employ epoch analysis, abbreviated as SEA. This comprehensive method allows them to evaluate

and detect anomalies in ULF electromagnetic disturbance signals related to seismic activity .

Table 6: Results K Comparison of two Methods

| Methods | Station Abbreviation | | | | | |
|----------|----------------------|-------|-------|------|-------|-------|
| | GOX | MB | MX | EMS | ZT | MSQS |
| IQR | 6.33 | 25.83 | 8.0 | 4.23 | 13.42 | 52.18 |
| i-Forest | 52.56 | 37.76 | 17.01 | 28.2 | 28.97 | 29.67 |

The research findings presented by the authors reveal that half of the chosen stations exhibit a correlation between signal anomalies and earthquakes. Moreover, their method demonstrates significantly superior performance compared to the traditional single-function IQR (Interquartile Range) method. The test results suggest that the approach introduced by the researchers holds the potential to be a much more effective choice in the future for identifying global anomaly points, particularly in the context of earthquake-related signal anomalies.

In [23] this study, the authors address the pressing issue of bacterial resistance to antibiotics, recognized as a global health crisis. Their research objective is to leverage machine learning to predict the minimum inhibitory concentration (MIC) for bacterial isolates based on the presence of resistance genes in their DNA. To achieve this, the authors employ a dataset that includes information about the presence of beta-lactamase genes, which were identified using whole genome sequencing. They also examine the correlation between the presence of these genes and the bacteria's sensitivity to beta-lactam antibiotics. The test results indicate that the proposed K-Nearest Neighbors method yields predictions of MIC with a similar level of accuracy as the Random Forest method, achieving predictions within a range of ± 12 -fold dilution. However, when it comes to predicting MIC, Random Forests achieve better F1-micro scores, demonstrating their effectiveness in this context.

In [24] the researchers in this study focus on employing the Isolation Forest (IF) method for detecting unusual behavior within business processes, encompassing errors and missing values. This led them to develop a method that integrates the degree of consistency within event

log functions from multiple perspectives, utilizing IF as a tool for anomaly detection.

In their study, the researchers introduce an algorithm implemented within the Scikit-learn Framework. They assess the algorithm's performance in terms of its ability to quickly recognize anomalous behavior and the quality of its model accuracy. The research findings suggest that the algorithm developed by the researchers is effective in detecting abnormal behavior in this specific type of data and notably enhances the quality of the Isolation Forest (IF) model.

In [25] this research, the authors highlight that the Isolation Forest (IF) method is well-suited for precise data but encounters challenges when dealing with imprecise, incomplete, ambiguous, or uncertain data. In response, they introduce a novel approach known as the Fuzzy Isolation Forest. This method is designed to address the inherent "ambiguity" present in such data. In their work, the researchers utilize an approach based on alpha-cuts, wherein data instances are assigned to tree nodes, and their membership values are compared with an alpha-cut threshold. Through a series of experiments, the authors define this new solution and demonstrate that their method can operate with nearly the same accuracy as when working with clean, well-defined data. Importantly, their approach proves to be as effective as the IF method in detecting anomalies within fuzzy data without significantly increasing processing time.

In [26] this work, the authors highlight an enhancement to the Isolation Forest (IF) algorithm through the use of the Finite Boundary (FB) algorithm. This improvement is particularly advantageous when working with datasets composed of variables that conform to a standard distribution. It provides an alternative to existing

algorithms like IF and Extended Isolation Forest (EIF) in such scenarios.

Table 7: AUC Values for ROC for Credit Card Dataset

| Features | IF | EIF | FBIF |
|---|-------|-------|-------|
| (V4, V10, V11, V12, V13, V14, V15) | 0.875 | 0.896 | 0.925 |
| (V4, V10, V11, V12, V13, V14, V15 CATEGORY AMOUNT) | 0.862 | 0.891 | 0.911 |
| (V4, V10, V11, V12, V13, V14, V15 CATEGORY AMOINT, CATEGORY TIME) | 0.833 | 0.879 | 0.908 |
| (ALL NUMERICAL WITHOUT AMOUNT, TIME) | 0.868 | 0.895 | 0.905 |
| (ALL NUMERICAL, CATEGORY AMOUNT, CATEGORY TIME) | 0.863 | 0.886 | 0.907 |

In their study, the authors introduce the FBIF algorithm, which leverages hyperspheres to form a branching boundary aimed at addressing data inconsistencies. This approach is especially effective when dealing with input features that exhibit properties akin to a standard distribution.

The researchers connect these hyperspheres to create a generalized branching decision boundary and assess the data for compliance with a standard distribution of input features.

The results obtained from a series of experiments indicate that the method proposed by the researchers significantly outperforms other methods such as FBIF with EIF and IF. This superior performance holds true even when the input data for the IF and EIF methods consists of a larger number of input features. In summary, their approach proves to be highly effective for anomaly detection, particularly when dealing with data characterized by standard distribution properties.

Table 8: AUC Values for PRC for Credit Card Dataset

| Features | IF | EIF | FBIF |
|---|-------|-------|-------|
| (V4, V10, V11, V12, V13, V14, V15) | 0.862 | 0.889 | 0.931 |
| (V4, V10, V11, V12, V13, V14, V15 CATEGORY AMOUNT) | 0.867 | 0.891 | 0.920 |
| (V4, V10, V11, V12, V13, V14, V15 CATEGORY AMOINT, CATEGORY TIME) | 0.843 | 0.882 | 0.922 |
| (ALL NUMERICAL WITHOUT AMOUNT, TIME) | 0.874 | 0.897 | 0.922 |
| (ALL NUMERICAL, CATEGORY AMOUNT, CATEGORY TIME) | 0.871 | 0.892 | 0.921 |

In [27] this study, the researchers aim to predict the accuracy of detecting credit card fraud, specifically through the analysis of sequential transaction data. They compare two methods: Light Gradient Booster and Isolation Forest (IF).

accuracy rate of 81.8%. This suggests that the Light Gradient Booster method is more effective in detecting credit card fraud in the context of sequential transaction data.

In this work, the researchers employ an algorithm, specifically Light Gradient Booster, with a sample size of 10, and they also use the Isolation Forest (IF) method with the same sample size of 10. Their objective is to provide a highly detailed assessment of the accuracy percentage in detecting fraudulent credit card transactions. The experimental results presented in the study indicate that the Light Gradient Booster algorithm achieves significantly higher accuracy, with a rate of 91.6%, compared to the Isolation Forest method, which achieves an

Table 8: The Accuracy of Predicting Fraudulent Credit Card Transactions Varies Based on Different Sample Sizes, with the Light Gradient Booster Algorithm Achieving an Accuracy of 91.6% and the Isolation Forest Method Achieving an Accuracy of 81.8%. the Choice of Sample Size Can Have a Significant Impact on the Performance of These Algorithms in Detecting Fraud

| S. No. | Random_State | Light GMB | Isolation Forest |
|--------|--------------|-----------|------------------|
| 1 | 100 | 94.50 | 83.50 |
| 2 | 200 | 95.00 | 82.60 |
| 3 | 300 | 92.00 | 82.69 |
| 4 | 400 | 93.00 | 80.20 |
| 5 | 500 | 92.61 | 79.56 |
| 6 | 600 | 89.25 | 82.090 |
| 7 | 700 | 91.60 | 81.80 |
| 8 | 800 | 90.11 | 81.90 |
| 9 | 900 | 88.11 | 83.00 |
| 10 | 100 | 90.26 | 79.50 |

The results of the conducted research indicate a statistically significant difference between the Light Gradient Booster (LGB) algorithm and the Isolation Forest (IF) algorithm, with a p-value of 0.0001 ($p < 0.05$) based on a two-sided analysis. This statistical significance suggests that there is a meaningful difference between the two algorithms.

Furthermore, the research findings confirm that the presented LGB algorithm outperforms the IF method, particularly in terms of accuracy in detecting fraudulent credit card transactions. Specifically, the LGB algorithm demonstrates a significantly higher accuracy rate for eight out of ten credit card transactions when compared to the IF method.

In [28] the researchers highlight that anomaly detection plays a crucial role in enabling automatic monitoring and identification of abnormal and unusual events. They emphasize that quick actions are essential for preventing potential failures. In this context, they reevaluate the classic Isolation Forest (IF) method, particularly when dealing with datasets that contain an exceptionally large number of data points. The researchers seek an alternative approach to calculate the division value, aiming to avoid dividing areas with very high data point density. This approach is designed to enhance the isolation capability of the method and improve its effectiveness in detecting anomalies within large datasets.

Table 9: Auc-Pr Results. the Second Column Shows the Value Obtained Using the Classic Isolation Forest. the Third Column Shows the Results Achieved Using the Proposed Model. Bold Texts Provide Leading Results

| Dataset | Isolation Forest | Proposed Method |
|-------------|------------------|-----------------|
| Anthyroid | 0.3042 0.323 | 0.3227 0.0195 |
| Breastw | 0.9707 0.0043 | 0.9715 0.0032 |
| Glass | 0.1133 0.0106 | 0.0931 0.0112 |
| Ionosphere | 0.8095 0.072 | 0.8021 0.0085 |
| Mammography | 0.2211 0.0357 | 0.2286 0.0476 |
| Pendigits | 0.02631 0.0639 | 0.2843 0.0574 |
| Pima | 0.5005 0.0089 | 0.5024 0.0135 |
| Satellite | 0.6583 0.0237 | 0.6651 0.0221 |
| Thyroid | 0.5257 0.0908 | 0.5461 0.0380 |

As previously emphasized, the researchers introduced a modification in the method by changing the approach to selecting the division value. This adjustment was aimed at avoiding

division in areas with high data point density. The research findings underscore that the proposed method consistently outperforms the results achieved using the traditional Isolation Forest (IF) method. Furthermore, it is noteworthy that these improved results are particularly significant when dealing with datasets that contain a large number of instances. Consequently, the comparative analysis between the standard IF method and the novel approach introduced by the researchers consistently indicates that the new method is superior, especially when applied to large datasets.

In [29] the researchers aimed to explore anomaly detection algorithms, with a specific focus on the Isolation Forest algorithm and the LOF algorithm, both renowned in the field of anomaly detection. They highlighted a challenge associated with these algorithms: determining which one is better suited for handling extensive datasets. Their study commenced by delving into the fundamental principles and operational mechanisms of these two algorithms. Subsequently, the researchers conducted practical experiments using a dataset to compare and assess the accuracy and stability of both algorithms in identifying data anomalies. As a result of their investigative work and the insights garnered from their experiments, the researchers concluded that the Isolation Forest algorithm exhibited greater effectiveness in detecting anomalies within data mining tasks. To address some of the limitations inherent to the Isolation Forest algorithm, they proposed several enhancements: Parameter Sensitivity Adjustment: The researchers suggested fine-tuning the algorithm's parameters to enhance its adaptability when processing datasets of varying sizes. This adjustment was aimed at bolstering overall performance and resilience. Drawing from Random Forest: Taking inspiration from the bagging-based Random Forest algorithm, the researchers sought to incorporate similar techniques to bolster the Isolation Forest's capability to effectively handle high-dimensional data. Integration of Simulated Annealing: The researchers recommended integrating a simulated annealing algorithm into the Isolation Forest. This addition would enable the algorithm to

selectively filter out trees with superior detection accuracy, ultimately elevating the algorithm's overall accuracy in anomaly detection. In summary, the researchers' objective was to not only compare the Isolation Forest and LOF algorithms for anomaly detection but also to propose valuable enhancements to the Isolation Forest algorithm. These refinements aimed to mitigate its limitations, rendering it more versatile and precise, especially when dealing with extensive and intricate datasets.

In [30] the researchers address the challenges associated with GPS data, which can be easily influenced by environmental factors, resulting in uneven data quality and impacting its utility. Their study focuses on utilizing the Isolation Forest (IF) technique to identify anomalies within GPS data. The IF technique operates by constructing a set of trees, initially using subsamples from the training dataset. Additional samples are then integrated into the isolated trees to calculate an anomaly score for each sampled point. For their experiments, the researchers employed real GPS data collected from sanitation vehicles and employed the Local Outlier Factor (LOF) technique for cross-validation. Detecting anomalies in GPS data poses several inherent challenges. Firstly, the likelihood of GPS drift is low, and historical data from moving targets provides limited actual abnormal data samples. Consequently, obtaining complete characteristics of all abnormal data is often unfeasible.

Additionally, GPS data frequently lacks explicit "normal" or "abnormal" labels, making the conversion of extensive GPS data into accurately labeled datasets challenging, particularly when manual labeling is required. To address these challenges, the researchers propose an unsupervised approach using the Isolation Forest method. This method enables the detection of anomalies within GPS datasets that are entirely unlabeled and contain a significantly larger proportion of normal data compared to abnormal data.

In [31] the researchers highlight that existing anomaly detection methods primarily rely on one-dimensional data for identification and

analysis. However, in their study, they introduce an algorithm designed for detecting anomalies within multidimensional data in an isolated forest framework. To achieve this, they employ multidimensional plane segmentation during the hyperplane segmentation construction, optimizing the process of generating an isolated forest.

The algorithm's effectiveness is demonstrated through comparative experiments involving several typical datasets, where it delivers promising results. The presented algorithm significantly enhances the analysis of multidimensional data by incorporating multidimensional plane segmentation, which improves segmentation efficiency and simplifies the identification

of anomalies. Furthermore, within the isolated tree generation process, the researchers implement a method for controlling the lengths of the left and right subtrees. This halts growth once specific proportion conditions are met, leading to the selection of an iTree, thereby enhancing the efficiency of the isolated forest algorithm for anomaly detection. In summary, the researchers approach, which centers on improving the construction of isolated forests and the utilization of multidimensional plane segmentation in hyperplane segmentation construction, leads to more efficient segmentation and facilitates the identification of anomalies within multidimensional datasets.

Table 10: Influence of the Division Value on the F1 Result

| Dataset | F1 score | | |
|------------|----------|-----------|---------------|
| | I Forest | E-iForest | M. D. iForest |
| Cancer | 0.5657 | 0.5672 | 0.5855 |
| Carida | 0.4126 | 0.4122 | 0.4174 |
| Wbc | 0.3413 | 0.3354 | 0.3972 |
| Prime | 0.5399 | 0.5388 | 0.5412 |
| Satimage-2 | 0.0924 | 0.0862 | 0.0973 |

The research findings demonstrate that the algorithm proposed by the researchers outperforms the original iForest algorithm by achieving significantly higher values. As the next step, the researchers plan to focus on enhancing the algorithm's computational efficiency.

In [32] the researchers in this study focus on addressing the challenges posed by the vast amount of data generated by wind farms. They introduce an approach using an autoencoder combined with the Isolation Forest (IF) method to tackle the complexities and incompleteness often encountered when processing wind energy data.

Their methodology involves segmenting wind speed data into bins and creating a distribution of

anomalous data, followed by the identification and removal of outliers using the autoencoder-based Isolation Forest. Subsequently, the data is reconstructed and cleaned. To validate the effectiveness of their approach, the researchers conducted experiments using real wind turbine data, utilizing the cleaned data for forecasting and comparing the results with actual forecasts. In summary, their study addresses outlier removal in wind speed-power curves, and the introduced forest cleaning algorithm, featuring an autoencoder and isolation forest, effectively removes outliers by leveraging wind speed intervals and the characteristics of abnormal data distribution.

Table 11: Comparison of Prediction Performance before and after Data Processing

| Algorithms | Date status | R2 | MSE |
|-----------------|------------------|--------|-------------|
| LSTM | Unprocessed data | 0.7160 | 342664.4400 |
| | Processed data | 0.9850 | 15178.6981 |
| XGBOOST | Unprocessed data | 0.6847 | 380499.2805 |
| | Processed data | 0.9867 | 14177.0987 |
| LinerRegression | Unprocessed data | 0.6785 | 387913.4428 |

| | | | |
|--------------|------------------|--------|-------------|
| | Processed data | 0.9704 | 31597.9429 |
| Lasso | Unprocessed data | 0.6785 | 387899.3228 |
| | Processed data | 0.9704 | 31617.9951 |
| SVR | Unprocessed data | 0.6825 | 383181.6014 |
| | Processed data | 0.9831 | 17980.5811 |
| Ridge | Unprocessed data | 0.6785 | 387913.2304 |
| | Processed data | 0.9704 | 31598.2600 |
| RandomForest | Unprocessed data | 0.6377 | 438113.0878 |
| | Processed data | 0.9844 | 16592.2207 |

Through a series of tests conducted on the Yalova wind energy dataset in Turkey, the researchers validate the effectiveness of their proposed method. They achieve this by comparing the prediction results using the data before and after the cleaning process. This verification is crucial for subsequent tasks such as wind power forecasting and modeling wind speed-power curves. Additionally, it provides valuable technical support for integrating wind energy into the grid and mitigating some of the adverse effects associated with wind energy generation.

In [34] this research, the authors highlight the significance of addressing fraud and anomalies within the telecommunications sector, emphasizing the potential of machine learning-based tools for detection. To tackle this challenge, they leverage Call Detail Record (CDR) data, comprising a substantial 417,000 call records across 217 different countries.

Initially, the researchers analyze telephone traffic, employing K-means clustering to categorize multi-valued categorical variables effectively. In the subsequent phase, they explore various models, including XGBoost, Extra Trees, Random Forest, and the Isolation Forest, as well as a novel model known as the Mixture of Experts. This new model calculates the average prediction probabilities of supervised models, contributing to anomaly detection. In the final step of their methodology, the researchers aggregate the results by summing the predictions from the five models, ultimately labeling connections based on their anomaly scores. The research findings reveal that approximately 1% of calls raise suspicion of fraudulent activity, aligning with industry reports on this issue. In another aspect of their work, the authors turn their focus to Fog Computing Networks (FCNs), taking into account factors such

as user dispersion, real-time data, and user privacy. They introduce a federated learning framework that incorporates malicious models.

This framework comprises a two-layer blockchain, consisting of the primary blockchain and a directed acyclic graph chain, which enhances data security within the network. To ensure the security of global models, the authors present an isolation forest-based malicious model detection algorithm. This algorithm effectively filters out malicious local models, facilitating global aggregation using the Stochastic Gradient Descent algorithm. The presented research results, including simulation outcomes, demonstrate the effectiveness of the proposed algorithm in enhancing security and privacy in Fog Computing Networks.

In [34] the authors of this study focus on the context of fog computing networks (FCNs), taking into careful consideration factors like user distribution, real-time data processing, and user privacy. In their research, they introduce a federated learning framework that hinges on the concept of malicious models. This innovative framework incorporates a two-layer blockchain system, consisting of the primary blockchain and an additional directed acyclic graph chain, seamlessly integrated into the network structure. The primary purpose of this blockchain architecture is to fortify data security within the FCNs, addressing concerns related to user privacy and secure data transmission.

Within their work, the researchers introduce an isolation forest-based algorithm designed to detect malicious models effectively. This algorithm plays a pivotal role in filtering out potentially malicious local models within the FCNs. Furthermore, it enables the global

aggregation of data while maintaining security standards. This aggregation process is facilitated through the application of the Stochastic Gradient Descent algorithm, ensuring the integrity and security of the global model. The research findings, which encompass results from extensive simulations, convincingly demonstrate the effectiveness of the proposed algorithm. It offers a robust solution to enhance security and privacy within fog computing networks, contributing to the advancement of this technology in various application domains.

In [35] researchers are paying attention to identifying places for in situ research in robot and human exploration. In the presented work, researchers analyze the usefulness of machine learning algorithms for outlier detection in order to identify interesting samples in field exploration sites based on remotely sensed observations. For the presented analyses, the authors used two sites of the Cucomungo alluvial fan in Death Valley in California and the Jezero crater on Mars. The study results indicated that outliers identified by the isolation forest algorithm in the satellite

datasets appeared to correspond to unique surface compositions or properties. The results indicate that this is a method that can support scientists by guiding the selection of a field exploration site.

In [36] the researchers are concerned with ensuring the reliability of sensor data in Internet of Things (IoT) applications, especially in detecting unusual conditions or anomalies like intrusions. They highlight the potential of machine learning-based anomaly detection in achieving this. In their study, they focus on addressing the issue of sensor data consistency within the context of hotspot detection using sensor pairs in a commercial IoT system designed for monitoring grain in large horizontal silos. The researchers aimed to assess the performance of traditional machine learning algorithms for anomaly detection, including the Localization Coefficient, Isolation Forest, and Single Class Support Vector Machine, in this specific context.

The study's findings reveal that the deep learning model with long-term memory (LSTM) introduced by the researchers outperforms the conventional machine learning methods [36].

Table 12: Results of the Traditional and Proposed Models

| Model | Features | Accuracy | Precision | Recall | F1 Score |
|--------------|--------------------------------|----------------|---------------|---------------|----------------|
| iForrest | Differences | 85.2% (0.001) | 84.2% (0.002) | 38.9% (0.002) | 0.532% (0.001) |
| | 2 Sensors Readings | 77.77% (0.001) | 46.6% (0.002) | 21.5% (0.002) | 0.294 (0.001) |
| | 2 Sensors Reading & Difference | 85.2% (0.001) | 84.2% (0.002) | 38.9% (0.002) | 0.532 (0.001) |
| LOF | Differences | 53.9% (0.0008) | 61.6% (0.001) | 54.8% (0.003) | 0.850 (0.001) |
| | 2 Sensors Readings | 57.6% (0.0006) | 59.9% (0.001) | 80.6% (0.002) | 0.688 (0.001) |
| | 2 Sensors Reading & Difference | 57.8% (0.0001) | 57.9% (0.001) | 99.5% (0.001) | 0.732 (0.001) |
| OneClass SVM | Differences | 52.4% (0.002) | 100% (0.001) | 2.0% (0.001) | 0.404 (0.002) |
| | 2 Sensors Readings | 51.5% (0.005) | 54.9% (0.002) | 1.2% (0.001) | 0.231 (0.002) |
| | 2 Sensors Reading & Difference | 51.7% (0.005) | 64.1% (0.002) | 1.4% (0.001) | 0.266 (0.002) |
| LSTM | 2 Sensors Readings | 90% (0.003) | 88% (0.002) | 90% (0.003) | 0.89 (0.004) |

In [37] the authors of the study draw attention to a significant surge in the volume of data being transmitted across publicly accessible computer networks. This proliferation of data has been accompanied by an escalating demand for novel and robust methods to safeguard against cyber threats and intrusions. In response to this pressing need, the researchers have directed their focus towards the crucial task of anomaly detection within the domain of cybersecurity. To

address this challenge, the authors have undertaken a comprehensive analysis of various anomaly detection methodologies, including DBSCAN, One-class SVM, LSTM, and Isolation Forest. Their objective is to evaluate the efficacy of these methods in identifying and mitigating cybersecurity threats effectively. The experimental results presented in their study yield valuable insights. They suggest that the individual classification algorithms examined may not be

well-suited for routine cybersecurity operations, thereby underscoring the imperative to explore avenues for potential performance enhancements in this critical area.

In [38] this work, the authors highlight the crucial role of flight data recorders (FDR) and cockpit voice recorders (CVR) in aviation accidents investigations, emphasizing their mandatory usage in aircraft. However, the authors propose a novel approach termed I-FDR (Intelligence Flight Data Recorder) that aims to enhance the utility of FDRs by providing continuous data mining capabilities. This innovation seeks to empower flight crews with improved situational awareness through real-time data analysis. Similar to traditional FDRs, I-FDR records essential flight parameters, but its distinctive feature lies in its ability to process and analyze this data continuously. To assess the effectiveness of I-FDR in assisting flight crews during critical situations, the authors explore the applicability of three unsupervised machine learning algorithms: DBSCAN (Density-Based Spatial Clustering of Applications with Noise), Isolation Forest, and LSTM (Long Short-Term Memory). The research results presented in the study provide valuable insights into the potential of these algorithms to contribute to flight crew management in hazardous scenarios

In [39] the authors emphasize the significance of Software-Defined Networking (SDN) as a solution for enhancing network security. SDN offers a holistic network view via a logically centralized component known as an SDN controller, effectively separating the control plane from the data plane. This division provides increased network control and introduces new possibilities for addressing emerging security threats, such as zero-day attacks. In their research, the authors propose a comprehensive machine learning (ML) framework designed for SDN environments. This framework incorporates unsupervised ML techniques and features a scalable method for collecting and selecting relevant network data attributes to facilitate the rapid detection of security threats. The key components of this framework include a Data Flow Collector (DFC) responsible for efficiently gathering network data features using the sFlow protocol, an Information Gathering Feature (IGF) selection module that identifies the most informative features, thus reducing training and testing complexities, and an innovative unsupervised ML module utilizing a novel outlier detection technique known as Isolation Forest (ML-IF) to swiftly and efficiently detect network security threats within the SDN context.

Table 13: Performance Metrics of Our Proposed Framework and State-of-the-Art ML/DL Models Based on the Available Measures Using the UNSW-NB15 Dataset

| Methods | Accuracy | DR | Time (second) |
|---------|----------|------|---------------|
| SSL | 0.86 | 0.85 | ON |
| RF | 0.93 | 0.92 | ON |
| FeedWSN | 0.92 | 0.91 | ON |
| OGM | 0.95 | 0.94 | ON |
| MHMM | 0.96 | 0.95 | ON |
| ML-IF | 0.97 | 0.96 | 38.33 |

The experimental results conducted on the UNSW-NB15 public network security dataset clearly demonstrate the superiority of the framework introduced in this research. This framework exhibits substantial enhancements in terms of both detection accuracy and processing speed when compared to contemporary solutions.

Furthermore, it achieves these improvements while effectively reducing computational

complexity, making it a highly promising and efficient approach to network security threat detection [39].

In [40] the authors of this study focus on chronic kidney disease (CKD), a complex condition that impacts kidney functions and structures, affecting millions of individuals worldwide and ranking among the leading causes of illness and death. In their research, the authors introduce an

innovative diagnostic approach for CKD, leveraging a 1D convolutional neural network (1D CNN) to address the limitations of existing methods while notably enhancing diagnostic accuracy. The outcomes of this study reveal that the model developed by the researchers achieved an impressive accuracy rate of 99.21%, surpassing current state-of-the-art techniques.

In [41] the authors of this study emphasize the significance of the industrial Internet of Things (IoT) and its successful integration with machine learning techniques in recent years. They address a notable limitation in many machine learning approaches, which is their ability to make predictions without providing explanations. This lack of interpretability poses a challenge for decision-makers who may struggle to comprehend and trust these predictions, potentially hindering the adoption of machine learning tools for practical use. To overcome this limitation, the authors propose an unsupervised anomaly detection system capable of not only identifying anomalies but also explaining the predictions, facilitating root cause analysis. They achieve this by combining the Isolation Forest method with a fast, model-independent interpretation technique known as Accelerated-agnostic Explanations (AcME). Their research findings highlight the effectiveness of AcME in providing explanations for predictions. The authors adapt AcME for anomaly detection, replacing the interpretation of Anomaly Detection with explanations for the anomaly results predicted by the Anomaly Detection algorithm, including the Isolation Forest. This approach proves to be both efficient and suitable for integration with decision support systems. In contrast, SHAP (Shapley additive explanations) is deemed impractical due to its high computational complexity.

In summary, the authors' work offers a promising solution for enhancing the interpretability of machine learning-based anomaly detection systems, particularly in the context of the industrial IoT, where actionable insights from predictions are crucial for informed decision-making.

In [42] their work, the authors introduce SPiForest as an innovative approach to outlier detection. They address a common issue faced by existing methods, where performance diminishes when dealing with outliers that appear similar to normal data in specific subspaces. To tackle this challenge, they propose SPiForest, which leverages principal component analysis (PCA) to identify principal components and estimate the probability density function (pdf) of each component. Additionally, they use inv-pdf, the inverse of the estimated pdf based on the dataset, to create support points across all attributes.

These support points define hyperplanes used to isolate outliers effectively. SPiForest offers two key advantages: it isolates outliers with fewer hyperplanes, leading to improved accuracy, and it excels at detecting outliers that may blend into subspaces due to their "few and different" nature. Comprehensive testing and comparative analyses demonstrate that SPiForest significantly enhances the area under the curve (AUC) compared to state-of-the-art methods. Notably, it improves AUC by up to 17.7% when compared to iF-based algorithms.

In [43] the authors highlight the significance of phasor measurement units (PMUs) in monitoring the state of smart grid systems. In their research, they explore various unsupervised learning techniques for detecting false data injection (FDI) attacks by identifying outliers. Initially, the study reveals that the moving average (MA) and density-based spatial clustering of applications with noise (DBSCAN) methods can accurately identify over 90% of the injected data points, while the isolation forest method detects almost 60% of these points. However, the research findings indicate that the implementation of the Amelia II imputation method, which imputes current data for multiple missing durations, yields impressive results. Specifically, the mean absolute percentage error (MAPE) for Amelia II is just 0.67, surpassing the performance of standard imputation methods.

Table 14: Results of DBSCAN in Identifying Injected Data

| Data set | % Total Inj. Pts Identified/Number of Inj. pts | Pts number. Detected Outside Injection Interval / Total Number of Data Pts |
|-------------|--|--|
| W1, 6 Sec. | 94.3% / 300 | 31/4080 |
| W2, 6 Sec. | 96.6% / 300 | 65/4080 |
| W1, 10 Sec. | 98.4% / 500 | 61/4080 |
| W2, 10 Sec. | 93.8% / 500 | 65/4080 |
| W1, 21 Sec. | 88.9% / 1050 | 22/4080 |
| W2, 21 Sec. | 91.4% / 1050 | 37/4080 |

In [44] the authors of this study address pressing contemporary health concerns, highlighting that currently, one in every four individuals aged over 25 is at risk of experiencing a stroke. Their primary research focus revolves around the prediction of stroke occurrence in patients, a crucial aspect for physicians to determine prognosis and offer targeted therapy within a limited timeframe. In their research, the authors have constructed an ensemble model, considering various basic, bagging, and boosting classifiers, which include support vector machine, Naive Bayes, decision tree, logistic regression, artificial neural network, random forest, XGBoost, LightGBM, and CatBoost. The experimental results presented in their work indicate that the final ensemble model, employing the Max Voting approach, achieved an impressive accuracy rate of 95.76%.

In [45] the authors of this study address the issue of assessing fatigue in athletes or patients, a task traditionally performed using costly laboratory equipment. In their research, they focus on investigating whether Inertial Measurement Units (IMUs) placed on the lower limbs, either individually or in combination, can effectively distinguish between states of fatigue and non-fatigue and predict the degree of fatigue. To accomplish this, the researchers created a multi-IMU based running fatigue dataset by recording inertial data during running sessions that led to fatigue. They conducted experiments to validate the performance of a Random Forest (RF) model and a Support Vector Machine (SVM) for classifying running fatigue and determining fatigue levels. The research findings reveal that the RF model outperformed the SVM in

classification accuracy. Moreover, as the number of sensors increased, the classification accuracy improved. Notably, the RF model achieved an accuracy of 87.21% when using IMU tibial data alone, while the highest classification accuracy of 91.10% was attained when combining tibial and femoral IMUs. The authors suggest that inertial sensors have the potential to objectively assess fatigue levels during running by detecting subtle biomechanical deviations in lower limb movements.

In [46] the authors of this study address the critical issue of intrusions into computer networks, which can disrupt network functionality and pose a significant threat to communication systems. Cyberattacks represent a substantial challenge in this context, compromising the privacy, authenticity, and availability of network resources. Intrusion Detection Systems (IDS) play a crucial role in identifying and mitigating these unauthorized actions or attacks. In their research, the authors propose a decision tree-based method for detecting network intrusions while enhancing data quality. They introduce their own model, which achieves impressive accuracy rates of 99.98% and 99.82% when tested on the CICIDS 2017 and NSL-KDD datasets, respectively. In comparison to existing models, this novel approach offers numerous advantages, particularly in terms of reducing the false alarm rate (FAR), improving detection rate (DR), and enhancing accuracy (ACC).

Table 15: Computation of Time

| Criterion | Decision Tree | Proposed |
|----------------------------|---------------|----------|
| Anomaly computation time | 120ms | 30ms |
| Signature computation time | 150ms | 50ms |
| Preprocessing time | 50ms | 15 ms |

In [47] the authors highlight the significance of IPv6 target generation in the context of rapid IPv6 scanning for online surveys. They identify a critical issue, which is the low hit rate resulting from improper space partitioning attributed to outlier addresses and myopic partition pointers. To address this challenge, the authors introduce 6Forest, a novel approach to IPv6 target generation based on ensemble learning. This approach offers global coverage and robustness against address outliers. Their experimental findings, based on eight extensive candidate datasets, demonstrate the effectiveness of 6Forest.

It outperforms state-of-the-art IPv6 scanning methods on a global scale, achieving an impressive up to 116.5% enhancement for low-cost scanning and a remarkable 15x improvement for high-cost scanning.

In [48] the researchers focused their attention on the common method for estimating Gaussian mixture models (GMM). They identified a challenge with GMM estimation, which is its susceptibility to outliers. This sensitivity to outliers can result in subpar estimation performance, depending on the dataset being analyzed. In their study, the authors suggest incorporating an outlier detection step into the Expectation-Maximization (EM) algorithm. This step assigns an anomaly score to each data sample in an unsupervised manner. The experimental results demonstrate the effectiveness of this proposed enhancement compared to other established imputation techniques, highlighting its potential benefits in GMM estimation.

In [49] the researchers have put forward an enhanced version of the Isolation Forest (IF) method by introducing two key elements. The first element involves assigning weights to the path taken by each feature, thereby generating a more informative anomaly score. The second element

modifies the aggregation function used to combine the results from individual trees within the forest. The researchers conducted extensive testing of their proposed method on various datasets. The outcomes of these tests and research highlight the significance of the first element, which enhances results at the individual tree level through the utilization of additional information. In contrast, the second contribution to the IF method involves a novel approach to aggregating results at the forest level, deviating from the original anomaly result and adopting a probabilistic tree interpretation.

In [50] the authors of this study explored various variants of the Isolation Forest (IF) method, including N-ary (NIF), fuzzy membership function-based (MIF), k-means clustering-based (KIF), two fuzzy clusters-enabled (CIF), and two fuzzy clusters with the addition of distance to the cluster center (prototype) (C₂DIF). Their research and evaluation of these IF-based methods focused on detecting road anomalies in real-world data, which is a critical challenge for maintaining modern economies and infrastructure. The results of the study demonstrated significant improvements in accuracy and a reduction in the false-positive rate compared to other state-of-the-art methods. These enhancements led to a 100% increase in sensitivity, showcasing the effectiveness of the proposed IF-based methods for road anomaly detection.

Table 16: Accuracy, Sensitivity and False Positive Rate of the Compared Algorithms

| Algorithm | Accuracy | Sensitivity | False Positive Rate |
|--------------|----------|-------------|---------------------|
| C2DIF | 94.94% | 100% | 5.16% |
| KIF | 94.94% | 100% | 5.16% |
| C2IF | 95.24% | 100% | 4.86% |
| MIG | 95.24% | 100% | 4.86% |
| NIF | 95.50% | 100% | 4.60% |
| IF | 95.47% | 96.77% | 4.56% |
| F-THRESH | 94.16% | 45.1% | 5.83% |
| MOD-Z-THRESH | 93.26% | 68.33% | 5.67% |

The research results clearly demonstrate the usefulness and effectiveness of IF-based techniques in a series of experiments. Among these techniques, NIF stands out with the highest accuracy of 95.5%, surpassing currently employed methods like F-THRESH (slightly above 94%) and MOD-Z-THRESH (approximately 93%). This signifies that IF-based techniques perform significantly better in terms of false positives, with the current rate being 4.5% compared to 6% with the previously used methods. The presented methods and their research results conclusively establish their superiority over the original IF method, making them more valuable and applicable to the provided datasets.

In [51] the researchers have tackled the challenge of dealing with a large volume of alerts generated by network security devices, which can lead to alert fatigue and slow response times. They have proposed a solution to alleviate this issue by using Extended Isolation Forest to identify and highlight anomalous alerts. This approach has significantly improved the quality of alerts for monitoring purposes.

While acknowledging that no algorithm is entirely foolproof, the research results demonstrate that their model effectively reduces the number of alerts received in the Security Operations Center (SOC) by an impressive 82.15%. This means that Security Analysts only need to focus on monitoring 17.85% of the 50,000 total alerts received from the Intrusion Detection System (IDS) system. This substantial reduction in the volume of alerts can lead to more efficient and effective threat detection and response in cybersecurity operations.

In [52] the researchers have directed their attention to the crucial task of land monitoring in agriculture to provide early warnings about land conditions and anomalies for farmers. They have introduced an anomaly detection model for terrain monitoring systems. In their approach, they utilized raw data from site monitoring and employed Isolation Forest (IF) to transform unlabeled data into labeled data. Subsequently, they developed an anomaly detection model using a Long Short-Term Memory (LSTM) autoencoder. The experimental results reveal promising outcomes for their method. The LSTM autoencoder demonstrated an accuracy of 0.95, precision of 0.96, recall of 0.99, and F1-score of 0.97. This translates to an overall accuracy of 95.72% for the proposed anomaly detection method. However, the researchers acknowledge certain limitations. They note that the quality of the area-monitoring data is not optimal, as illustrated in the PCA function graph. Additionally, the ROC curve showed less than satisfactory results, indicating that the anomaly detection method has high false positive and false negative rates. In light of these limitations, the researchers collectively agree that there is a need to re-evaluate and refine the proposed model to enhance its performance and robustness.

In [53] the researchers have emphasized the importance of detecting anomalies in the context of Industry 4.0, particularly during production processes. Their study involves a comparison of various algorithms designed for detecting anomalies in time series data collected from automotive sensors. In their investigation, the researchers evaluated several algorithms,

including the interquartile range (IQR), isolation forest, particle swarm optimization (PSO), and k-means clustering. They specifically focused on identifying anomalies within automotive systems, both during the initial phase of a vehicle's usage and throughout its entire lifecycle. The authors highlighted the significance of the vast amount of data generated by sensors inside vehicles, which amounts to over a gigabyte per second. These sensors are interconnected through the vehicle's network, consisting of electronic control units (ECU) and the Controller Area Network (CAN bus). Each ECU processes input from its sensors, executes specific commands, and monitors the vehicle's normal state while detecting any abnormal behavior. Based on their research findings, the authors concluded that, for the task of unsupervised anomaly detection in time series data from vehicle sensors, the isolation forest algorithm outperformed the IQR and PSO+K-Means algorithms in terms of accuracy and effectiveness.

In [54] the researchers address the challenge of anomaly detection in the context of heterogeneous and correlated multivariate data, and they do so without assuming prior knowledge of statistical correlation. Notably, they employ a copula-based approach to assess the statistical correlation between different data modalities. In their methodology, they utilize an unsupervised learning (UL) framework to identify anomalies by working with data points extracted from a copula-based joint distribution. Specifically, they explore various Gaussian copula techniques, including R-Vine, D-Vine, and C-Vine, in combination with anomaly detection algorithms such as isolation forest, one-class SVM, local outlier factor, elliptic envelope, and UL autoencoder. The results of their experiments indicate that the proposed framework significantly outperforms direct training methods in terms of detection accuracy, showcasing its effectiveness in addressing the problem of anomaly detection in correlated and heterogeneous multivariate data.

In [55] the researchers focus on the detection of anomalies in time series data, as this data type is crucial for analyzing historical behavior and

forecasting future trends. They emphasize that in domains such as satellite data, taxi rides, stock exchanges, online transactions, and more, the volume of generated data is so vast that manual processing becomes infeasible. In their study, the researchers employ deep learning models, including ARIMA, Isolation Forest, and LSTM-based autoencoders, to identify anomalies in datasets. The dataset used in their research pertains to daily closing prices, aiming to determine whether these prices are correct or not. The dataset is subjected to analysis using the aforementioned models. The test results reveal that ARIMA, LSTM, and Isolation Forest achieved accuracy rates of 90.13%, 84.98%, and 88.88%, respectively. The researchers suggest that further improvements can be made to these models by incorporating multidimensional time series data and optimizing them using various parameters for enhanced anomaly detection in time series datasets.

In [56] the researchers focus their attention on the issue of fraud in credit and debit card transactions, aiming to develop methods for effectively detecting online fraud when using credit cards. In their study, they explore various aspects of machine learning for fraud detection, including techniques such as the Local Outlier Factor (LOF), Isolation Forest (IF), and Convolutional Neural Networks (CNN). The research findings reveal that the authors achieved a high accuracy rate of 99% for both deep learning (CNN) and supervised machine learning techniques. Notably, the Isolation Forest method detected 73 errors, while the Local Outlier Factor identified 97 errors. Isolation Forest achieved an accuracy rate of 99.74%, and LOF achieved an accuracy rate of 99.65%. When comparing key metrics such as F1 score, precision, and recall across the three models (CNN, LOF, and IF), it became evident that the Convolutional Neural Network (CNN) outperformed the other methods.

Therefore, the conclusion drawn is that the Convolutional Neural Network method excels in identifying fraud cases, demonstrating that neural networks outperform traditional machine learning models in this context.

In [57] the researchers introduce a novel similarity search method called "Random Separations" designed to detect unknown variants of known malware in network traffic, with a focus on identifying threats. This method demonstrates superior performance when compared to other approaches, including unsupervised methods like isolation forest and lightweight online anomaly detectors, supervised approaches like random forest, and traditional similarity search algorithms such as kNN. The study involves the evaluation of eight high-risk malware families in various known-to-unknown ratios. The authors' proposed method, Random Separations, incorporates elements inspired by Random Forest and Isolation Forest. Notably, this algorithm introduces innovative features that contribute to its effectiveness in identifying threats in network traffic data.

In [58] the authors introduce a novel condition monitoring technique for wind turbine pitch systems, with a primary focus on reducing data storage requirements and computational intensity. This technique leverages the isolation forest anomaly detection model. In their approach, researchers train the new technique using data from a specific time period for each turbine. Subsequently, they evaluate each following month of data individually. The study's findings suggest that, in most cases, one month of data was sufficient for anomaly detection, while some instances required 3 or 4 months of data.

Remarkably, this technique demonstrated the capability to detect impending failures up to 3

months earlier and identify abnormal activity approximately 10 to 12 months before the actual failure occurs.

In [59] the authors introduce a versatile approach to constructing a gallery of multiple shots for observed reference identities. This method involves L2 norm descriptor matching for gallery retrieval, utilizing descriptors generated by a closed-set re-identification system. The gallery is continuously updated by replacing outliers with newly matched descriptors. To identify outliers, the authors employ the Isolation Forest algorithm, enhancing the gallery's resilience to incorrect assignments. Experimental results demonstrate that this approach surpasses state-of-the-art methods, as evidenced by improved TTR/FTR metrics. Additionally, the researchers note that this method performs effectively in controlled environments.

In [60] the authors emphasize the critical role of data accuracy from sensors in maintaining the functionality of mechanical equipment, particularly in practical industrial settings. In their work, researchers introduce an outlier detection algorithm based on a 1D neural network that offers depth separation through extended convolution. The research results demonstrate that this proposed method can enhance the accuracy of outlier identification and boasts state-of-the-art capabilities when compared to existing methods, including principal component analysis, k-means, isolation forest, local outlier, and one-class support vector machine.

Table 17: Prediction Results of Different Methods

| | Precision (%) | Accuracy (%) | AUC |
|---------------------|---------------|--------------|--------|
| PCA | 87.50 | 99.26 | 0.9375 |
| iForest | 100.00 | 99.26 | 0.9961 |
| k-means | 62.50 | 97.78 | 0.8125 |
| LOF | 50.00 | 93.33 | 0.7303 |
| OCSVM | 50.00 | 57.78 | 0.4828 |
| The Proposed Method | 100.00 | 100.00 | 1.00 |

In [61] the authors highlight the significance of outlier detection in data mining, a crucial task applicable in various domains such as fraud detection, malicious behavior monitoring, and health diagnostics. In their work, the authors

introduce PIF (Privacy-preserving Isolation Forest), designed to detect outlier values across multiple distributed data providers while offering both high performance and accuracy.

Additionally, PIF provides certain security guarantees. Research results indicate that PIF can achieve an average AUC (Area Under the Curve, a common metric for model performance)

comparable to the existing iForest method while maintaining linear time complexity, thus preserving privacy without sacrificing efficiency.

Table 18: Comparison of Runtime (S) per Dataset and Algorithm

| | H-Solution | | | V-Solution | | | iForest | | | LOF |
|-----------------|------------|-------|-------|------------|--------|--------|---------|-------|-------|---------|
| | Train | Test | Total | Train | Test | Total | Train | Test | Total | |
| http (KDDCUP99) | 0.014 | 7,189 | 7.203 | 0.034 | 69,486 | 69,520 | 0.008 | 7,495 | 7,503 | ON |
| ForestCover | 0.014 | 3,761 | 3.77 | 0.034 | 34,474 | 34,508 | 0.009 | 3.774 | 3.782 | ON |
| Smtp | 0.012 | 1.266 | 1.178 | 0.034 | 12,995 | 130.29 | 0.007 | 1,320 | 1.327 | ON |
| shuttle | 0.014 | 0.695 | 0.709 | 0.036 | 7,133 | 7,169 | 0.007 | 0.709 | 0.715 | 126,943 |
| mammography | 0.013 | 0.145 | 0.158 | 0.034 | 1.487 | 1.521 | 0.006 | 0.150 | 0.153 | 64,938 |
| Satellite | 0.021 | 0.091 | 0.112 | 0.032 | 0.909 | 0.0941 | 0.007 | 0.090 | 0.097 | 23,231 |
| Pima | 0.013 | 0.012 | 0.025 | 0.030 | 0.102 | 0.133 | 0.006 | 0.011 | 0.017 | 0.231 |
| Breastw | 0.014 | 0.009 | 0.023 | 0.029 | 0.096 | 0.125 | 0.006 | 0.009 | 0.015 | 0.240 |
| Arrhythmia | 0.078 | 0.006 | 0.084 | 0.030 | 0.062 | 0.092 | 0.006 | 0.006 | 0.012 | 0.475 |
| Ionosphere | 0.018 | 0.005 | 0.024 | 0.035 | 0.049 | 0.084 | 0.007 | 0.005 | 0.012 | 0.07 |

In [62] the authors address security concerns related to malicious users gaining unauthorized access to Virtual Machines (VMs) in cloud computing environments. They propose a proactive monitoring and anomaly detection model for VM resource usage. To develop this model, they utilize machine learning algorithms such as Isolation Forest and OCSVM (One-Class Support Vector Machine), training and testing the model on sampled VM load traces, which include various resource metrics. The research results indicate that OCSVM achieved an average F1 score of 0.97 for hourly time series and 0.89 for daily time series, while Isolation Forest achieved an average F1 score of 0.93 for hourly time series and 0.80 for daily time series. Both algorithms show promise for the presented model, but OCSVM exhibited a higher classification success rate compared to Isolation Forest.

In [63] the authors of this study conducted a comparative analysis of fault detection techniques, specifically Local Outlier Factor and Isolation Forest, and introduced a new methodology called Standardized Mahalanobis Distance. Their focus was on detecting faults in bearings and rotating machinery using data from vibration sensors. The research results indicate that the Standardized Mahalanobis Distance methodology outperforms both the Local Outlier Factor and Isolation Forest in detecting voltage drop faults in rotating machinery, particularly

when the abnormal voltage value is not close to the nominal value. Additionally, this methodology demonstrated the capability to identify outliers before the occurrence of outer race errors in the bearing, making it a valuable tool for early fault detection.

In [64] the authors of this study highlighted the emerging field of Artificial Intelligence of Things (AIoT) as a notable trend in the context of Industry 4.0, which refers to the fourth industrial revolution characterized by the integration of digital technologies into various industrial processes. Additionally, they emphasized the importance of data privacy within this context, which is a critical consideration as more and more devices and systems become interconnected and generate vast amounts of data. Protecting the privacy of this data is crucial for maintaining the integrity and security of AIoT systems.

The authors of this study were particularly interested in addressing the issue of malicious models in the context of Federated Artificial Intelligence of Things (AIoT) with learning capabilities. They proposed a model called D2MIF, which is based on an isolation forest (iforest) and is designed to detect malicious models within the federated AIoT system. Their research findings demonstrated that the D2MIF model effectively identifies and mitigates the presence of malicious models, leading to a

significant improvement in the overall accuracy of the global model in the federated AIoT system with learning capabilities.

In [65] this study, the authors conducted a comparative analysis of methods for detecting and isolating damage in a chemical process, focusing on strengthening and bagging techniques. These techniques involve training multiple weak classifiers and combining their predictions to enhance the accuracy of damage detection. The study applied boosting methods like AdaBoost and gradient boosting, as well as bagging using a random forest approach, with decision trees as the base classifier. The research findings indicated that the strengthening and bagging methods outperformed conventional techniques when applied to the benchmark Tennessee Eastman process. Both boosting and bagging methods demonstrated significant improvements in performance. Specifically, the random forest (RF) algorithm stood out for its ease of implementation, parameter tuning, and predictive power assessment, making it a practical choice for damage detection in chemical processes. The RF algorithm's out-of-bag (OOB) accuracy provided a reliable measure of its predictive capabilities without requiring separate test samples or cross-validation.

In [66] the authors of this study focus on the problem of errors in electronic spreadsheets, which are widely used in organizations for data analysis and decision-making tasks. They are particularly interested in predicting whether a specific part of a spreadsheet contains errors. To address this issue, they propose a novel approach that combines a wide range of spreadsheet metrics with modern machine learning algorithms. In their approach, the authors employ supervised learning algorithms to create error predictors. These predictors utilize data from various spreadsheet metrics included in their catalog. The experimental results presented in the study demonstrate that, in many cases, random forest classifiers are effective at predicting whether a given spreadsheet formula contains errors with high accuracy. This highlights the potential of their method for identifying errors in spreadsheets.

In [67] the authors of this study address the growing security threats and risks associated with the rapid development of IoT (Internet of Things) in smart regions and cities, particularly in environments like smart healthcare and smart homes/offices. Their specific focus is on detecting tampering with IoT security sensors in an office setting.

To tackle this problem, the authors employ machine learning techniques in two ways: They use real-time traffic patterns to train an isolation forest-based unsupervised machine learning method for anomaly detection. They create labels based on traffic patterns and apply a supervised decision tree method within their anomaly detection system using machine learning (AD-ML). The research results demonstrate that both proposed models achieve a high accuracy rate of 84% with the isolation forest silhouette metric. Additionally, the supervised machine learning model, based on 10 cross-validations for decision trees, achieved the highest classification accuracy of 91.62% with the lowest false positive rate. These findings highlight the effectiveness of their approaches in detecting sensor tampering in office IoT environments.

In [68] the authors of this study acknowledge that machine learning (ML) and artificial intelligence (AI) methods have been widely applied to enhance research outcomes in building energy management, particularly when dealing with large datasets. Their specific focus in this work is on conducting a comparative study of various unsupervised fault detection approaches for heating, ventilation, and air conditioning (HVAC) systems, with a particular emphasis on scenarios involving a limited number of faulty data points. In this research, the authors evaluate three distinct methods: Isolation Forest, One-Class Support Vector Machine (OCSVM), Long Short-Term Memory (LSTM) Autoencoders. The experimental results indicate the following: LSTM outperformed all other techniques in detecting faulty data points, achieving an average precision of around 80%. Isolation Forest performed well when dealing with a small number of erroneous data points. However, its precision decreased as the number of data points increased. OCSVM

exhibited a different pattern, where precision increased with an increasing number of faulty data points until it reached 96 points. Afterward, a regressive trend was observed. These findings suggest that LSTM autoencoders are particularly effective for fault detection in HVAC systems, especially when dealing with limited faulty data points. However, the choice of method may depend on the specific characteristics of the dataset and the number of erroneous data points involved.

In [69] the authors of this study tackle the evolving challenges associated with chronic kidney disease (CKD), which has varying incidence rates, prevalence, and progression patterns, especially in regions with diverse social determinants of health. Their primary objective is to develop an intelligent system capable of efficiently classifying patients into two categories: "CKD" or "Non-CKD." Such a system could

significantly aid healthcare professionals in expeditiously diagnosing patients, especially when dealing with a substantial caseload. In their research endeavor, the authors employ a range of unsupervised machine learning algorithms, including: K-Means Clustering, DB-Scan, Isolation Forest (I-Forest), Autoencoder. Additionally, they integrate these algorithms with various feature selection methods to optimize their performance. Notably, the integration of feature reduction methods with the K-Means Clustering algorithm yields remarkable results, achieving an impressive overall classification accuracy of 99% in distinguishing between CKD and Non-CKD clinical data. This research signifies the potential of their intelligent system, which combines K-Means Clustering with feature selection techniques, to significantly enhance the accuracy of patient classification in the context of chronic kidney disease diagnosis.

Table 19: Validation Scores for Reduced Features

| | Kmeans | Dbscan | Autoencoder | Iforest |
|------------------------|---------|----------|-------------|---------|
| Recall | 1 | 0.956 | 0.556 | 0.952 |
| Precision | 0.984 | 1 | 0.965 | 0.859 |
| F1 score | 0.992 | 0.978 | 0.706 | 0.903 |
| Accuracy score | 0.99 | 0.973 | 0.71 | 0.873 |
| Mutualinfo score | 0.926 | 0.84 | 0.258 | 0.441 |
| Adjustedrand score | 0.96 | 0.893 | 0.171 | 0.55 |
| Vmeasure score | 0.926 | 0.84 | 0.258 | 0.441 |
| Silhouette score | 0.321 | 0.327 | 0.211 | 0.178 |
| Calinskiharabasz score | 195,001 | 2001,258 | 138.11 | 109,964 |
| Daviesbouldin score | 1.055 | 1.115 | 1,396 | 1,223 |
| TP | 250 | 239 | 139 | 238 |
| TN | 146 | 150 | 145 | 111 |
| FB | 4 | 0 | 5 | 39 |
| FN | 0 | 11 | 111 | 12 |

In [70] the authors highlight the transformative impact of Internet of Things (IoT), cloud computing, and artificial intelligence (AI) on the healthcare sector, ushering in the era of smart healthcare. Their research centers on introducing an innovative disease diagnosis model that harnesses the convergence of AI and IoT, with a specific focus on diagnosing heart diseases and diabetes. In this novel approach, the authors leverage the CSO-CLSTM model, which stands for "Cascaded Long Short-Term Memory" and

incorporates the Crow Search Optimization (CSO) algorithm. Additionally, they employ the isolation forest technique (iForest) to effectively eliminate outliers from the dataset. The research outcomes are noteworthy, demonstrating the CSO-LSTM model's impressive diagnostic capabilities. It achieves a remarkable accuracy rate of 96.16% in diagnosing heart diseases and an even higher accuracy rate of 97.26% for diabetes diagnosis. These results underscore the potential of the proposed CSO-LSTM model as a valuable tool for

disease diagnosis within intelligent healthcare systems.

In [71] the authors highlight the critical issue of faults occurring in heating, ventilation, and air conditioning (HVAC) chiller systems, emphasizing the negative consequences such as energy wastage, user discomfort, reduced equipment lifespan, and system unreliability. They emphasize the importance of early anomaly detection to prevent these issues and conserve energy. In their research, the authors introduce a data-driven approach designed to identify common faults in chiller systems. The proposed method utilizes Kernel Principal Component

Analysis (KPCA) to capture the system's normal operating conditions. KPCA proves to be highly effective in handling non-linear phenomena, thanks to the incorporation of a Gaussian kernel. Moreover, a self-tuning procedure ensures optimal accuracy while maintaining strong generalization capabilities. Experimental findings underscore the superiority of the KPCA approach compared to linear PCA. Furthermore, the KPCA method outperforms other anomaly detection techniques, including local outlier factor, one-class support vector machine, and isolation forest. These results demonstrate the effectiveness of KPCA in fault detection within HVAC chiller systems.

Table 20: Imulated Dataset—Classification Results: Comparison Between Linear and Kernel PCA

| | | | Normal classified (%) | Faults classified (%) | BA |
|------|------|--------|--------------------------|--------------------------|-------|
| Rcaf | PCA | Normal | 89.3 | 10.7 | 53.55 |
| | | Faults | 82.2 | 17.8 | |
| revf | KPCA | Normal | 88.2 | 11.8 | 85.50 |
| | | Faults | 17.2 | 82.8 | |
| revf | PCA | Normal | 92.3 | 7.7 | 57.7 |
| | | Faults | 76.9 | 23.11 | |
| | KPCA | Normal | 91.1 | 8.9 | 93.20 |

In [72] the authors underscore the importance of security in the transmission of information through Wireless Sensor Networks (WSNs). They emphasize that effective anomaly detection is a crucial element in ensuring the security of IoT systems. To address these concerns, they introduce a novel method named BS-iForest, which is based on an isolation forest variant designed specifically for wireless sensor networks.

One distinctive aspect of this approach is its utilization of a subset of data filtered by a boxplot for training and constructing trees. Notably, the authors deliberately do not select isolation trees with higher accuracy during the training phase to build the foundational forest anomaly detector.

Subsequently, this base forest anomaly detector is employed to assess data outliers in subsequent periods, facilitating efficient anomaly detection in the WSN context.

In [73] the authors introduce an innovative approach to enhance anomaly detection in

datasets where conventional methods, such as the isolation forest (IF), struggle due to the unique characteristics of the data. Their solution involves leveraging a neural network trained for data reconstruction, coupled with the IF algorithm, which is known for its ability to identify outliers in datasets. In this method, an autoencoder is developed, where the neural network learns to generate a compact representation of the input data. Subsequently, the IF algorithm is applied to the reconstructed data to identify anomalies. This combined approach effectively improves the process of detecting anomalies in multidimensional datasets, addressing the challenges posed by the specific context and nature of the data.

In [74] the authors present modifications to the attention-based Isolation Forest (ABIF), which is based on Nadaraya-Watson regression. These modifications aim to enhance anomaly detection. The central idea behind their approach is to assign attention weights to each tree path using learnable parameters that depend on both the

instance being evaluated and the trees themselves. Importantly, the authors propose this solution without relying on gradient-based algorithms. Their numerical experiments involve synthetic and real datasets, demonstrating that ABIForest outperforms other methods. Notably, the test results confirm that ABIForest can detect

additional anomalies that the original Isolation Forest failed to identify. The authors' straightforward attention-based model mechanism represents an initial step toward incorporating various forms of attention mechanisms into Isolation Forest. This inclusion holds promise for advancing research in anomaly detection.

Table 21: F1- Score Measures of the Original iForest and ABIForest as Functions of the Number of Anomalous Instancesnanom

| The Ionosphere Dataset | | | | | |
|------------------------|-------|-------|-------|-------|-------|
| nano | 10 | 20 | 40 | 50 | 60 |
| iForest | 0.655 | 0.960 | 0.694 | 0.687 | 0.681 |
| ABIForest | 0.692 | 0.709 | 0.711 | 0.695 | 0.687 |

In [75] the authors highlight the global shift towards distributed energy resources, particularly solar energy generated from photovoltaics, as a key renewable energy source. However, they stress the importance of detecting anomalies in individual photovoltaic panels due to their potential impact on performance and safety, including fire hazards. To address this, the researchers have developed techniques for accurately and effectively identifying anomalies in photovoltaic systems. Their work primarily focuses on the performance analysis of the isolation forest technique for anomaly detection in photovoltaic systems and the use of a rule-based fault localization technique to pinpoint faulty panel events. Through a series of experiments, the isolation forest method successfully detected around 453 anomalies among 45,740 observations, with approximately six panels indicating system failures. The researchers emphasize that any unusual behavior in the system can trigger the rule-based fault detection method. The research results demonstrate the stability and effectiveness of the presented

approach in various challenging situations. The accuracy of the fault detection method is notably high, approximately 0.9886, indicating its suitability for identifying faults in photovoltaic systems.

In [76] the authors emphasize the growing significance of the Internet of Things (IoT) and the need for improved cybersecurity measures to counter emerging cyberattacks in this context. Their research aims to develop a malware traffic detection architecture that operates by analyzing summarized statistical packet data, rather than processing information from entire packets. This approach offers the advantage of enabling the analysis of a vast number of IoT devices while reducing the requirements for data storage and computational power. To achieve this goal, the authors employ two machine learning techniques, namely isolation forest and k-means clustering, in their architecture to identify malware traffic patterns. The research results aim to provide insights into the effectiveness of these techniques in detecting and mitigating malware threats within the specified architecture.

Table 22: Comparison Table

| | Isolation | K-means |
|---------------------|-----------|---------|
| TPR | ABOUT | x |
| FPR | ABOUT | ABOUT |
| MCC | ABOUT | ABOUT |
| C&C detection | ABOUT | ABOUT |
| Host scan detection | ABOUT | ABOUT |
| DoS detection | ABOUT | ABOUT |
| speed | ABOUT | ABOUT |

In [77] the authors emphasize the significance of the Building Internet of Energy (BIE) in optimizing energy consumption, cost reduction, and enhancing building transformations. They stress that combining artificial intelligence with BIE is essential for effectively analyzing big data and enabling intelligent decision-making. In their research, the authors introduce an intelligent approach for detecting anomalies in energy consumption patterns. Unlike traditional methods that focus on consumption values, this approach analyzes the shape of daily energy consumption curves. The authors employ a two-stage hybrid approach that combines supervised and unsupervised learning techniques. They utilize eXtreme Gradient Boosting (XGBoost) to create a regression model, enabling the classification of consumption anomalies on weekdays using rule-based algorithms and residuals. Subsequently, they employ the isolation forest algorithm for unsupervised anomaly detection. The results of their study demonstrate that this approach achieves an anomaly detection accuracy rate of 95.93%.

In [78] the authors focus on the critical role of networks in contemporary life and the increasing importance of digital security research. Their work centers on the development of an effective machine learning-based Intrusion Detection System (IDS) for detecting HTTP/2.0 slow Denial of Service (DoS) attacks. They begin by utilizing real attack-based datasets to understand the model better. In their study, the researchers explore the capabilities of three single-class classifier algorithms: Support Vector Machine (SVM), Isolation Forest (IF), and Minimum Covariant Determinant (MCD). The test results indicate that one of these classifier algorithms outperforms the others across various measurements, including accuracy (0.99), sensitivity (0.99), and specificity (0.99).

In [79] the researchers are primarily concerned with identifying and addressing corrupted data, which can result from various unethical and illegal sources. They emphasize the importance of developing a highly effective method for detecting and properly handling corrupted data within a

dataset. In their work, they introduce PAACDA, which stands for Proximity-Based Adamic Adar Corruption Detection Algorithm. Their approach focuses specifically on detecting corrupted data rather than outliers, and they consolidate the results accordingly.

In their study, the authors introduce a novel PAACDA algorithm, which demonstrates superior performance compared to other unsupervised learning benchmarks. They conducted evaluations against 15 popular baselines, including K-means clustering, isolation forest, and LOF (Local Outlier Factor). The results indicate an accuracy of 96.35% for clustered data and an impressive 99.04% for linear data.

In [80] their research, the authors present a theoretical framework that examines the effectiveness of isolation methods from a distributional perspective. They propose that algorithms capable of fitting a mixture of distributions, where the average path length of observations approximates the mixture coefficient, can provide valuable insights. Building on this concept, they introduce a Generalized Isolation Forest (GIF), which goes beyond the traditional use of average path length when combining random trees. Through an extensive evaluation of over 350,000 experiments, the researchers demonstrate that GIF outperforms other methods across various datasets while maintaining comparable runtime. Their work deepens the theoretical understanding of isolation-based techniques and introduces a novel algorithm for improved outlier detection.

Table 23: ROC AUC Score for Isolation-Based Methods (EIF, GIF, IF, and SciF) and Proximity-Based Methods (iNNE, LeSiNN, aNNE). Results From [4] Are Also Included Which Evaluate 12 Additional Proximity-Based Methods. Bold Highlights the Best Isolation-Based Out

| | Gif | Eif | If | Science fiction | Anne | Other | Lesinn | [3] |
|------------------|--------|--------|--------|-----------------|--------|--------|--------|--------|
| Annthyroid | 0.6474 | 0.6263 | 0.708 | 0.518 | 0.5980 | 0.5363 | 0.5903 | 0.7666 |
| Cardiotocography | 0.9285 | 0.7923 | 0.7944 | 0.7066 | 0.6546 | 0.8172 | 0.6601 | 0.8279 |
| Creditfraud | 0.9421 | 0.9522 | 0.9513 | 0.9133 | 0.9447 | 0.9516 | 0.9529 | 0.9624 |
| Forestcover | 0.9408 | 0.92 | 0.9236 | 0.7013 | 0.7474 | 0.9568 | 0.7988 | 0.8847 |
| KDDCup99 | 0.9831 | 0.9909 | 0.9895 | 0.9905 | 0.9778 | 0.9859 | 0.9152 | 0.9904 |
| mammography | 0.8706 | 0.8119 | 0.8175 | 0.5922 | 0.7345 | 0.8018 | 0.7683 | 0.8435 |
| PageBlocks | 0.9342 | 0.9271 | 0.9281 | 0.8616 | 0.8408 | 0.9461 | 0.8904 | 0.96 |
| PenDigits | 0.9644 | 0.9205 | 0.9154 | 0.8803 | 0.9369 | 0.8236 | 0.9413 | 0.9903 |
| Pi has | 0.8355 | 0.7453 | 0.7415 | 0.6415 | 0.6021 | 0.7272 | 0.6805 | 0.7759 |
| Satellite | 0.8575 | 0.741 | 0.7277 | 0.5893 | 0.6791 | 0.745 | 0.6602 | 0.7315 |
| ShuUle | 0.85% | 0.802 | 0.868 | 0.9847 | 0.7051 | 0.8074 | 0.7259 | 0.9484 |
| Spambase | 0.7508 | 0.7889 | 0.8147 | 0.8216 | 0.7038 | 0.7409 | 0.7160 | 0.7914 |
| Waveform | 0.9115 | 0.7289 | 0.7192 | 0.7115 | 0.6701 | 0.7578 | 0.6527 | 0.7713 |
| Veela | 0.5683 | 0.4042 | 0.5188 | 0.4796 | 0.5282 | 0.5363 | 0.5903 | 0.7858 |

In [81] their research, the authors introduce an enhanced SA-iForest method for data anomaly detection. Their primary objective is to address the issues of low accuracy and execution efficiency associated with general data anomaly detection algorithms based on the Isolation Forest (IF) method. Their approach focuses on selective integration, prioritizing precision and value

differentiation. To achieve this, the authors utilize the simulated annealing algorithm to select isolation trees with a high capability for detecting anomalies, optimizing the forest in the process. Their proposed approach results in substantial improvements in algorithm performance by enhancing the construction of the isolated forest.

Table 24: Execution Time of Different Data Sets

| Date set name | SA-iForest | iForest | LOF |
|-------------------|------------|---------|---------|
| Breast | 0.11 | 0.23 | 1.14 |
| Diabest | 0.24 | 0.29 | 2.07 |
| Unblanched | 0.27 | 0.38 | 3.63 |
| Http | 8.83 | 32.67 | 9186.15 |
| Shuttle | 3.21 | 9.36 | 736.08 |
| Pendigits | 1.4 | 4.88 | 376.12 |
| Messidor_features | 0.89 | 1.16 | 4.54 |

The research findings indicate that both SA-iForest and iForest algorithms exhibit significantly lower execution times compared to the LOF algorithm. This reduction in execution time is attributed to the fact that SA-iForest and iForest do not involve distance or density calculations, which are not utilized for anomaly detection. Consequently, eliminating distance and density-based methods results in reduced computational costs. The researchers highlight that SA-iForest achieves a lower computational complexity constant. In the SA-iForest algorithm, tree selection is based on simulated annealing for improved detection performance. As a result, it does not construct the entire set of trees used for

anomaly detection, unlike the original iForest, which builds all the trees for detection purposes. By removing some detection efficiency and employing the fast convergence of the simulated annealing optimization algorithm, SA-iForest enhances network intrusion anomaly detection performance. The presented test results affirm that SA-iForest exhibits lower execution times than iForest and demonstrates significant improvements in overall generalization and prediction performance.

In [82] the authors of this research focus on enhancing the security of computer networks,

particularly by detecting intrusions at an early stage to minimize their impact, which is critical for an institution's financial well-being and reputation. They emphasize the two primary intrusion detection systems (IDS): signature-based and anomaly-based IDS. Signature-based IDS: These systems identify intrusions by comparing network traffic patterns to a database of known intrusion signatures. If a match is found, it signals an intrusion. Anomaly-based IDS: In contrast, anomaly-based IDS systems identify intrusions by establishing a baseline of normal network behavior. Deviations from this baseline are flagged as potential intrusions. In their research, the authors employ real-time traffic data from the University of Virginia network to evaluate the performance of two anomaly-based intrusion detection methods: Local Outlier Factor (LOF) and Isolation Forest (iForest). They aim to understand the similarities and differences in the results produced by each approach. The results of the research indicate that iForest scores exhibit greater specificity across all data points compared to LOF scores. This suggests that, especially when anomalies are considered rare and distinct data points, the Isolation Forest method performs better at identifying anomalies than LOF.

In [83] the authors of this research highlight the growing volume of hydrological data and the

challenges associated with applying anomaly detection algorithms efficiently. They point out that current algorithms often suffer from low time efficiency and produce an excessive number of anomaly points, making it difficult for decision-makers to extract meaningful insights. To address these issues, the researchers present an application of the Isolation Forest algorithm for hydrological pattern representation data. They further propose an Isolation Forest-based hydrological time series anomaly pattern detection algorithm, which they validate through a series of experiments. One significant challenge they tackle is the difficulty in determining the partition threshold in the Isolation Forest and deriving the top-k anomalies. To overcome this, they integrate the k-means clustering algorithm and the nearest neighbor algorithm into the Isolation Forest method. This enhancement reduces the subjectivity associated with manually setting thresholds and improves result stability.

The authors apply these algorithms to data collected from the Chuhe River Basin and compare their improved Isolation Forest method with other anomaly detection algorithms in terms of accuracy and time complexity. The results demonstrate that the improved Isolation Forest algorithm is well-suited for characterizing large hydrological time series and achieves high accuracy in anomaly detection.

Table 25: Comparison of Four Algorithms AUC and Run Time

| Algorithm name | AUC | Time/ms |
|------------------|--------|---------|
| Improved iForest | 0.9239 | 39 |
| NLOF | 0.9089 | 278 |
| LOF | 0.8891 | 475 |
| Improved k-means | 0.7267 | 1136 |

In [84] the authors of this research highlight the significant volume of electricity data collected from the distribution network and the potential for incorrect data to occur, which can adversely impact data analysis. In response, they propose an Attribute-Linked Isolation Forest algorithm designed to detect anomalous electrical data. In their algorithm, attributes are regrouped based on attribute associations, and they modify the partition generation method. By using data

samples to traverse the isolation forest, abnormal data is identified based on the anomaly score.

Furthermore, they suggest that these anomalies can be corrected using the modified Wavelet Neural Network method. The authors emphasize that their method effectively and quickly detects incorrect data and demonstrates high accuracy in both identifying and correcting such data anomalies, addressing a critical issue in working with electrical data.

Table 26: Anomaly Data and Correction Results

| Original Anomaly Data | | | Corrected Data | | |
|-----------------------|-------------------------|---------|---------------------|---------------------------|-------------|
| Active Pow Wher (k) | Reactive Power (kVar/h) | Voltage | Active Power (kW/h) | Reactive Power (kVar/ bh) | Voltage (V) |
| 10.314 | 0.007 | 234.84 | 4.216 | 0.418 | 234.84 |
| 5.36 | 3,663 | 233.63 | 5.36 | 0.436 | 233.63 |
| 5,388 | 0.502 | 543.70 | 5,388 | 0.502 | 233.74 |
| 3.52 | 0.522 | 544.18 | 3.52 | 0.522 | 235.02 |
| 3.702 | 0.02 | 33.32 | 3.702 | 0.52 | 235/09 |
| 3.7 | 4.34 | 235.22 | 3.7 | 0.52 | 235.22 |
| 0.067 | 0.51 | 233.99 | 3.668 | 0.51 | 233.99 |
| 3.662 | 0.01 | 233.86 | 3.662 | 0.51 | 233.86 |
| 4.448 | 0.498 | 657.55 | 4.448 | 0.498 | 232.86 |
| 5.412 | 0.47 | 352.66 | 5.412 | 0.47 | 232.78 |
| 1,670 | 0.478 | 232.99 | 5.224 | 0.478 | 232.99 |
| 5.268 | 3.398 | 232.91 | 5.268 | 0.398 | 232.91 |
| 4,054 | 0.422 | 499.08 | 4,054 | 0.422 | 235.24 |
| 3,384 | 2,677 | 237.14 | 3,384 | 0.282 | 237.14 |
| 6.27 | 0.152 | 236.73 | 3.27 | 0.152 | 236.73 |
| 6.43 | 0.156 | 237/06 | 3.43 | 0.156 | 237/06 |
| 3.266 | 0 | 437.13 | 3.266 | 0.237 | 437.13 |
| 3,728 | 0 | 677.10 | 3,728 | 0.235 | 435.84 |

The results presented by the researchers show that their proposed method detected 15 anomalies, with a false detection rate of 0.08%. In comparison, using the original Isolation Forest (IF) method, a total of 21 anomalous data points were detected, resulting in a higher false detection rate of 0.16%. This indicates that the researchers' method outperforms the traditional IF approach, achieving higher efficiency and accuracy in anomaly detection while mitigating the issues of redundancy.

In [85] the researchers focused on software defect prediction (SDP) in software engineering, which aims to assess the quality and reliability of software. They introduced an innovative SDP method based on the isolation forest (IF) algorithm, with a particular emphasis on feature selection. Their experiments were conducted on five real NASA datasets. The researchers found that the selection of features for building an isolation tree had a significant impact on prediction performance. Carefully chosen features were more effective than randomly selected ones.

The results indicated that their proposed method, which leverages the isolation forest algorithm, could help mitigate issues associated with

unbalanced training data in software defect prediction.

In [86] the researchers addressed the challenge of ensuring proper identification in modern applications, especially when dealing with unlabeled data. Their focus was on practical applications, and they aimed to compare various unsupervised anomaly detection techniques using performance metrics like precision, recall, F-score, and the area under the curve (AUC). They experimented with techniques such as One-Class Support Vector Machine (OneClassSVM), Local Outlier Factor (LOF), Isolation Forest (IF), and Elliptic Envelope (EE) using datasets from shuttles and satellites. The research results showed that these unsupervised anomaly detection techniques had varying levels of performance on the given datasets. However, the specific findings or performance metrics achieved were not provided in the summary.

Table 27: Area Under the Curve for Shuttle Dataset

| Alhorithm | AUC |
|---------------|------|
| One Class SVM | 0.95 |
| LOF | 0.87 |
| IF | 0.99 |
| EE | 0.94 |
| SVM | 0.85 |

In [87] the researchers addressed the challenge of analyzing a vast amount of raw data from social media, which contains ironic statements. Manual analysis of this data is challenging due to its rapid growth and limitations posed by character constraints and typographical errors on social media platforms. Traditional classification methods were found to be insufficient for this task. Therefore, the researchers treated irony detection in online social media as a classification problem and assessed the performance of various supervised machine learning methods on real data. The machine learning algorithms applied for irony detection included Bayesian Network (BayesNet), OneR, Stochastic Gradient Descent (SGD), Logistic Model Tree (LMT), Multi-Layer Perceptron (MLP), Radial Basis Function Networks (RBF), Voted Perceptron, IBk, Randomizable Filtered Classifier (RFC), Isolation

Forest, Fuzzy Lattice Reasoning (FLR), and Bagging algorithms. The researchers conducted comprehensive evaluations of these methods. The results of the experiments varied depending on the evaluation criteria and the allocation of training and test data. The Voted Perceptron, IBk, and RFC algorithms performed best across various evaluation metrics. Specifically, when the entire dataset was used as training data, IBk and RFC achieved the best results. However, when 70% of the dataset was allocated as training data and the remainder as test data, Voted Perceptron outperformed the other algorithms in terms of most evaluation metrics, except for accuracy. The researchers also conducted 10-fold cross-validation and reported findings, although specific details about those results were not provided in the summary.

Table 28: Results of Experiment III

| Classification algorithms | Evaluation Metrics | | | |
|---------------------------|--------------------|-----------|--------|-----------|
| | Accuracy | Precision | Recall | F-measure |
| BayesNet | 0.669 | 0.670 | 0.669 | 0.669 |
| OneR | 0.674 | 0.674 | 0.674 | 0.674 |
| SGD | 0.674 | 0.674 | 0.674 | 0.674 |
| LMT | 0.671 | 0.671 | 0.671 | 0.671 |
| MLP | 0.671 | 0.671 | 0.671 | 0.671 |
| RBF | 0.660 | 0.660 | 0.660 | 0.660 |
| VotedPerceptron | 0.676 | 0.676 | 0.676 | 0.676 |
| IBk | 0.668 | 0.668 | 0.668 | 0.667 |
| RFCs | 0.669 | 0.668 | 0.669 | 0.667 |
| IsolationForest | 0.446 | 0.448 | 0.446 | 0.428 |
| FLR | 0.504 | 0.506 | 0.505 | 0.505 |
| Bagging | 0.672 | 0.672 | 0.672 | 0.672 |

In [88] the researchers highlight the importance of monitoring the production process and detecting anomalies to maintain a high level of quality in the final product. In their study, they propose an integrated stack-based anomaly detection method called SBIOD. This method combines five different learning components: HBOS (histogram-based outlier detection), LOF

(local outlier factor), iForest (isolation forest), DT (decision tree), and LR (logistic regression). In addition to these components, the researchers introduce a feature normalization step, and they incorporate the MLP (Multi-Layer Perceptron) classifier to analyze feature values. The goal is to improve the performance of anomaly detection while maintaining practical value. The presented

research results indicate that the SBIOD method significantly enhances the AUC (Area Under the Curve) metric, which is a common measure of the performance of classification models. This

suggests that the SBIOD approach is effective in detecting anomalies in the production process and is suitable for practical applications where maintaining product quality is crucial.

Table 29: Results from Different Models for Sugar

| Algorithms | Auc before cross validation | Auc after cross validation | Recall |
|---------------------|-----------------------------|----------------------------|--------|
| HBOS | 0.634 | 0.623474 | 0.04 |
| LOF | 0.462 | 0.409413 | 0.01 |
| iFores1 | 0.534 | 0.479122 | 0.26 |
| DT | 0.956 | 0.994589 | 0.92 |
| Logistic Regression | 0.764 | 0.970663 | 0.54 |
| SBIOD | 0.95 | 0.9828 | 0.94 |

In [89] the researchers focus on addressing abuses by call center agents and present a machine learning model for abuse detection. They base this model on features extracted from audio recordings of phone calls. To identify the most effective approach, they train, evaluate, and compare three different models: one-class support vector machine, isolation forest, and multivariate Gaussian models. The study's results reveal that a combination of the recursive SVM feature elimination scheme and the isolation forest algorithm performs best for detecting three out of four types of abuse in the phone call recordings. This indicates that this specific combination of techniques is the most suitable for identifying instances of abuse in call center interactions.

In [90] the authors emphasize the challenges in processing ticket purchase applications due to their sheer volume, making management and analysis difficult. To address this issue, they propose the use of both unsupervised and supervised learning techniques to identify invalid tickets. These techniques include autoencoders, angle-based outlier detection, gradient boosting, isolation forests, and neural networks. The authors note that by identifying outliers within the ticket applications, government entities can significantly improve their efficiency in handling these requests, ultimately enhancing their relationship with city residents. The study's results indicate that unsupervised approaches were most effective when examining the dataset for reports with unusual attribute combinations. Conversely, supervised approaches performed best in identifying tickets with unreasonable

turnaround times. Furthermore, the authors highlight that the presence of true outliers can lead to better evaluations of various anomaly detection algorithms, thereby improving their overall functionality. This, in turn, allows for more accurate detection of future outliers in similar datasets.

In [91] the authors highlight the significance of product reviews and comments on popular auction websites, as they play a crucial role in influencing buying and selling decisions. They specifically address the issue of fake reviews, which can mislead consumers with fraudulent information, leading to financial losses. In their research, the authors propose a method for detecting fake reviews based on the review records associated with products. Their approach involves extracting product review records into a temporal feature vector. They then develop an isolation forest algorithm to detect outlier reviews, with a focus on identifying differences in product review patterns to pinpoint outlier reviews. This method aims to enhance the authenticity and reliability of product reviews on auction websites.

In [92] the authors introduce a novel outlier detection method based on machine learning in their research. They utilize the isolation forest algorithm to compute the outlier coefficient. In the subsequent step, they establish an outlier detection model, essentially transforming the problem of outlier detection into a binary classification learning model. According to the results presented in their research, this new

method for detecting outliers is competitive with the Z-score method, as it achieves superior results in terms of accuracy and effectiveness. This suggests that their approach offers a promising alternative for outlier detection in various applications.

In [93] their research, the authors focus on the application and analysis of four machine learning (ML) techniques for diagnosing failures in vehicle fleet tracking modules. They specifically compare various sampling methods during the training and testing process using real data from DDMX, a

company involved in vehicle fleet tracking. The authors create 16 models using Random Forest (RF), Naive Bayes (NB), Support Vector Machine (SVM), and Multi-Layer Perceptron (MLP) techniques. Their results indicate that these techniques achieve high precision rates of 99.76% and 99.68%, respectively, in detecting and isolating errors in the provided dataset. They suggest that the models developed in their work can serve as prototypes for remote fault diagnosis, showcasing their potential for practical applications in vehicle fleet management and maintenance.

Table 30: Time to Train and Evaluate the Models (S)

| | | RF | | N.B | | SVM | | MLP | |
|-------|------|-------|------|-------|------|--------|------|-------|------|
| | | Train | Test | Train | Test | Train | Test | Train | Test |
| Set 1 | U.S | 0.77 | 0.31 | 0.04 | 0.31 | 269.00 | 0.50 | 3.06 | 0.36 |
| | AXIS | 1.00 | 0.20 | 0.01 | 0.31 | 565.00 | 0.54 | 3.27 | 0.35 |
| Set 2 | U.S | 0.53 | 0.31 | 0.08 | 0.32 | 102.00 | 0.55 | 0.54 | 0.28 |
| | AXIS | 1.21 | 0.38 | 0.11 | 0.47 | 0.865 | 0.4 | 2.51 | 0.27 |

In [94] their research, the authors highlight the application of machine learning techniques in developing intelligent solutions for detecting anomalies in various computer and communication systems. Their specific focus in this study is to assess the performance of both supervised (such as K-Nearest Neighbors - KNN and Support Vector Machine - SVM) and unsupervised (like Isolation Forest and K-Means) algorithms for intrusion detection, using the UNSW-NB12 dataset. The research results presented in their study show that the supervised SVM Gaussian fine algorithm achieved an impressive accuracy rate of 92%. This high accuracy indicates the algorithm's capability to effectively classify normal and abnormal data.

However, they also note that unsupervised algorithms like the K-Means algorithm are proficient in grouping data and determining the appropriate number of clusters. Nevertheless, they point out that the UNSW-NB12 dataset used in their study exhibits high data density, which can pose challenges for clustering algorithms.

IV. CONCLUSION

In this paper it has been done a brief review of existing outlier detection methods that have been

developed in the last 5 years. Methods based on outlier detection are widely applicable to various data sets. Because there are many anomaly detection models.

From the analysis of the above texts emerges a fascinating picture of the role of Isolation Forests in the domain of anomaly detection and outlier observation. This tool, based on decision trees, demonstrates its potential across various contexts and domains, making a valuable contribution to the fields of data analysis and cybersecurity. In the realm of computer security, Isolation Forests play a pivotal role in Intrusion Detection Systems (IDS). Research suggests that it is a promising tool for identifying unauthorized activities and attacks. Importantly, Isolation Forests have the ability to detect new types of threats that have not yet been classified in databases, which is crucial for protecting sensitive data and maintaining the integrity and reliability of computer systems. In the area of fault diagnostics and device monitoring, Isolation Forests are equally valuable.

They enable the detection of errors and abnormalities in various mechanical and technical systems. Industries where failures can lead to significant financial losses and decreased efficiency benefit from the quick detection and

response to anomalies that this tool provides. In the field of large-scale data analysis, Isolation Forests open doors to discovering atypical patterns and outlier observations. They allow for the identification of hidden dependencies within data, which is essential for making informed business decisions. This tool offers the opportunity for in-depth data exploration and a better understanding of its structure, which is invaluable in today's world where data is a source of competitive advantage. Concerning fraud detection and abuse prevention, Isolation Forests prove useful in identifying fake product reviews or abuses in customer service. This is important not only from the perspective of consumer protection but also for companies aiming to maintain the integrity and quality of their products and services. It provides an avenue for safeguarding a company's reputation and building trust with consumers. It's worth emphasizing that Isolation Forests come with several advantages, such as higher precision, efficiency, and speed compared to other anomaly detection techniques. However, it's essential to remember that the effectiveness of this algorithm may depend on the specific data and problem context. Therefore, ongoing research and experimentation are crucial to tailor this technique to various applications and challenges.

In conclusion, Isolation Forests remain an incredibly valuable tool in data analysis, cybersecurity, and outlier detection across multiple domains. Research and development in this technique have the potential to bring further enhancements and broaden its applications in the future. This tool continues to evolve, opening new possibilities for researchers and professionals engaged in data analysis and cybersecurity. Isolation Forests serve as a foundation for innovative advancements in the field of anomaly detection, enhancing our ability to understand and protect the digital world.

REFERENCES

1. Puggini, L., & McLoone, S. (2018). An enhanced variable selection and Isolation Forest based methodology for anomaly detection with OES data. *Engineering Applications of Artificial Intelligence*, 67, 126-135.
2. Ounacer, S., Ait El Bour, H., Oubrahim, Y., Ghoumari, M. Y., & Azzouazi, M. (2018). Using Isolation Forest in anomaly detection: the case of credit card transactions. *Periodicals of Engineering and Natural Sciences*, 6 (2), 394-400.
3. Li, X., Cai, Y., & Zhu, W. (2018). Power Data Cleaning Method Based on Isolation Forest and LSTM Neural Network. W X. Sun, Z. Pan i E. Bertino (Red.), *Cloud Computing and Security* (s. 11067). Springer. doi: 10.1007/978-3-030-00018-9_47.
4. Habler, E., & Shabtai, A. (2018). Using LSTM encoder-decoder algorithm for detecting anomalous ADS-B messages. *Computers & Security*, 78, 155-173.
5. Liu, Z., Liu, X., Ma, J., & Gao, H. (2018). An Optimized Computational Framework for Isolation Forest. *Mathematical Problems in Engineering*, 2018, Article ID 2318763, 13 pages. doi: 10.1155/2018/2318763.
6. Chun-Hui, X., Chen, S., Cong-Xiao, B., & Xing, L. (2018). Anomaly Detection in Network Management System Based on Isolation Forest. W 2018 4th Annual International Conference on Network and Information Systems for Computers (ICNISC) (s. 56-60). IEEE. doi: 10.1109/ICNISC.2018.00019.
7. Tao, X., Peng, Y., Zhao, F., Zhao, P., & Wang, Y. (2018). A parallel algorithm for network traffic anomaly detection based on Isolation Forest. *International Journal of Distributed Sensor Networks*, 14(11). doi: 10.1177/1550147718814471.
8. Liao, L., & Luo, B. (2018). Entropy Isolation Forest Based on Dimension Entropy for Anomaly Detection. *Communications in Computer and Information Science*.
9. Filippov, A. I., Iuzbashev, A. V., & Kurnev, A. S. (2018). User authentication via touch pattern recognition based on isolation forest. W 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EICONRUS) (s. 1485-1489). IEEE. doi: 10.1109/EICONRUS.2018.8317378.
10. Kurnianingsih, L. E., Nugroho, E., Widyawan, L., Lazuardi, L., & Prabuwono, A. S. (2018).

Detection of Anomalous Vital Sign of Elderly Using Hybrid K-Means Clustering and Isolation Forest. *W TENCON 2018 - 2018 IEEE Region 10 Conference* (s. 0913-0918). IEEE. doi: [10.1109/TENCON.2018.8650457](https://doi.org/10.1109/TENCON.2018.8650457).

11. Wu, Y. -JA Zhang, & X. Tang. (2018). Isolation Forest Based Method for Low-Quality Synchrophasor Measurements and Early Events Detection. *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 1-7. DOI: [10.1109/SmartGridComm.2018.8587434](https://doi.org/10.1109/SmartGridComm.2018.8587434).

12. Togbe, M. U. et al. (2020). "Anomaly Detection for Data Streams Based on Isolation Forest Using Scikit-Multiflow." In: et al. *Computational Science and Its Applications – ICCSA 2020. ICCSA 2020. Lecture Notes in Computer Science ()*, vol 12252. Springer, Cham. DOI: [10.1007/978-3-030-58811-3_2](https://doi.org/10.1007/978-3-030-58811-3_2).

13. Togbe, M.U., Chabchoub, Y., Boly, A., Chiky, R.: "Etude comparative des méthodes de détection d'anomalies." *Revue des Nouvelles Technologies de l'Information Extraction et Gestion des Connaissances, RNTI-E-36*, 109–120 (2020).

14. Hara, Y., Fukuyama, Y., Murakami, K., Iizaka, T., & Matsui, T. (2020). Fault Detection of Hydroelectric Generators using Isolation Forest. *2020 59th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, Chiang Mai, Thailand. DOI: [10.23919/SICE48898.2020.9240331](https://doi.org/10.23919/SICE48898.2020.9240331).

15. Ma, H., Ghojogh, B., Samad, M. N., Zheng, D., & Crowley, M. (2020). Isolation Mondrian Forest for Batch and Online Anomaly Detection. *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Toronto, ON, Canada. DOI: [10.1109/SMC42975.2020.9283073](https://doi.org/10.1109/SMC42975.2020.9283073).

16. Karczmarek, P., Kiersztyn, A., & Pedrycz, W. (2020). Fuzzy Set-Based Isolation Forest. *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Glasgow, UK. DOI: [10.1109/FUZZ48607.2020.9177718](https://doi.org/10.1109/FUZZ48607.2020.9177718).

17. Liu, S., Ji, Z., & Wang, Y. (2020). Improving Anomaly Detection Fusion Method of Rotating Machinery Based on ANN and Isolation Forest. *2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, Chongqing, China. DOI: [10.1109/CVIDL51233.2020.00-23](https://doi.org/10.1109/CVIDL51233.2020.00-23).

18. Fitriyani, N. L., Syafrudin, M., Alfian, G., Fatwanto, A., Qolbiyani, S. L., & Rhee, J. (2020). Prediction Model for Type 2 Diabetes using Stacked Ensemble Classifiers. *2020 International Conference on Decision Aid Sciences and Application (DASA)*, Sakheer, Bahrain. DOI: [10.1109/DASA51403.2020.9317090](https://doi.org/10.1109/DASA51403.2020.9317090).

19. Park, J., Surabhi, V. R., Krishnamurthy, P., Garg, S., Karri, R., & Khorrami, F. (2020). Anomaly Detection in Embedded Systems Using Power and Memory Side Channels. *2020 IEEE European Test Symposium (ETS)*, Tallinn, Estonia. DOI: [10.1109/ETS48528.2020.9131596](https://doi.org/10.1109/ETS48528.2020.9131596).

20. Meneghetti, L., Terzi, M., Del Favero, S., Susto, G. A., & Cobelli, C. (2020). Data-Driven Anomaly Recognition for Unsupervised Model-Free Fault Detection in Artificial Pancreas. *IEEE Transactions on Control Systems Technology*, 1, 33-47. doi: [10.1109/TCST.2018.2885963](https://doi.org/10.1109/TCST.2018.2885963).

21. Dridi, A., Boucetta, C., Hammami, S. E., Afifi, H., & Mounbla, H. (2021). STAD: Spatio-Temporal Anomaly Detection Mechanism for Mobile Network Management. *IEEE Transactions on Network and Service Management*, 18(1), 894-906. doi: [10.1109/TNSM.2020.3048131](https://doi.org/10.1109/TNSM.2020.3048131).

22. Liu, C., Yong, S., Wang, X., & Zhang, X. (2020). A Multi-feature Anomaly Detection Method Based on AETA ULF Electromagnetic Disturbance Signal. In *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)* (pp. 1103-1108). Chongqing, China. doi: [10.1109/ITNEC48623.2020.9085032](https://doi.org/10.1109/ITNEC48623.2020.9085032).

23. Kromer-Edwards, C., Castanheira, M., & Oliveira, S. (2020). Year, Location, and Species Information In Predicting MIC Values with Beta-Lactamase Genes. In *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 1383-1390). Seoul, Korea (South). doi: [10.1109/BIBM49941.2020.9313331](https://doi.org/10.1109/BIBM49941.2020.9313331).

24. Fang, N., Fang, X., & Lu, K. (2022). Anomalous Behavior Detection Based on the Isolation Forest Model with Multiple Perspective Business Processes. *Electronics*, 11, 3640. doi: 10.3390/electronics11213640.

25. Chater, M., Borgi, A., Taieb, M. T., Sfar-Gandoura, K., & Landoulsi, M. I. (2022). Fuzzy Isolation Forest for Anomaly Detection. *Procedia Computer Science*, 207, 916-925. doi: 10.1016/j.procs.2022.09.147.

26. Choudhury, J., & Shi, C. (2022). Enhanced Performance of Finite Boundary Isolation Forest (FBIF) for Datasets with Standard Distribution Properties. In 2022 International Conference on Electrical, Computer and Energy Technologies (ICECET) (pp. 1-5). Prague, Czech Republic. doi: 10.1109/ICECET55527.2022.9873022.

27. Reddy, P. R., & Kumar, A. S. (2022). Credit Card Fraudulent Transactions Prediction Using Novel Sequential Transactions by Comparing Light Gradient Booster Algorithm Over Isolation Forest Algorithm. In 2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM) (pp. 563-567). Gautam Buddha Nagar, India. doi: 10.1109/ICIPTM54933.2022.9754211.

28. Marcelli, E., Barbariol, T., Savarino, V., Beghi, A., & Susto, G. A. (2022). A Revised Isolation Forest Procedure for Anomaly Detection with a High Number of Data Points. In 2022 IEEE 23rd Latin American Test Symposium (LATS) (pp. 1-5). Montevideo, Uruguay. doi: 10.1109/LATS57337.2022.9936964.

29. Fan, L., Ma, J., Tian, J., Li, T., & Wang, H. (2021). Comparative Study of Isolation Forest and LOF Algorithm in Anomaly Detection of Data Mining. In 2021 International Conference on Big Data, Artificial Intelligence and Risk Management (ICBAR) (pp. 1-5). Shanghai, China. doi: 10.1109/ICBAR55169.2021.00008.

30. Feng, F., Liu, Z., & Zhang, J. (2022). Detection of GPS Abnormal Data of Sanitation Vehicles Based on Isolation Forest Algorithm. In 2022 International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACA) (pp. 460-463). Zakopane, Poland. doi: 10.1109/ICDACA57211.2022.00097J.

31. Su and J. Li, "An Anomaly Detection Algorithm for Multi-dimensional Segmentation Plane Isolation Forest," 2022 IEEE 5th International Conference on Computer and Communication Engineering Technology (CCET), Beijing, China, 2022, pp. 89-93, doi: 10.1109/CCET55412.2022.9906369.

32. Yu, P., & Jia, L. (2022). Wind Power Data Cleaning Based on Autoencoder-Isolation Forest. In 2022 7th International Conference on Power and Renewable Energy (ICPRE) (pp. 803-808). Shanghai, China. doi: 10.1109/ICPRE55555.2022.9960342.

33. Kilinc, H. H. (2022). Anomaly Pattern Analysis Based on Machine Learning on Real Telecommunication Data. In 2022 7th International Conference on Computer Science and Engineering (UBMK) (pp. 43-48). Diyarbakir, Turkey. doi: 10.1109/UBMK55850.2022.9919564.

34. Huang, X., Ren, Y., He, Y., & Chen, Q. (2022). Malicious Models-based Federated Learning in Fog Computing Networks. In 2022 14th International Conference on Wireless Communications and Signal Processing (WCSP) (pp. 192-196). Nanjing, China. doi: 10.1109/WCS55476.2022.10039266.

35. Kerner, H. R., & Adler, J. B. (2022). Guiding Field Exploration on Earth and Mars with Outlier Detection. In IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium (pp. 5333-5336). Kuala Lumpur, Malaysia. doi: 10.1109/IGARSS46834.2022.9884366.

36. Zulkernan, I., Ahmed, N., Elmeligy, A., Abdelnaby, A., & Sheta, N. (2022). IoT Sensor Data Consistency using Deep Learning. In 2022 IEEE International Conference on Internet of Things and Intelligence Systems (IoTaIS) (pp. 198-203). BALI, Indonesia. doi: 10.1109/IoTaIS56727.2022.9975955.

37. Gajda, J., Kwiecień, J., & Chmiel, W. (2022). Machine learning methods for anomaly detection in computer networks. In 2022 26th International Conference on Methods and Models in Automation and Robotics (MMAR)

(pp. 276-281). Międzyzdroje, Poland. doi: 10.1109/MMAR55195.2022.9874341.

38. Malaek, S. M., & Alipour, E. (2022). Intelligent Flight-Data-Recorders; a Step Toward a New Generation of Learning Aircraft. In 2022 8th International Conference on Control, Decision and Information Technologies (CoDIT) (pp. 1545-1549). Istanbul, Turkey. doi: 10.1109/CoDIT55151.2022.9804136.

39. El Houda, Z. A., Hafid, A. S., & Khoukhi, L. (2021). A Novel Machine Learning Framework for Advanced Attack Detection using SDN. In 2021 IEEE Global Communications Conference (GLOBECOM) (pp. 1-6). Madrid, Spain. doi: 10.1109/GLOBECOM46510.2021.9685643.

40. Kafi, H. M., Miah, A. S. M., Shin, J., & Siddique, M. N. (2022). A Lite-Weight Clinical Features Based Chronic Kidney Disease Diagnosis System Using 1D Convolutional Neural Network. In 2022 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE) (pp. 1-5). Gazipur, Bangladesh. doi: 10.1109/ICAEEE54957.2022.9836398.

41. Anello, E., et al. (2022). Anomaly Detection for the Industrial Internet of Things: an Unsupervised Approach for Fast Root Cause Analysis. In 2022 IEEE Conference on Control Technology and Applications (CCTA) (pp. 1366-1371). Trieste, Italy. doi: 10.1109/CCTA49430.2022.9966158.

42. Yang, X., Zhuang, Y., Shi, M., Cao, X., Chen, D., & Tang, Y. (Year). SPiForest: An Anomaly Detecting Algorithm Using Space Partition Constructed by Probability Density-Based Inverse Sampling. *IEEE Transactions on Neural Networks and Learning Systems*. doi: 10.1109/TNNLS.2022.3223342.

43. Sugunaraj, N., Vrtis, J., Snyder, J., & Ranganathan, P. (Year). False Data Injection Modeling & Detection for Phasor Measurement Units. In 2022 North American Power Symposium (NAPS) (pp. 1-6). Salt Lake City, UT, USA. doi: 10.1109/NAPS56150.2022.10012137.

44. Premisha, P., Prasanth, S., Kanagarathnam, M., & Banujan, K. (Year). An Ensemble Machine Learning Approach for Stroke Prediction. In 2022 International Research Conference on Smart Computing and Systems Engineering (SCSE) (pp. 165-170). Colombo, Sri Lanka. doi: 10.1109/SCSE56529.2022.9905215.

45. Wang, G., Mao, X., Zhang, Q., & Lu, A. (Year). Fatigue Detection in Running with Inertial Measurement Unit and Machine Learning. In 2022 10th International Conference on Bioinformatics and Computational Biology (ICBCB) (pp. 85-90). Hangzhou, China. doi: 10.1109/ICBCB55259.2022.9802471.

46. Singh, R., & Deorari, R. (Year). Enhancing Collaborative Intrusion detection networks against insider attack using supervised learning technique. In 2022 IEEE 2nd Mysore Sub Section International Conference (Mysuru Con) (pp. 1-6). Mysuru, India. doi: 10.1109/MysuruCon55714.2022.9972599.

47. Yang, T., Cai, Z., Hou, B., & Zhou, T. (Year). 6Forest: An Ensemble Learning-based Approach to Target Generation for Internet-wide IPv6 Scanning. In IEEE INFOCOM 2022 -IEEE Conference on Computer Communications (pp. 1679-1688). London, United Kingdom. doi: 10.1109/INFOCOM48880.2022.9796925.

48. Mouret, F., Albughdadi, M., Duthoit, S., Kouamé, D., & Tourneret, J.-Y. (Year). Robust Estimation of Gaussian Mixture Models Using Anomaly Scores and Bayesian Information Criterion for Missing Value Imputation. In 2022 30th European Signal Processing Conference (EUSIPCO) (pp. 827-831). Belgrade, Serbia. doi: 10.23919/EUSIPCO55093.2022.9909815.

49. Mensi, A., Bicego, M. (Year). Enhanced anomaly scores for isolation forests. *Pattern Recognition*, Volume 120, December 2021, 108115. doi: doi.org/10.1016/j.patcog.2021.108115.

50. Badurowicz, M., Karczmarek, P., & Montusiewicz, J. (Year). Fuzzy Extensions of Isolation Forests for Road Anomaly Detection. In 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (pp. 1-6). Luxembourg, Luxembourg. doi: 10.1109/FUZZ45933.2021.9494469.

51. Ahmed, T., Shah, A., Kolla, M., & Yellasiri, R. (Year). Reduction of Alert Fatigue using Extended Isolation Forest. In 2021 International Conference on Forensics, Analytics, Big Data, Security (FABS) (pp. 1-5). Bengaluru, India. doi: 10.1109/FABS52071.2021.9702617.

52. Priyanto, C. Y., Hendry, & Purnomo, H. D. (Year). Combination of Isolation Forest and LSTM Autoencoder for Anomaly Detection. In 2021 2nd International Conference on Innovative and Creative Information Technology (ICITech) (pp. 35-38). Salatiga, Indonesia. doi: 10.1109/ICITech50181.2021.9590143.

53. Derse, C., Baghdadi, M., Hegazy, O., Sensoz, U., Gezer, H. N., & Nil, M. (Year). An Anomaly Detection Study on Automotive Sensor Data Time Series for Vehicle Applications. In 2021 Sixteenth International Conference on Ecological Vehicles and Renewable Energies (EVER) (pp. 1-5). Monte-Carlo, Monaco. doi: 10.1109/EVER52347.2021.9456629.

54. Damodaran, S., Padmanabhan, R., Maahin, R., & Gurugopinath, S. (Year). A Copula-Driven Unsupervised Learning Framework for Anomaly Detection with Multivariate Heterogeneous Data. In 2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP) (pp. 1-6). Gold Coast, Australia. doi: 10.1109/MLSP52302.2021.9596359.

55. Karthik, S., Supreetha, H. V., & Sandhya, S. (Year). Detection of Anomalies in Time Series Data. In 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS) (pp. 1-5). Bangalore, India. doi: 10.1109/CSITSS54238.2021.9683715.

56. Lopes, A. P., Parshionikar, S., Kale, A., Sharma, N., & Varghese, A. A. (Year). Comparative Analysis of Deep Learning Techniques For Credit Card Fraud Detection. In 2021 International Conference on Advances in Computing, Communication, and Control (ICAC3) (pp. 1-5). Mumbai, India. doi: 10.1109/ICAC353642.2021.9697205.

57. Komárek, T., Brabec, J., Škarda, Č., & Somol, P. (Year). Threat Hunting as a Similarity Search Problem on Multi-positive and Unlabeled Data. In 2021 IEEE International Conference on Big Data (Big Data) (pp. 2098-2103). Orlando, FL, USA. doi: 10.1109/BigData52589.2021.9671958.

58. McKinnon, C., Carroll, J., McDonald, A., Koukoura, S., & Plumley, C. (Year). Investigation of anomaly detection technique for wind turbine pitch systems. The 9th Renewable Power Generation Conference (RPG Dublin Online 2021) (pp. 277-282). Online Conference. doi: 10.1049/icp.2021.1401.

59. Witzig, P., Upenik, E., & Ebrahimi, T. (Year). Open-Set Person Re-Identification Through Error Resilient Recurring Gallery Building. In 2021 IEEE International Conference on Image Processing (ICIP) (pp. 245-249). Anchorage, AK, USA. doi: 10.1109/ICIP42928.2021.9506241.

60. Chen, R., Yang, Y., & Xia, M. (Year). Anomaly Detection of Sensor Data Based on 1D Depth Separable Dilated Convolution Neural Network. In 2021 International Conference on Networking, Communications and Information Technology (NetCIT) (pp. 240-244). Manchester, United Kingdom. doi: 10.1109/NetCIT54147.2021.00055.

61. Lu, G., Duan, C., Zhou, G., Ding, X., & Liu, Y. (Year). Privacy-Preserving Outlier Detection with High Efficiency over Distributed Datasets. IEEE INFOCOM 2021 - IEEE Conference on Computer Communications (pp. 1-10). Vancouver, BC, Canada. doi: 10.1109/INFOCOM42981.2021.9488710P.

62. Ntambu and S. A. Adeshina, "Machine Learning-Based Anomalies Detection in Cloud Virtual Machine Resource Usage," 2021 1st International Conference on Multidisciplinary Engineering and Applied Science (ICMEAS), Abuja, Nigeria, 2021, pp. 1-6, doi: 10.1109/ICMEAS52683.2021.9692308.

63. Kotsopoulos, T., "Fault Detection on Bearings and Rotating Machines based on Vibration Sensors Data," (2021) 2021 IEEE International Conference on Progress in Informatics and Computing (PIC), Shanghai, China, pp. 474-483. doi: 10.1109/PIC53636.2021.9686999.

64. Liu, W., "D2MIF: A Malicious Model Detection Mechanism for Federated-Learning - Empowered Artificial Intelligence of Things," (2023) IEEE Internet of Things Journal, 10(3), 2141-2151. doi: 10.1109/JIOT.2021.3081606.

65. Shrivastava, R., "Comparative study of boosting and bagging based methods for fault detection in a chemical process," (2021) 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, pp. 674-679. doi: 10.1109/ICAIS50930.2021.9395905.

66. Koch, P., Schekotihin, K., Jannach, D., Hofer, B., & Wotawa, F., "Metric-Based Fault Prediction for Spreadsheets," (2021) IEEE Transactions on Software Engineering, 47(10), 2195-2207. doi: 10.1109/TSE.2019.2944604.

67. Pathak, A. K., Saguna, S., Mitra, K., & Åhlund, C., "Anomaly Detection using Machine Learning to Discover Sensor Tampering in IoT Systems," (2021) ICC 2021 - IEEE International Conference on Communications, Montreal, QC, Canada, pp. 1-6. doi: 10.1109/ICC42927.2021.9500825.

68. Alghanmi, A., Yunusa-Kaltungo, A., & Edwards, R., "A Comparative Study of Faults Detection Techniques on HVAC Systems," (2021) 2021 IEEE PES/IAS PowerAfrica, Nairobi, Kenya, pp. 1-5. doi: 10.1109/PowerAfrica52236.2021.9543158.

69. Antony, L., "A Comprehensive Unsupervised Framework for Chronic Kidney Disease Prediction," (2021) IEEE Access, 9, 126481-126501. doi: 10.1109/ACCESS.2021.3109168.

70. Mansour, R. F., Amraoui, A. E., Nouaouri, I., Díaz, V. G., Gupta, D., & Kumar, S., "Artificial Intelligence and Internet of Things Enabled Disease Diagnosis Model for Smart Healthcare Systems," (2021) IEEE Access, 9, 45137-45146. doi: 10.1109/ACCESS.2021.3066365.

71. Simmini, F., Rampazzo, M., Peterle, F., Susto, G. A., & Beghi, A., "A Self-Tuning KPCA-Based Approach to Fault Detection in Chiller Systems," (2022) IEEE Transactions on Control Systems Technology, 30 (4), 1359-1374, July. doi: 10.1109/TCST.2021.3107200.

72. Chen, J., Zhang, J., Qian, R., Yuan, Junfeng, & Ren, Y. (2023). An Anomaly Detection Method for Wireless Sensor Networks Based on the Improved Isolation Forest. *Applied Sciences*, 13, 702. doi: 10.3390/app13020702.

73. Almansoori, M., & Telek, M. (2023). Anomaly Detection using a combination of Autoencoder and Isolation Forest. In 1st Workshop on Intelligent Infocommunication Networks, Systems and Services (WI2NS2) (pp. 25-30). Budapest. doi: 10.3311/WINS2023-005.

74. Utkin, L., Ageev, A., Konstantinov, A., & Muliukha, V. (2023). Improved Anomaly Detection by Using the Attention-Based Isolation Forest. *Algorithms*, 16, 19. <https://doi.org/10.3390/a16010019>

75. Kabir, S., Shufian, A., & Zishan, M. S. R. (2023). Isolation Forest Based Anomaly Detection and Fault Localization for Solar PV System. 2023 3rd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), Dhaka, Bangladesh, pp. 341-345. doi: 10.1109/ICREST57604.2023.10070033.

76. Baviskar, P. V., Singh, G., & Patil, V. N. (2023). Design of Machine Learning-Based Malware Detection Methodologies in the Internet of Things Environment. 2023 International Conference for Advancement in Technology (ICONAT), Goa, India, pp. 1-6. doi: 10.1109/ICONAT57137.2023.10080517.

77. Himeur, Y., Fadli, F., & Amira, A. (2022). A Two-Stage Energy Anomaly Detection for Edge-based Building Internet of Things (BIoT) Applications. 2022 5th International Conference on Signal Processing and Information Security (ICSPIS), Dubai, United Arab Emirates, pp. 180-185. doi: 10.1109/ICSPIS57063.2022.10002641.

78. Anand, N., & Saifulla, M. A. (2023). An efficient IDS for slow rate HTTP/2.0 DoS Attacks using one class classification. 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), Lonavla, India, pp. 1-9. doi: 10.1109/I2CT57861.2023.10126162.

79. Bannur, C., Bhat, C., Singh, K., Kulkarni, S. A., & Doddamani, M. (2023). PAACDA: Comprehensive Data Corruption Detection Algorithm. *IEEE Access*, 11, 24908-24934. doi:10.1109/ACCESS.2023.3253022.

80. Buschjäger, S., Honysz, P. -J., & Morik, K. (2020). Generalized Isolation Forest: Some Theory and More Applications Extended Abstract. 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), Sydney, NSW, Australia, pp. 793-794. doi: 10.1109/DSAA49011.2020.00120.

81. Xu, D., Wang, Y., Meng, Y., & Zhang, Z. (2017). An Improved Data Anomaly Detection Method Based on Isolation Forest. 2017 10th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, pp. 287-291. doi: 10.1109/ISCID.2017.202.

82. Wielgosz, M., Skoczen, A., & Wiatr, K. (2018). Looking for a Correct Solution of Anomaly Detection in the LHC Machine Protection System. 2018 International Conference on Signals and Electronic Systems (ICSES), Kraków, Poland, pp. 257-262. doi: 10.1109/ICSES.2018.8507291Y.

83. Qin, Y., & Lou, Y. (2019). Hydrological Time Series Anomaly Pattern Detection based on Isolation Forest. 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chengdu, China, pp. 1706-1710. doi: 10.1109/ITNEC.2019.8729405.

84. Luo, S., Luan, L., Cui, Y., Chai, X., Wang, Z., & Kong, Y. (2019). An Attribute Associated Isolation Forest Algorithm for Detecting Anomalous Electro-data. 2019 Chinese Control Conference (CCC), Guangzhou, China, pp. 3788-3792. doi: 10.23919/ChiCC.2019.8866495.

85. Ding, Z., Mo, Y., & Pan, Z. (2019). A Novel Software Defect Prediction Method Based on Isolation Forest. 2019 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2 MSE), Zhangjiajie, China, pp. 882-887. doi: 10.1109/QR2MSE46217.2019.9021215.

86. Shriram, S., & Sivasankar, E. (2019). Anomaly Detection on Shuttle data using Unsupervised Learning Techniques. 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, pp. 221-225. doi: 10.1109/ICCIKE47802.2019.9004325.

87. Baloglu, U. B., Alatas, B., & Bingol, H. (2019). Assessment of Supervised Learning Algorithms for Irony Detection in Online Social Media. 2019 1st International Informatics and Software Engineering Conference (UBMYK), Ankara, Turkey, pp. 1-5. doi: 10.1109/UBMYK48245.2019.8965580.

88. Shu, H., Zhao, X., Luo, H., & Li, C. (2019). Research on Stacking-Based Integrated Algorithm of Anomaly Detection in Production Process. 2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), Shenzhen, China, pp. 85-90. doi: 10.1109/HPBDIS.2019.8735447.

89. Iheme, L.O., & Ozan, Ş. (2019). Feature Selection for Anomaly Detection in Call Center Data. 2019 11th International Conference on Electrical and Electronics Engineering (ELECO), Bursa, Turkey, pp. 926-929. doi: 10.23919/ELECO47770.2019.8990454.

90. Komárek, T., Brabec, J., Škarda, Č., & Somol, P. (2021). Threat Hunting as a Similarity Search Problem on Multi-positive and Unlabeled Data. 2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, pp. 2098-2103. doi: 10.1109/BigData52589.2021.9671958.

91. Liu, W., He, J., Han, S., Cai, F., Yang, Z., & Zhu, N. (Year). A Method for the Detection of Fake Reviews Based on Temporal Features of Reviews and Comments.

92. Lv, Y., Cui, Y., Zhang, X., Cai, M., Gu, X., & Xiong, Z. (2019). A New Outlier Detection Method Based on Machine Learning. 2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Chongqing, China, pp. 1-7. doi: 10.1109/ICSI DP47821.2019.9173217.

93. Sepulvene, L. H. M., et al. (2019). Analysis of Machine Learning Techniques in Fault Diagnosis of Vehicle Fleet Tracking Modules. 2019 8th Brazilian Conference on Intelligent Systems (BRACIS), Salvador, Brazil, pp. 759-764. doi: 10.1109/BRACIS.2019.00136.

94. Portela, F. G., Mendoza, F. A., & Benavides, L. C. (2019). Evaluation of the performance of

supervised and unsupervised Machine learning techniques for intrusion detection. 2019 IEEE International Conference on Applied Science and Advanced Technology (iCASAT), Queretaro, Mexico, pp. 1-8. doi: 10.1109/iCASAT48251.2019.9069538

This page is intentionally left blank



Scan to know paper details and
author's profile

Frequency Compensation Design of Three-Stage CMOS OTA Amplifier for RTD Applications

Tayebeh Asiyabi & Seid Jafar Hosseinipouya

Islamic Azad University

ABSTRACT

In this paper, a new three-stage CMOS operational transconductance amplifier (OTA) for RTD applications is proposed. The presented structure eliminates the feedforward path and strengthens the feedback path of compensation network together. In addition, the presented circuit uses only a small compensation capacitor on the order of 1 pF, which alleviates the chip area. The circuit is simulated using a 0.18- μ m N-well CMOS technology process showing high performance and power efficiency regarding to figure of merit (FOM) factors defined in the literature. The frequency response of the proposed amplifier shows that the DC gain is about 120 dB and bandwidth is 18.8 MHz with a phase margin of 88° and power dissipation of 544 μ W.

Keywords: efficiency coefficient, differential block frequency compensator, RTD, amplifier, low power.

Classification: LCC Code: TK7871

Language: English



Great Britain
Journals Press

LJP Copyright ID: 392935
Print ISSN: 2631-8474
Online ISSN: 2631-8482

London Journal of Engineering Research

Volume 24 | Issue 2 | Compilation 1.0



Frequency Compensation Design of Three-Stage CMOS OTA Amplifier for RTD Applications

Tayebeh Asiyabi^a & Seid Jafar Hosseinipouya^a

ABSTRACT

In this paper, a new three-stage CMOS operational transconductance amplifier (OTA) for RTD applications is proposed. The presented structure eliminates the feedforward path and strengthens the feedback path of compensation network together. In addition, the presented circuit uses only a small compensation capacitor on the order of 1 pF, which alleviates the chip area. The circuit is simulated using a 0.18- μ m N-well CMOS technology process showing high performance and power efficiency regarding to figure of merit (FOM) factors defined in the literature. The frequency response of the proposed amplifier shows that the DC gain is about 120 dB and bandwidth is 18.8 MHz with a phase margin of 88° and power dissipation of 544 μ W.

Keywords: efficiency coefficient, differential block frequency compensator, RTD, amplifier, low power.

Author a: Dez Dam & Power Plant Operation & Generation CO, Iran.

a: Department of Electrical Engineering, Izeh Branch, Islamic Azad University, Izeh, Iran.

I. INTRODUCTION

Process parameters in the industry or power plant have to be continuously monitored and controlled. Temperature is one of the most important parameters that must be controlled by the RTD, which is used to adjust the signal amplitude from the amplifier at the output of the Wheatstone bridge. Additionally, implementing an RTD of platinum is too expensive for the most applications. As a result, only small RTD elements are made out of platinum. To use the resistance value of RTD element, this resistance must be measured to the device. For this purpose, a copper

wire with an insulating coating is used to connect the RTD element to the measuring device. The formation of platinum, copper also has electrical resistance. The resistance along the copper wire can affect the resistance value measured by the device connected to the RTD. In two-wire RTDs, there is no resistance to the copper wire to reduce the resistance and be measured against the copper wire. Consequently, two-wire RTDs are commonly used for approximate temperature measurement.

Three-wire RTDs are mostly used in industrial applications. These RTDs usually use the Wheatstone bridge measurement circuit to compensate the resistance of the copper wire. The response time of the temperature element is tested in all power plants and other industries. The transient response during the operation of the sensor is particular important to meet the requirements of the technical specifications and to enhance the signal range. An RTD is a transducer that converts temperature into an electrical signal.

If high energy is needed to move the drive unit, the controller command is amplified by the amplifier and then sent to the drive. Due to scaling down of channel length and power supply voltage, standard configurations such as cascode are not useful in OTA design based on CMOS technology, while multi-stage amplifiers can be used [11-13]. However, by increasing the DC gain, such arrangements show low phase margin (PM) and high impedance nodes, which affect the frequency stability of the amplifier requiring the invention and use of techniques frequency compensation [1]. In the case of a single-stage amplifier, one of the conventional frequency compensation methods that is specifically used is RNMC in which the first two stages must be negative. This technique is remarkable due to its design and without significant energy consumption. However, this method adds a zero

frequency offset in the right half (RHP) to the amplifier, which causes stability problems [2-5]. During the last year, several techniques such as current and voltage buffer [6], forward paths [7-9], AFFC [10] and ACBC [11] have been invented to solve that problem mentioned earlier, but they are not bound to achieve stability.

Moreover, the condition and improvement of gain bandwidth (GBW) did not succeed at the same time. This is due to the presence of RHP in the conversion function. In general, NMC and RNMC are two common methods for frequency compensation of three-layer OTAs. These two methods use two capacitors as a compensation network. In this paper, we present a new compensator method that uses a differential feedback path to remove the leading path and increase the frequency response to PM and GBW.

The frequency response analysis of the proposed circuit shows that the values of the compensating network elements can be controlled and adjusted for the phase margin and bandwidth. A detailed mathematical analysis of the proposed method is

presented in the second part, along with operational theories, stability conditions, and design issues. The simulation results of the proposed topology and comparison among existing techniques are given in the third section. In the end, the conclusions will be presented in the last section.

II. RECOMMENDED TECHNIQUE

Figure 1 shows the structure of the proposed method. The output resistance and parasitic capacitor of each floor are specified by R_{i-3} and C_{i-3} respectively. C_L is also a charge capacitor. In this structure, a differential feedback layer is used to achieve two goals. The first differential feedback layer forms the compensation network, and in the second layer, the output signal is driven from this layer, so it can reach the maximum DC gain. These two issues make the proposed method very effective to reach the desired GBW and PM. The proposed circuit is shown in Figure 1, where the compensation capacitor that has the Miller effect must be in the loop with a negative gain.

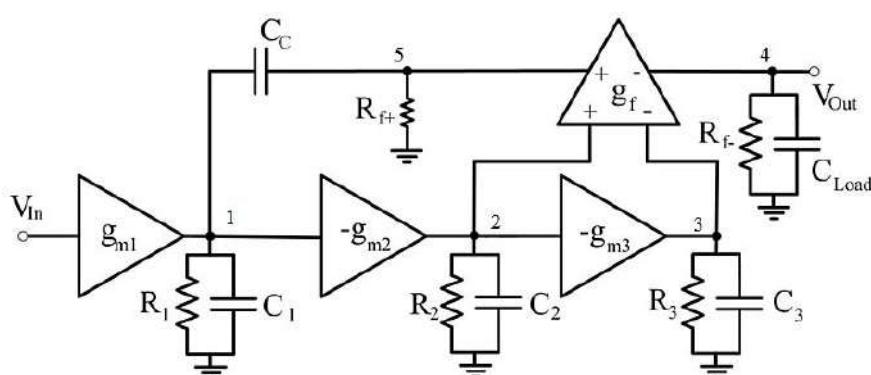


Figure 1: Structure of the proposed technique

According to the proposed structure in Figure 1, we can obtain the transformation function indicated in (1).

$$H_1 = \frac{(C_C g_{m1} R_1 R_f) S - g_{m1} g_{m2} g_f R_1 R_2 R_f - g_{m1} g_{m2} g_{m3} g_f R_1 R_2 R_3 R_f}{C_C C_L R_1 R_f S^2 + (C_C R_1 + C_C R_f + C_L R_f + C_C g_{m2} g_{m3} g_f R_1 R_2 R_3 R_f) S + 1} \quad (1)$$

Note that the parasitic indices show that the capacitor and resistance are placed between the corresponding nodes and the ground. V_i is the voltage of the node, C_i and R_i indicate the total capacity and total resistance of the i th node,

respectively. In addition, g_{mi} shows the transmission conductivity of i -th layer.

III. CIRCUIT DESIGN

Figure 2 shows the transistor surface of the proposed circuit. The circuit includes a

differential input stage, while the second and third stages include common source amplifiers.

The fourth floor is another differential amplifier that forms the output floor.

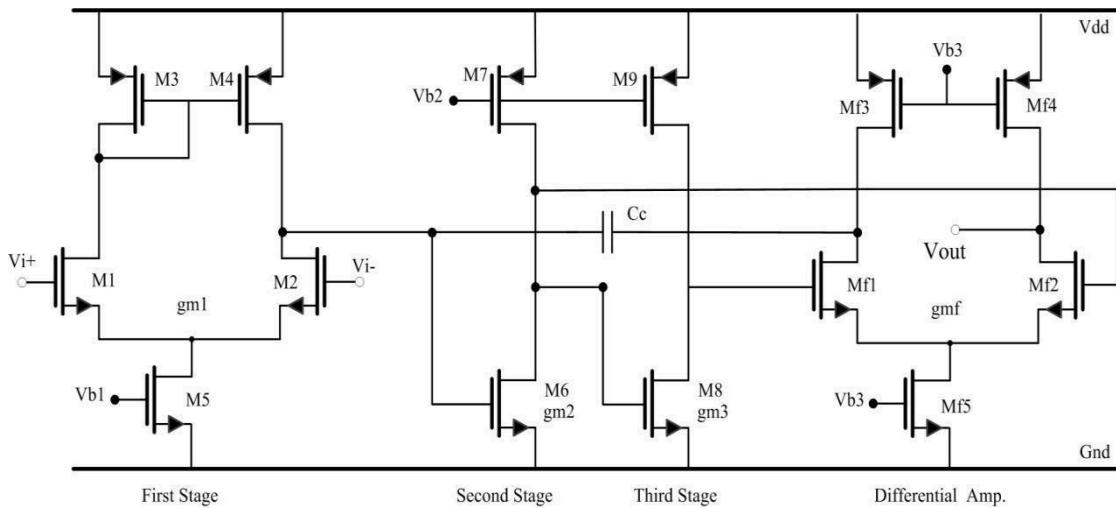


Figure 2: Transistor Level of the Proposed Technique

The input differential layer consists of transistors M1-M5 with active load (M3, M4), while the second layer is a common source amplifier (M6) and current mirror M7, which form the transfer conduction gm_1 . The third layer (M8, M9) is the same as the previous layer, which form the transmission conduction gm_2 . The output stage (Mf1-Mf5), which is driven by the third stage, is with gm_f . The compensating capacitor (CC) from the output differential amplifier to the output of the first stage is located in the loop with negative feedback. Assuming C_i , C_c , C_l and ignoring the dynamics of high frequencies and considering GBW is much larger compared to the poles of equation (1) and also considering that the gain of each floor is greater than unity, the conversion function The open loop of the amplifier (H_2) will be simplified as equation (2).

$$H_2 = \frac{(C_c g_{m1} R_1 R_f) S - g_{m1} g_{m2} g_{m3} g_f R_1 R_2 R_3 R_f}{C_c C_L R_1 R_f S^2 + (C_c g_{m2} g_{m3} g_f R_1 R_2 R_3 R_f) S + 1} \quad (2)$$

As a result, the dominant and non-dominant poles as well as the zero of the transformation function are presented in (3), (4) and (5), respectively.

$$P_1 = \frac{1}{C_c g_{m2} g_{m3} g_f R_1 R_2 R_3 R_f} \quad (3)$$

$$P_2 = \frac{g_{m2} g_{m3} g_f R_2 R_3}{C_L} \quad (4)$$

$$Z = \frac{g_{m2} g_{m3} g_f R_2 R_3}{C_c} \quad (5)$$

Additionally, closed loop transfer function with unit feedback gain is calculated as (6).

$$H_{CL3} = \frac{(C_c g_{m1} R_1 R_f) S - g_{m1} g_{m2} g_f R_1 R_2 R_f - g_{m1} g_{m2} g_{m3} g_f R_1 R_2 R_3 R_f}{C_c C_L R_1 R_f S^2 + (C_c g_{m2} g_{m3} g_f R_1 R_2 R_3 R_f) S + g_{m1} g_{m2} g_{m3} g_f R_1 R_2 R_3 R_f} \quad (6)$$

Figure 3 shows the frequency response of the proposed circuit. According to this figure, the simplified transformation function (5) follows the exact transformation function with appropriate accuracy. In fact, after removing the non-dominant coefficients, the simplified transformation function is obtained. These non-dominant coefficients usually appear due to

the effects of parasitic capacitors. Therefore, another advantage of the presented method is that this compensation method is independent of parasitic capacitors and the estimated conversion function accurately leads to the improvement of the frequency response.

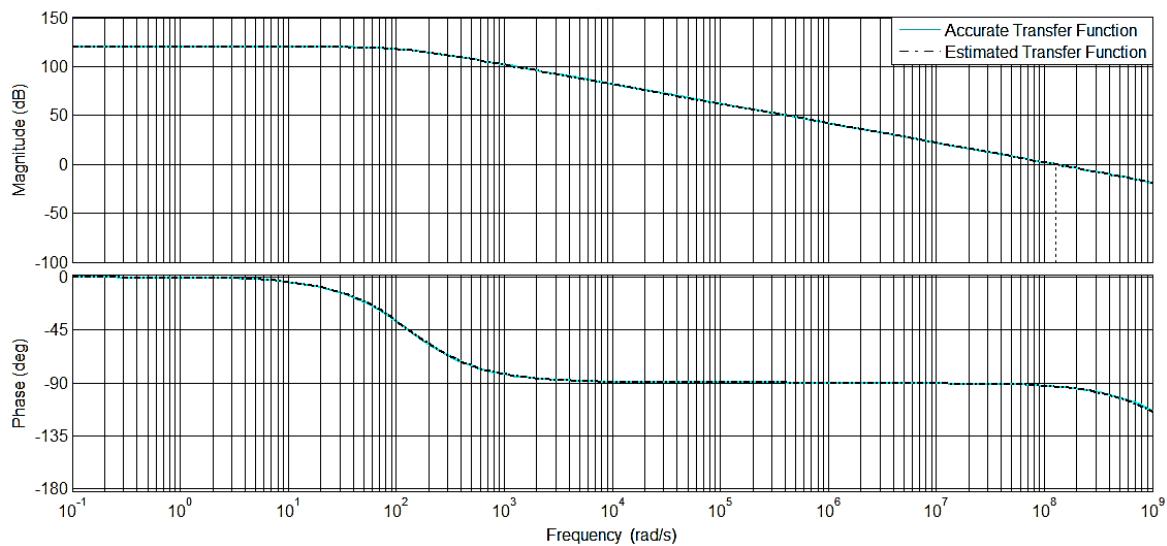


Figure 3: Frequency Response of the Proposed Circuit

IV. IMPLEMENTING OF THE PROPOSED CIRCUIT IN RTD APPLICATIONS

Most of the instruments and equipment used to control an industrial process are usually installed in the control room at a distance from the process; on the other hand, the measuring element is usually installed on the process or at a distance

close to it. Therefore, the signal resulting from the measured quantity must be reliably sent to the control room, this is done by the transmitter. Figure 4 shows the transmitter connection to the monitor of the measuring device in various types of industrial equipment or process control.

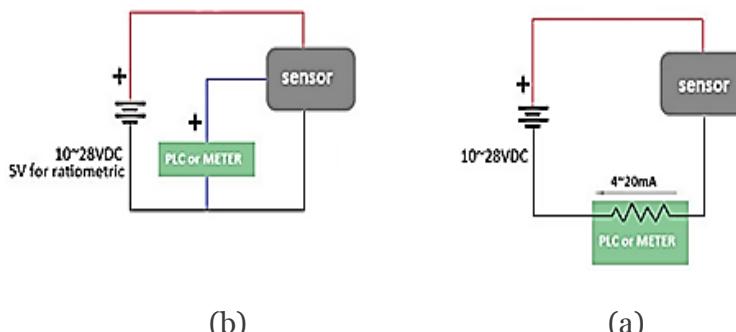


Figure 4: Connecting the Transmitter to the PLC: a) Two-Wire Circuit, b) Three-Wire Circuit

Transducer output current is usually between 4 and 20 mA. By adding a resistive bridge to the inputs of the operational amplifier, the resulting circuit will be able to both add and subtract the voltage applied to the corresponding inputs. Figure 5 shows the completed proposed circuit.

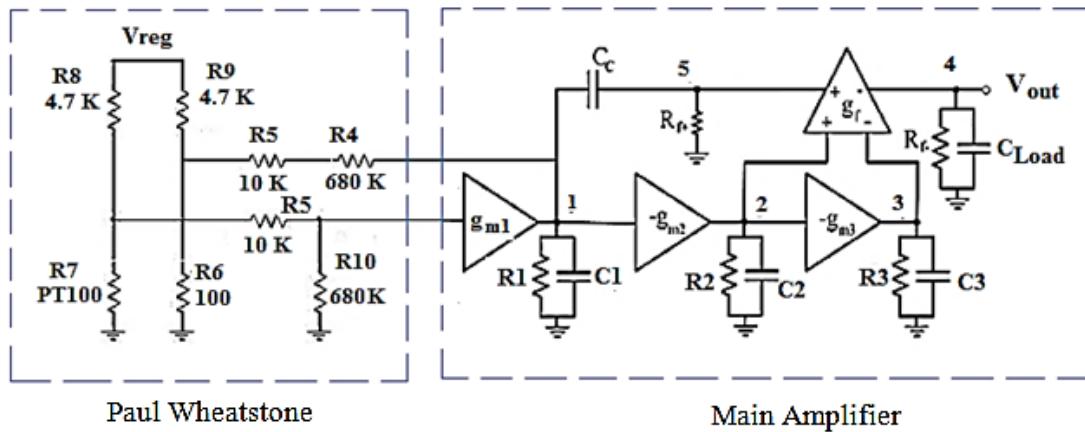


Figure 5: Proposed Test Circuit

This test is performed in power plants using a heating current of about 40 mA. To provide the necessary current for the test, the power supply is adjusted to the Wheatstone bridge, so the high current should be between 30 and 50 mA. The RTD is only stable up to 30mA in one process. On the other hand, if the RTD is in a process that has large temperature fluctuations, a current higher than 50 mA is required to improve the

signal-to-noise ratio (S/N). For this reason, an amplifier is used at the output of the Wheatstone bridge to enhance the signal amplitude. As a result, the output voltage of the Wheatstone bridge (V) changes linearly according to the RTD resistance (8R) during the test. Figure 6 shows the test response indicating that the transient response can be improved by using the proposed compensation technique.

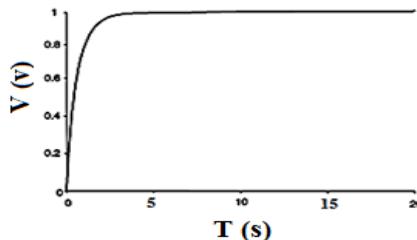


Figure 6: Output Time Response

To compare the results with other compensation methods, we used the following two definitions of FOM (figure of merit) coefficients defined in [3].

$$FOM_S = \frac{\omega_{GBW} \times C_L}{Power} \left(\frac{MH_z \cdot \mu F}{mW} \right) \quad (6)$$

$$FOM_L = \frac{SR \times C_L}{Power} \left(\frac{V}{\mu s} \cdot \frac{PF}{mW} \right) \quad (7)$$

The FOM_S and FOM_L indicate improvement in small signal and large signal characteristics, respectively. Since in (6) and (7) the compensation capacitors are not included, we define a new FOM to consider their effects. Because the most of the chip area is consumed by

passive elements [19] and the compensation capacitors are only passive elements in the OTA (through in many practical applications the output load is outside the chip), the new FOM is shown in (8). The performance comparison OTAs is summarized in Table 1.

$$FOM_{NEW} = \frac{\omega_{GBW} \times C_L^2}{Power \times C_{ctot}} \left(\frac{MH_z \cdot \mu F}{mW} \right) \quad (8)$$

Table 1: Performance Comparison of Different Topologies

| | DC Gain ((dB) | Load (pF) | Power (μW) | GBW (MHz) | Compensation Capacitor ((pF) | Slew Rate (V/μS) | PM (°) | FOM _s | FOM _L | FOM _{NEW} |
|------------------|---------------|------------|------------|-------------|------------------------------|------------------|-----------|------------------|------------------|--------------------|
| [NMC [13] | 100 | 100 | 345 | 0.22 | 110 | 0.25 | 68.3 | 0.06 | 0.07 | 0.054 |
| NMCNR [[2] | 100 | 100 | 345 | 0.32 | 78 | 0.30 | 70.5 | 0.09 | 0.08 | 0.115 |
| [DPZC [2] | 100 | 100 | 345 | 0.40 | 49.5 | 0.39 | 90.5 | 0.11 | 0.11 | 0.222 |
| [NGCC [12] | 100< | 100 | 365 | 0.25 | 96 | 0.33 | 69.1 | 0.06 | 0.09 | 0.062 |
| [NMCF [2] | 102 | 100 | 345 | 0.67 | 34 | 0.57 | 69.6 | 0.19 | 0.16 | 0.558 |
| NMCFNR [[2] | 100< | 100 | 345 | 0.80 | 28.7 | 0.63 | 72.1 | 0.23 | 0.18 | 0.801 |
| [DFCFC [10] | 100< | 100 | 372 | 0.96 | 35 | 0.80 | 66.6 | 0.25 | 0.23 | 0.714 |
| [AFFC [11] | 100< | 100 | 424 | 2.60 | 15 | 12 | 70.4 | 0.61 | 2.83 | 4.066 |
| [ACBC [9] | 100< | 100 | 365 | 2.06 | 18 | 1.22 | 69.6 | 0.56 | 0.33 | 3.111 |
| This work | 120 | 100 | 515 | 18.8 | 1.1 | 4.6 | 88 | 3.65 | 0.89 | 331 |

V. CONCLUSION

In this article, a new network compensation method has been presented to enhance frequency response of three-stage amplifiers for RTD applications. The proposed structure eliminates feedforward path and simultaneously strengthens the feedback path of compensation network while the Wheatstone bridge output voltage changes linearly according to RTD alters. The results prove the correctness of circuit theory analysis having superior performance with respect to small signal and large signal characteristics.

REFERENCE

1. G. Palumbo and S. Pennisi, Feedback Amplifiers, "Theory and Design. Boston", MA: Kluwer, 2002.
2. Grasso, A. D., Palumbo, G. and Pennisi, S., "Analytical comparison of frequency compensation techniques in three-stage amplifiers". Int. J. Circ. Theor. Appl., vol. 36, pp. 53–80, 2008.
3. Grasso, A. D., Marano, D., Palumbo, G. and Pennisi, S., "Analytical comparison of reversed nested Miller frequency compensation techniques". Int. J. Circ. Theor. Appl., vol. 38: 709–737, 2010
4. R. Mita, G. Palumbo, S. Pennisi, "Design guidelines for reversed nested Miller compensation in three-stage amplifiers", IEEE Trans. Circuits and Systems-II, Vol. 50, pp. 227-233, 2003.
5. K.P. Ho, C.F. Chan, C.S. Choy, K.P. Pun, "Reversed nested Miller compensation with voltage buffer and nulling resistor", IEEE J. Solid-State Circuits, vol. 38, pp. 1735-1738, 2003.
6. A.D. Grasso, G. Palumbo, S. Pennisi, "Improved reversed nested miller frequency compensation technique with voltage buffer and resistor", IEEE Trans. Circuits and Systems II: Express Briefs, Vol. 54, pp. 382-386, 2007.
7. A.D. Grasso, G. Palumbo, S. Pennisi, "Advances in reversed nested Miller compensation", IEEE Trans. Circuits Systems I: Regular Papers, Vol. 54, pp. 1459-1470, 2007.
8. F. Zu, S. Yan, J. Hu, E. Sanchez-Sinencio, "Feedforward reversed nested Miller compensation techniques for three-stage amplifiers", IEEE Int. Symp. on Circuits and Systems, Kobe, Japan, Vol. 1, pp. 2575-2578, 2005.
9. Peng X, Sansen W, "AC boosting compensation scheme for low-power multi-stage amplifiers", IEEE J Solid-State Circuits 2004;vol. 39, no. 11, pp. 2074–2079.

10. Leung AKN, Mok PK, Ki WH, Sin JK, "Damping-factor-control frequency compensation technique for low-voltage low-power large capacitive load applications. In: Solid-State Circuits Conference, 1999". Digest of Technical Papers. ISSCC.1999 IEEE International. IEEE, pp. 158–9, 2010.
11. Lee H, Mok PK, "Active-feedback frequency-compensation technique for low-power multistage amplifiers". IEEE J Solid-State Circuits 2003, vol. 38, no. 3, pp. 511–20.
12. You, F., Embabi, S. H., & Sanchez-Sinencio, E.. "Multistage amplifier topologies with nested G m-C compensation. *Solid-State Circuits*", *IEEE Journal of*, vol. 32, no. 12, pp. 2000-2011, 1997.
13. Ka Nang Leung, Mok PKT, "Nested Miller compensation in low-power CMOS design", IEEE Trans Circuits Syst II: Analog Digit Signal Process, vol. 48, no. 4, pp. 388–94, 2001.

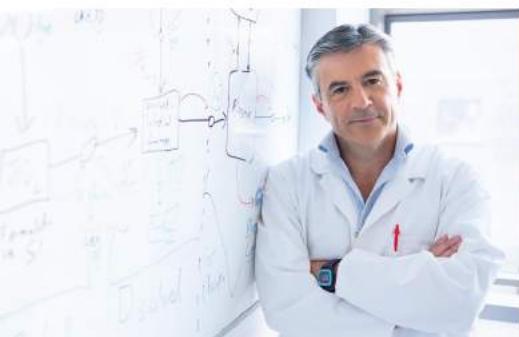
Great Britain Journal Press Membership

For Authors, subscribers, Boards and organizations



Great Britain Journals Press membership is an elite community of scholars, researchers, scientists, professionals and institutions associated with all the major disciplines. Great Britain memberships are for individuals, research institutions, and universities. Authors, subscribers, Editorial Board members, Advisory Board members, and organizations are all part of member network.

Read more and apply for membership here:
<https://journalspress.com/journals/membership>



For Authors



For Institutions



For Subscribers

Author Membership provide access to scientific innovation, next generation tools, access to conferences/seminars/symposiums/webinars, networking opportunities, and privileged benefits. Authors may submit research manuscript or paper without being an existing member of GBJP. Once a non-member author submits a research paper he/she becomes a part of "Provisional Author Membership".

Society flourish when two institutions Come together." Organizations, research institutes, and universities can join GBJP Subscription membership or privileged "Fellow Membership" membership facilitating researchers to publish their work with us, become peer reviewers and join us on Advisory Board.

Subscribe to distinguished STM (scientific, technical, and medical) publisher. Subscription membership is available for individuals universities and institutions (print & online). Subscribers can access journals from our libraries, published in different formats like Printed Hardcopy, Interactive PDFs, EPUBs, eBooks, indexable documents and the author managed dynamic live web page articles, LaTeX, PDFs etc.



GO GREEN AND HELP
SAVE THE ENVIRONMENT

JOURNAL AVAILABLE IN

PRINTED VERSION, INTERACTIVE PDFS, EPUBS, EBOOKS, INDEXABLE DOCUMENTS AND THE AUTHOR MANAGED DYNAMIC LIVE WEB PAGE ARTICLES, LATEX, PDFS, RESTRUCTURED TEXT, TEXTILE, HTML, DOCBOOK, MEDIAWIKI MARKUP, TWIKI MARKUP, OPML, EMACS ORG-MODE & OTHER



support@journalspress.com
www.journalspress.com



*THIS JOURNAL SUPPORT AUGMENTED REALITY APPS AND SOFTWARES