online issn: 1307-6892

irc 2022
XVI. international research conference
proceedings

open science index 16 2022

july 19-20, 2022 helsinki finland
international scholarly and scientific research & innovation

Open Science

Open Science Philosophy

Open science encompasses unrestricted access to scientific research articles, access to data from public research, and collaborative research enabled by information and communication technology tools, models, and incentives. Broadening access to scientific research publications and data is at the heart of open science. The objective of open science is to make research outputs and its potential benefits available to the entire world and in the hands of as many as possible:

- Open science promotes a more accurate verification of scientific research results. Scientific inquiry and discovery can be sped up by combining the tools of science and information technologies. Open science will benefit society and researchers by providing faster, easier, and more efficient availability of research outputs.
- Open science reduces duplication in collecting, creating, transferring, and re-using scientific material.
- Open science increases productivity in an era of tight budgets.
- Open science results in great innovation potential and increased consumer choice from public research.
- Open science promotes public trust in science. Greater citizen engagement leads to active participation in scientific experiments and data collection.

Open Science Index

The Open Science Index (OSI) currently provides access to over thirty thousand full-text journal articles and is working with member and non-member organizations to review policies to promote and assess open science. As part of the open science philosophy, and by making open science a reality; OSI is conducting an assessment of the impact of open science principles and restructuring the guidelines for access to scientific research. As digitalization continues to accelerate science, Open science and big data hold enormous promise and present new challenges for policymakers, scientific institutions, and individual researchers.

OSI is helping the global scientific research community discover, evaluate, and access high-quality research output. Renowned for its editorially curated and refereed collection of the highest-quality publications, OSI has always been and will remain free-of-charge.

OSI provides an efficient and thorough discovery process to the open science research database and provides links and free access to full-text articles. There are 50 open access journal categories that are curated and refereed by international scientific committees, the in-house editorial team, and trusted partners. Since its inception in 2007, OSI has made more than thirty thousand peer-reviewed open access full-text journal articles (PDF versions) freely available online without cost, barriers, or restrictions.

Open Science Access

With the Open Science Index, researchers can discover and access trusted peer-reviewed open access full-text scientific research articles with confidence. OSI helps researchers find appropriate non-profit open access journals to publish their work.

OSI gives one-click access to online full-text PDFs and expands the reach to global society by giving users free access from anywhere around the globe. Through cutting-edge open science collaboration, in an innovative public partnership, the non-profit OSI is devoted to making science open and reusable.

To learn more, visit online at waset.org

Open Science

Open Society

An open society allows individuals to change their roles and to benefit from corresponding changes in status. Open science depends to a greater or lesser extent on digital technologies and innovations in structural processes by an open society. When realized, open science research and innovation can create investment opportunities for new and better products and services and therefore increase competitiveness and employment. Open science research and innovation is a key component of thematic open science priorities. Central to the open science digital infrastructure is enabling industry to benefit from digital technology and to underpin scientific advances through the development of an open society. Open science research and innovation can also contribute to society as a global actor because scientific relations can flourish even where global relations are strained. Open science has a critical role across many areas of decision making in providing evidence that helps understand the risks and benefits of different open science choices. Digital technology is making the conduct of open science and innovation more collaborative, more global, and more open to global citizens. Open society must embrace these changes and reinforce its position as the leading power for science, for new ideas, and for investing sustainably in the future.

It is apparent in open society that the way science works is fundamentally changing, and an equally significant transformation is taking place in how organizations and societies innovate. The advent of digital technology is making research and innovation more open, collaborative, and global. These exchanges are leading open society to develop open science and to set goals for research and innovation priority. Open science goals are materializing in the development of scientific research and innovation platforms and greater acceptance of scientific data generated by open science research. Open science research and innovation do not need help from open society to come up with great ideas, but the level of success ideas ultimately reach is undoubtedly influenced by regulation, financing, public support, and market access. Open society is playing a crucial role in improving all these success factors.

Open Science

Open science represents a new approach to the scientific process based on cooperative work and new ways of diffusing knowledge by using digital technologies and collaborative tools. These innovations capture a systemic change to the way science and research have been carried out for the last fifty years. Science is shifting from the standard practice of publishing research results in scientific publications after the research and reviews are completed. The shift is towards sharing and using all available knowledge at an earlier stage in the research process. Open science is to science what digital technology is to social and economic transactions: allowing end users to be producers of ideas, relations, and services and in doing so, enabling new working models, new social relationships and leading to a new modus operandi for science. Open science is as important and disruptive as e-commerce has been for the retail industry. Just like e-commerce, the open science research paradigm shift affects the whole business cycle of doing science and research. From the selection of research subjects to the carrying out of research, to its use and re-use, to the role of universities, and that of publishers are all dramatically changed. Just as the internet and globalization have profoundly changed the way we do business, interact socially, consume culture, and buy goods, these changes are now profoundly impacting how one does research and science.

The discussion on broadening the footprint of science and on novel ways to produce and spread knowledge gradually evolved from two global trends: Open Access and Open Source. The former refers to online, peer-reviewed scholarly outputs, which are free to read, with limited or no copyright and licensing restrictions, while open source refers to software created without any proprietary restriction and which can be accessed and freely used. Although open access became primarily associated with a particular publishing

or scientific dissemination practice, open access already sought to induce a broader practice that includes the general re-use of all kinds of research products, not just publications or data. It is only more recently that open science has coalesced into the concept of a transformed scientific practice, shifting the focus of researchers' activity from publishing as fast as possible to sharing knowledge as early as possible. Open science is defined as the idea that scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process. As a result, the way science is done in the future will look significantly different from the way it is done now. Open science is the ongoing evolution in the modus operandi of doing research and organizing science. This evolution is enabled by digital technology and is driven by both the globalization of the scientific community and increasing public demand to address the societal challenges of our times. Open science entails the ongoing transitions in the way research is performed, researchers collaborate, knowledge is shared, and science is organized.

Open science impacts the entire research cycle, from the inception of research to its publication, and on how this cycle is organized. The outer circle reflects the new interconnected nature of open science, while the inner circle shows the entire scientific process, from the conceptualization of research ideas to publishing. Each step in the scientific process is linked to ongoing changes brought about by open science, including the emergence of alternative systems to establish a scientific reputation; changes in the way quality and impact of research are evaluated; the growing use of scientific blogs; open annotation; and open access to data and publications. All institutions involved in science are affected, including research organizations, research councils, and funding bodies. The trends are irreversible, and they have already grown well beyond individual projects. Theses changes predominantly result from a bottom-up process driven by a growing number of researchers who increasingly employ social media in their research and initiate globally coordinated research projects while sharing results at an early stage in the research process.

Open science is encompassed in five schools of thought:

- the infrastructure school, concerned with technological architecture
- the public school, concerned with the accessibility of knowledge creation
- the measurement school, concerned with alternative impact assessment
- the democratic school, concerned with access to knowledge
- the pragmatic school, concerned with collaborative research

According to the measurement school, the reputation and evaluation of individual researchers are still mainly based on citation-based metrics. The h-index is an author-level metric that attempts to measure both the productivity and citation impact of the publications of a scientist or scholar. The impact factor is a measure reflecting the average number of citations to articles published in an academic journal and is used as a proxy for the relative importance of a journal.

Numerous criticisms have been made of citation-based metrics, primarily when used, and often misused, to assess the performance of individual researchers. These metrics:

- are often not applicable at the individual level
- do not take into account the broader social and economic function of scientific research
- are not adapted to the increased scale of research
- cannot recognize new types of work that researchers are performing

Web-based metrics for measuring research output, popularized as altmetrics, have recently received much attention: some measure the impact at the article level, others make it possible to assess the many outcomes of research in addition to the number of scientific articles and references. The current reputation and evaluation system has to adapt to the new dynamics of open science and acknowledge and incentivize

engagement in open science. Researchers engaging in open science have growing expectations that their work, including intermediate products such as research data, will be better rewarded or taken into account in their career development. Vice-versa, the use, and reuse of open data will require appropriate codes of conduct requiring, for example, the proper acknowledgment of the original creator of the data.

These ongoing changes are progressively transforming scientific practices with innovative tools to facilitate communication, collaboration, and data analysis. Researchers that increasingly work together to create knowledge can employ online tools and create a shared space where creative conversation and collaboration can occur. As a result, the problem-solving process can be faster, and the range of problems that can be solved can be expanded. The ecosystem underpinning open science is evolving very rapidly. Social network platforms for researchers already attract millions of users and are being used to begin and validate more research projects.

Furthermore, the trends towards open access are redefining the framework conditions for science and thus have an impact on how open innovation is produced by encouraging a more dynamic circulation of knowledge. It can enable more science-based startups to emerge thanks to the exploitation of openly accessible research results. Open science, however, does not mean free science. It is essential to ensure that intellectual property is protected before making knowledge publicly available in order to subsequently attract investments that can help translate research results into innovation. If this is taken into account, fuller and broader access to scientific publications and research data can help to accelerate innovation. Investments that boost research and innovation in open science would benefit society with fewer barriers to knowledge transfer, open access to scientific research, and greater mobility of researchers. In this context, open access can help overcome the barriers that innovative organizations face in accessing the results of research funded by the public.

Open innovation

An open society is the largest producer of knowledge, but the phenomenon of open science is changing every aspect of the scientific method by becoming more open, inclusive, and interdisciplinary. Ensuring open society is at the forefront of open science means promoting open access to scientific data and publications alongside the highest standards of research integrity. There are few forces in this globe as engaging and unifying as science. The universal language of science maintains open channels of communication globally. Open society can maximize its gains through maintaining its presence at the highest level of scientific endeavor, and by promoting a competitive edge in the knowledge society of the information age. The ideas and initiatives described in this publication can stimulate anyone interested in open science research and innovation. It is designed to encourage debate and lead to new ideas on what and open society should do, should not do, or do differently.

An open society can lead to a research powerhouse; however, open society rarely succeeds in turning research into innovation and in getting research results to the global market. Open society must improve at making the most of its innovation talent, and that is where open innovation comes into play. The basic premise of open innovation is to open up the innovation process to all active players so that knowledge can circulate more freely and be transformed into products and services that create new markets while fostering a stronger culture of entrepreneurship. Open innovation is defined as the use of purposive inflows and outflows of knowledge to accelerate internal innovation. This original notion of open innovation was primarily based on transferring knowledge, expertise, and even resources from one company or research institution to another. This notion assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as they seek to improve their performance. The concept of open innovation is continually evolving and is moving from linear, bilateral transactions and collaborations

towards dynamic, networked, multi-collaborative innovation ecosystems. This means that a specific innovation can no longer be seen as the result of predefined and isolated innovation activities but rather as the outcome of a complex co-creation process involving knowledge flows across the entire economic and social environment. This co-creation takes place in different parts of the innovation ecosystem and requires knowledge exchange and absorptive capacities from all the actors involved, whether businesses, academia, financial institutions, public authorities, or citizens.

Open innovation is a broad term, which encompasses several different nuances and approaches. Two main elements underpin the most recent conceptions of open innovation: the users are in the spotlight and invention becomes an innovation only if users become a part of the value creation process. Notions such as user innovation emphasize the role of citizens and users in the innovation processes as distributed' sources of knowledge. This kind of public engagement is one of the aims of open science research and innovation. The term 'open' in these contexts has also been used as a synonym for 'user-centric'; creating a well-functioning ecosystem that allows co-creation and becomes essential for open innovation. In this ecosystem, relevant stakeholders are collaborating along and across industry and sector-specific value chains to co-create solutions for socio-economic and business challenges. One important element to keep in mind when discussing open innovation is that it cannot be defined in absolutely precise terms. It may be better to think of it as a point on a continuum where there is a range of context-dependent innovation activities at different stages, from research to development through to commercialization, and where some activities are more open than others. Open innovation is gaining momentum thanks to new large-scale trends such as digitalization and the mass participation and collaboration in innovation that it enables. The speed and scale of digitalization are accelerating and transforming the way one designs, develops, and manufactures products, the way one delivers services, and the products and services themselves. It is enabling innovative processes and new ways of doing business, introducing new cross-sector value chains and infrastructures.

Open society must ensure that it capitalizes on the benefits that these developments promise for citizens in terms of tackling societal challenges and boosting business and industry. Drawing on these trends, and with the aim of helping build an open innovation ecosystem in open society, the open society's concept of open innovation is characterized by:

- combining the power of ideas and knowledge from different actors to co-create new products and find solutions to societal needs
- creating shared economic and social value, including a citizen and user-centric approach
- capitalizing on the implications of trends such as digitalization, mass participation, and collaboration

In order to encourage the transition from linear knowledge transfer towards more dynamic knowledge circulation, experts agree that it is essential to create and support an open innovation ecosystem that facilitates the translation of knowledge into socio-economic value. In addition to the formal supply-side elements such as research skills, excellent science, funding and intellectual property management, there is also a need to concentrate on the demand side aspects of knowledge circulation, making sure that scientific work corresponds to the needs of the users and that knowledge is findable, accessible, interpretable and reusable. Open access to research results aims to make science more reliable, efficient, and responsive and is the springboard for increased innovation opportunities, e.g. by enabling more science-based startups to emerge. Prioritizing open science does not, however, automatically ensure that research results and scientific knowledge are commercialized or transformed into socio-economic value. In order for this to happen, open innovation must help to connect and exploit the results of open science and facilitate the faster translation of discoveries into societal use and economic value.

Open Science

Collaborations with global partners represent important sources of knowledge circulation. The globalization of research and innovation is not a new phenomenon, but it has intensified in the last decade, particularly in terms of collaborative research, international technology production, and worldwide mobility of researchers and innovative entrepreneurs. Global collaboration plays a significant role both in improving the competitiveness of open innovation ecosystems and in fostering new knowledge production worldwide. It ensures access to a broader set of competencies, resources, and skills wherever they are located, and it yields positive impacts in terms of scientific quality and research results. Collaboration enables global standard-setting, allows global challenges to be tackled more effectively, and facilitates participation in global value chains and new and emerging markets.

To learn more, visit online at waset.org

Open Science

Scholarly Research Review

The scholarly research review is a multidimensional evaluation procedure in which standard peer review models can be adapted in line with the ethos of scientific research, including accessible identities between reviewer and author, publishing review reports and enabling greater participation in the peer review process. Scholarly research review methods are employed to maintain standards of quality, improve performance, provide credibility, and determine suitability for publication. *Responsible Peer Review Procedure:* Responsible peer review ensures that scholarly research meets accepted disciplinary standards and ensures the dissemination of only relevant findings, free from bias, unwarranted claims, and unacceptable interpretations. Principles of responsible peer review:

• Honesty in all aspects of research
• Accountability in the conduct of research
• Professional courtesy and fairness in working with others
• Good stewardship of research on behalf of others

The responsibilities of peer review apply to scholarly researchers at all stages of peer review: Fairness, Transparency, Independence, Appropriateness and Balance, Participation, Confidentiality, Impartiality, Timeliness, Quality and Excellence, Professionalism, and Duty to Report.

Scholarly Research Review Traits:
• Scholarly Research Review Identities: Authors and reviewers are aware of each other's identity

• Scholarly Research Review Reports: Review reports are published alongside the relevant article

• Scholarly Research Review Participation: The wider academic community is able to contribute to the review process

• Scholarly Research Review Interaction: Direct reciprocal discussion between author(s) and reviewers, and/or between reviewers, is allowed and encouraged

• Scholarly Research Pre-review Manuscripts: Manuscripts are made immediately available in advance of any formal peer review procedures

• Scholarly Research Review Final-version Reviewing: Editorial revision of the language and format is conducted on the final version of the manuscript for publication

• Scholarly Research Review Platforms: The scholarly research review process is independent of the final publication of the manuscript and it is facilitated by a different organizational entity than the venue of publication

All submitted manuscripts are subject to the scholarly research review process, in which there are three stages of evaluation for consideration: pre-review manuscripts, chair-review presentation, and final-review manuscripts. All submitted full text papers, that may still be withstand the editorial review process, are presented in the conference proceedings. Manuscripts are tracked and all actions are logged by internal and external reviewers according to publication policy. External reviewers' editorial analysis consists of the evaluation reports of the conference session chairs and participants in addition to online internal and external reviewers' reports. Based on completion of the scholarly research review process, those manuscripts meeting the publication standards are published 10 days after the event date.

To learn more, visit online at waset.org

TABLE OF CONTENTS

An Integrated Approach to Assessing Urban Nature as an Indicator to Mitigate Urban Heat Island Effect: A Case Study of Lahore, Pakistan

Muhammad Nasar-u-Minallah*, Dagmar Haase, Salman Qureshi

Landscape Ecology Lab, Institute of Geography, Humboldt University of Berlin, Germany

*Corresponding Email: muhammad.nasar-u-minallah@hu-berlin.de

Abstract: Rapid urbanization significantly change land use, urban nature, land surface vegetation cover, and heat distribution, leading to the formation of urban heat island (UHI) effect and affecting the healthy growth of cities and the comfort of human living style. Past information and present changes in Land Surface Temperature (LST) and urban landscapes could be useful to geographers, environmentalists, and urban planners in an attempt to shape the urban development process and mitigate the effects of urban heat islands (UHI). This study aims at using Satellite Remote Sensing (SRS) and GIS techniques to develop an approach for assessing the urban nature and UHI effects in Lahore, Pakistan. The study employed the Radiative Transfer Method (RTM) in estimating LST to assess the SUHI effect during the interval of 20 years (2000-2020). The assessment was performed by the available Landsat 7/ETM+ and Landsat 8/OIL_TIRs data for the years 2000, 2010, and 2020 respectively. Pearson's correlation and normalized mutual information were applied to investigate the relationship between green space characteristics and LST. The result of this work revealed that the influence of urban heat island is not always at the city centers but sometimes in the outskirt where a lot of development activities were going on towards the direction of expansion of Lahore, Pakistan. The present study explores the usage of image processing and spatial analysis in the drive towards achieving urban greening of Lahore and a sustainable urban environment in terms of urban planning, policy, and decision making and promoting the healthy and sustainable urban environment of the city.

Keywords: Urban nature, Urban Heat Islands, Urban Green Space, Land use, Lahore

Canopy Temperature Acquired from Daytime and Nighttime Aerial Data as an Indicator of Trees' Health Status

Agata Zakrzewska, Dominik Kopeć, Adrian Ochtyra

***Abstract*—**The growing number of new cameras, sensors, and research methods allow for a broader application of thermal data in remote sensing vegetation studies. The aim of this research was to check whether it is possible to use thermal infrared data with a spectral range (3.6-4.9 μm) obtained during the day and the night to assess the health condition of selected species of deciduous trees in an urban environment. For this purpose, research was carried out in the city center of Warsaw (Poland) in 2020. During the airborne data acquisition, thermal data, laser scanning, and ortophotomap images were collected. Synchronously with airborne data, ground reference data were obtained for 617 studied species (*Acer platanoides, Acer pseudoplatanus, Aesculus hippocastanum, Tilia cordata,* and *Tilia* × *euchlora*) in different health condition states. The results were as follows: (i) healthy trees are cooler than trees in poor condition and dying both in the daytime and nighttime data; (ii) the difference in the canopy temperatures between healthy and dying trees was 1.06°C of mean value on the nighttime data and 3.28°C of mean value on the daytime data; (iii) condition classes significantly differentiate on both daytime and nighttime thermal data, but only on daytime data all condition classes differed statistically significantly from each other. In conclusion, the aerial thermal data can be considered as an alternative to hyperspectral data, method of assessing the health condition of trees in an urban environment. Especially data obtained during the day, which can differentiate condition classes better than data obtained at night. The method based on thermal infrared and laser scanning data fusion could be a quick and efficient solution for identifying trees in poor health that should be visually checked in the field.

***Keywords*—** middle wave infrared, thermal imagery, tree discoloration, urban trees

A. Z. Department of Biogeography, Paleoecology and Nature Protection, Faculty of Biology and Environmental Protection, University of Lodz, Łódź, Poland (e-mail: agata.zakrzewska@biol.uni.lodz.pl)

D. K. Department of Biogeography, Paleoecology and Nature Protection, Faculty of Biology and Environmental Protection, University of Lodz, Łódź, Poland (e-mail: dominik.kopec@biol.uni.lodz.pl)

A. O. Department of Geoinformatics, Cartography and Remote Sensing, Faculty of Geography and Regional Studies, Warsaw University, Warsaw, Poland (e-mail: adrian.ochtyra@uw.edu.pl)

Building Envelope Engineering and Typologies for Complex Architectures: Composition and Functional Methodologies

Massimiliano Nastri

Abstract**—**The study examines the façade systems according to the constitutive and typological characters, as well as the functional and applicative requirements such the expressive, constructive and interactive criteria towards the environmental, perceptive and energy conditions. The envelope systems are understood as instruments of mediation, interchange and dynamic interaction between environmental conditions. The façades are observed for the sustainable concept of eco-efficient envelopes, selective and multi-purpose filters, adaptable and adjustable according to the environmental performance. In particular, the study examines the contents and the methodological and systemic guidelines concerning building façades, which are expressed in terms of morpho-typological, environmental, interactive and energy characters: the design, production and manufacturing of building envelope systems, aimed at performing morpho-typological, perceptual, physical and environmental functions are dealt with the evolved curtain wall configuration and with the geometric and material continuity of "light façades". The evolved configuration of curtain wall is achieved according to the conditions arising from the widespread "technical opening" offered by the contemporary industrial production of façade systems.

Keywords**—**Typologies of façades, environmental and energy sustainability. interaction and perceptive mediation. technical skins.

I. Introduction

THE The study around the field of envelope elements considers the surfaces applied to façade systems in their characters of morphological expression and interaction with external environmental loads, observing, above all, the procedures for controlling thermal and luminous transmission. Therefore, the study of the envelope explicates the physical, material and performance contents according to the criteria of action towards both energy and environmental conditions, and ergonomic conditions through the procedures of reflection, capture and diffusion of stresses external or internal to the built spaces. This is done through "passive" transformation, aimed at accumulating and distributing the energy produced by solar radiation without the use of plant equipment, or "active" transformation, with the contribution of devices (in the form of "collectors") aimed at integrating and conveying heat, natural light or convection phenomena related to air flows. The technologies applied and, in particular, the components and devices of the envelope are assumed with respect to the processes of "permeability" (determining the criteria of energy, environmental and perceptive control of a "selective" and dynamic type), with the possibility of regulating the flows and conveying them in the overall operation. The envelope, as an apparatus of mediation and reaction towards external loads, is structured in the combination with the calibration of properties and performance (according to a "selective" orientation), with the technical design and with the coherent application to settlement needs and requirements (as an environmentally conscious design activity). Enclosure systems are specified in the constitution of integrated functional components with the objective of receiving, guiding and selecting environmental stresses to realize calibrated ergonomic conditions in built spaces. For this, the systems are endowed with engineering performances (such as multiple environmental performances), are articulated in the form of environmentally responsive walls (capable of actively "responding" to environmental stresses through perceptual and "organic contact" with climatic conditions) and engineered walls (such as equipment that can be operated by mechanical devices), aimed at regulating the transmission of heat, light and natural ventilation, together with the attenuation of wind and acoustic loads.

II. The Typological, Functional and Operational Definition of Envelope Systems

The study deals with the envelope systems conceived as light prefabricated external walls (made in the laboratory and assembled on site), identified within the perimeter walls, in the form of homogeneous and uniform vertical enclosures in the external surfaces. The envelope systems are examined according to the flat vertical composition and the combination of a series of functions, such as the delimitation and protection of built spaces, visibility, regulation and control of thermal incidence, natural lighting and ventilation, with respect to climatic conditions and stresses (mechanical, thermo-hygrometric, wind and acoustic). The envelope systems (which do not contribute to the main load-bearing performance) are considered as external enclosure units of the architectural organism, built by frames (mainly produced of metal, but also of wood or plastic materials) made of vertical and horizontal structural elements (which themselves are connected and anchored with interface devices to the main load-bearing structures, which are generally frame structures). These units, established as non-load-bearing external walls (and, therefore, supported by a vertical and/or horizontal support grid), consist of technical elements (with a modular production base)

Massimiliano Nastri is Associate Professor of Building Technology, Material Balance Lab+Engineering, Department ABC - Architecture, Built Environment and Construction Engineering, Politecnico di Milano, Italy (e-mail: massimiliano.nastri@polimi.it).

prefabricated and joined to the load-bearing structures of the architectural organism through mechanical assembly techniques, aiming at separating the interiors from the external environment.

The envelope systems are composed within structural frames capable of supporting, accommodating, combining, engaging and hooking the subordinate arrangements defined by the formulation of enclosure components. Therefore, the elaboration of contemporary envelope systems relates to curtain walls and to the extension of the curtain wall applied to vertically developed buildings. This typology indicates an external non-load-bearing enclosure made of panels, of glass or other material, connected to a support frame that is assembled to the main structural elements (vertical and horizontal) of the architectural organism. The composition of the façade systems through the design, production and application of lightweight prefabricated components is outlined in the form of the external curtain wall made on steel or reinforced concrete framed structures, as a unit that can be separated from the whole construction and has no load-bearing function (Fig. 1) [1].

Fig. 1 Pierpaolo Ricatti, *Diesel* Headquarters, Breganze (Vicenza) (©Courtesy of Schüco).

According to the general definition, concerning the vertical perimeter enclosures executed by means of light prefabricated components, the envelope is mainly made up of closing elements of the voids between the elements of the load-bearing framework, its morphology essentially depends on the materials and products used, on the laying methods and on the possible

forms of surface finishing. Furthermore, during the construction of the building envelope, different types of connection can be used between the parts making up the envelope itself (internal connections) and between this and the load-bearing structure of the building (external connections). The types and methods of connection vary according to the materials and products used, the structural typology of the building, and the number of functional layers of which the envelope is composed. In general, in contemporary construction, the elements of the envelope may be inserted between the grids of the buildings external structure, or they may be leaning against the exterior, directly attached to the beams and columns, or tied to a different, specially prepared support structure. The evolved configuration of the *curtain wall* takes shape according to the conditions, the impulse and the supply deriving from the widespread technical openness of the window and door industry for façade systems (already defined by the processes of *components approach* or *componenting*, with the use of aggregative rules for the assembly of pieces), which identifies a sphere characterized both by the multiplicity of combinations, and by the synergies between technical elements and materials of different productive origin, which must support the criteria of flexible relationship between the contents referring to the structural and closing elements, to the connection and functional devices, up to the assembling methods (according to the environmental and technological performances that are going to be examined and obtained). Within the design, production and construction scenario, the envelope systems are examined with respect to:

- the role of transition between internal and external space, independently, on a morpho-typological level, with respect to the intended uses and according to the combination of performance contents (as *useful skin*) and external aspects (as *ornamental packaging skin*);
- the constitution of the components in an integrated form, characterized by specialization processes aimed at assuming the overall quality at different levels, according to the procedures of structural, connective, geometric and dimensional coordination, in order to allow both the mechanical assembly methods and the application to multiple building types (Fig. 2) [2], [3].

Fig. 2 EFA - Esperienze Forme Architettura, *Chiesi Farmaceutici* Headquarters, Parma (©Courtesy of EFA).

Therefore, the envelope systems participate in the assignment of the expressive contents of architecture, so that the external composition of the architectural organism coincides with the functional and constructive solutions adopted for the vertical enclosures, through the reduction of material incidence, the perceptive dematerialization and the enhancement of the meanings related to lightness and transparency. Moreover, in relation to the results of

contemporary research, the examination of the technical and executive configuration of the envelope systems also addresses the innovative opportunities of morpho-typological expression correlated to aspects of a functional nature. This leads the design and application scenario to constant experimentation and considers the closing elements in their progressive acquisition of tools and performance solutions aimed at increasing performance with respect to mechanical and environmental stress as well as energy, light and air exchanges. The concept of envelope systems made from prefabricated components, therefore, includes the development of membranes separated from the supporting frames, with the necessary performance to provide protection and improve the ergonomics of interiors.

The building envelope is determined, on a conceptual and executive level, in the building façade, understood as:

- the external boundary with a representative function in the urban space with respect to the procedures of spatial composition: This is done through the (structural or formal) frameworks constituted by the frames and the corresponding backgrounds that are included;
- the integration tool towards the fruition and functional characters of both the context and the environmental, atmospheric and perceptive characters (Fig. 3) [4].

Fig. 3 OMI Architects, *Anaconda Cut* Building, Manchester (©Courtesy of Focchi).

In this perspective, the conception, production and construction of the envelope systems, by using semi-finished elements or small finished components to be assembled on site, is configured as an activity of selection, adaptation and connection of products. In this regard, the convergence between design culture and industrial culture is outlined according to the flexibility of the production chains, the innovation oriented to the flexibility (but also to the specialization) of the products and the provision of new performances, while respecting the principles of multi-material relationship and specialized stratification: This, on the whole, determines the fine-tuning of functions according to specific needs, compared to products that manifest morphological neutrality and numerous possibilities of use, articulation and joining. According to this approach, the compositional elaboration of the envelope systems (in the form of orders and rules conforming to the use of planar, modular and often industrially manufactured elements) is examined through:

- the application of traditional materials, on the basis of references and allusions to the physical properties of their

surfaces, grading and modulations on the façade level;

- the use of morpho-typological rules which are mediated with respect to the prefabricated components, the connection methods of which establish the logic of correlation both expressive and executive;
- the poetics of construction aimed at defining the semantic criteria of the frames and enclosures in conformity with the expression of the principles and modes of relationship between the parts and the materials. In this regard, the elaboration of the envelope systems considers the procedures of mechanical assembly as a strategy aimed at specifying the design and executive practice no longer according to a fixed and immutable configuration, but rather as an activity that identifies a series of feasible solutions, allowing, in some cases, the reversibility of the construction (Fig. 4) [5].

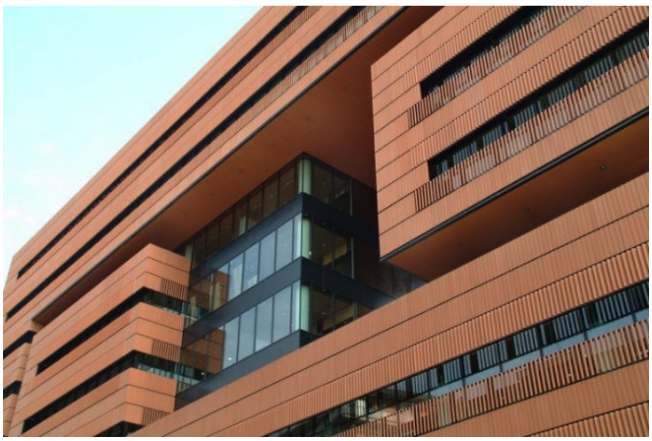


Fig. 4 Mario Botta and Giancarlo Marzorati, *Campari* Headquarters, Sesto S. Giovanni (Milan) (©Courtesy of Giancarlo Marzorati).

The characteristics of building envelope systems implies the following conditions:

- the adaptability of building products, within industrial logics characterized by high flexibility and construction strategies capable of encompassing different execution techniques. In particular, adaptability refers to the degree of variability of the technical elements, proportional to their degree of aggregation flexibility, that is, to their degree of autonomy in modifying their interactions with those belonging to other elements, which are conceived in an open manner through the use of solutions designed to allow subsequent integration;
- modularity, which, as a design and construction strategy, requires interchangeability and the possibility of combining technical elements, based on operational and procedural references such as standardization and modular coordination. In particular, modular coordination constitutes a method for determining the dimensions of technical elements and the structures they create, also in order to realize the conditions for a market of interchangeable and dimensionally selected building products. The module, therefore, is no longer the exclusive function of the measure, but becomes a function of the object understood not as a finished element, but as a basic element, moving from the "module-measure" to the "module-object" for both the frames and the enclosure apparatus;
- flexibility, which proposes the adaptation, by means of extension or convertibility, of characteristics already present in the system or in the technical elements, both structural and enclosure (Fig. 5) [6].

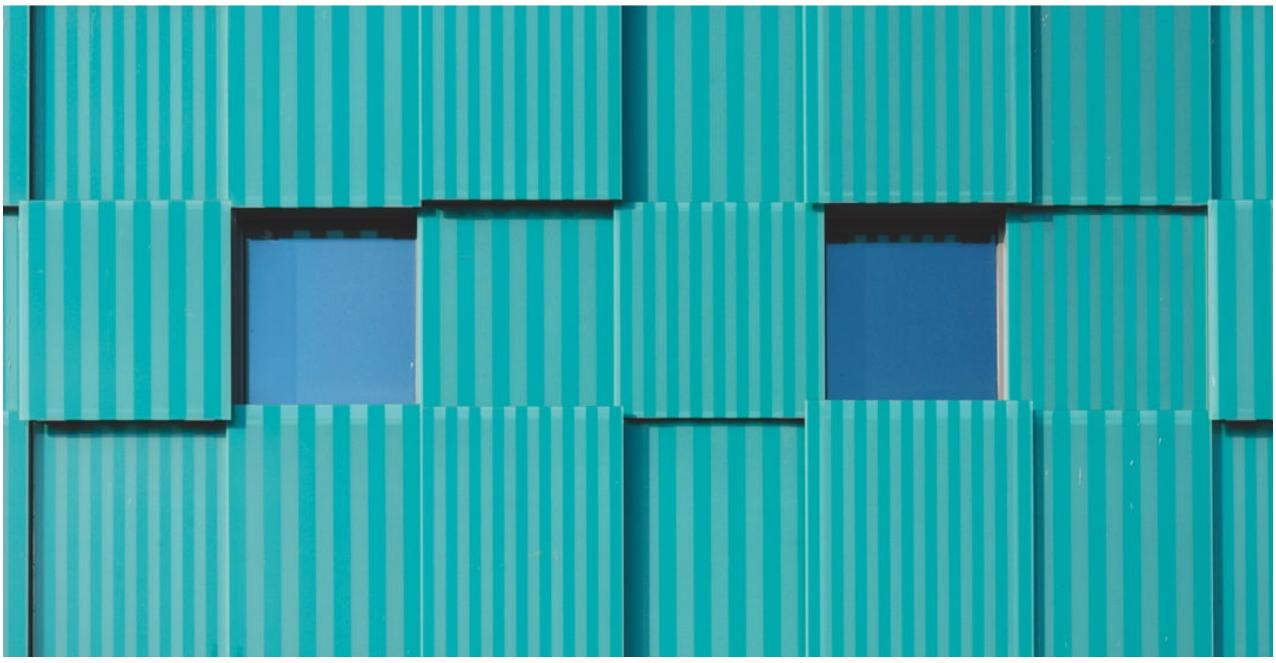


Fig. 5 Giancarlo Marzorati, *Barceló Hotel*, Milan (©Courtesy of Giancarlo Marzorati).

III. The Productive and Executive Articulation of Envelope Systems

The façade systems, designed and produced according to general needs, requirements and performance or for individual applications, are defined as light prefabricated external walls, executed in the form of curtain walls, cellular curtain walls, double skin curtain walls, structural curtain walls, suspended curtain walls and curtain walls with point fixing. In particular, both mullion and transom metal frame and structural and cellular façade types are specified by fixed and/or openable transparent closing elements, also combined with opaque elements (or *spandrels*, consisting of stratified panels, covered with pre-painted metal laminates, or panels with a ventilated air chamber and covered with glass or stone).

The façade systems are mainly considered with respect to the formulation of the structural apparatuses and connection devices, capable of resisting static and dynamic loads (such as wind stresses) and the deformations of the main load-bearing structures of the architectural organism (in order to support the movements due to differential settlements, to the inflections of pillars and beams, to horizontal loads). The structural apparatuses and connection devices are also identified in relation to their ability to respond to expansion and contraction phenomena, such as:

- thermal type, in order to absorb the stresses caused by temperature differences between indoor and outdoor, which can lead to deformations of the closing elements;
- hygroscopic, in order to allow differential movements of the enclosure elements due to humidity variations.

The load-bearing structural apparatus (consisting of the vertical and horizontal frames), the interface and connection devices, the components and the technical enclosure elements are then specified with respect to the requirements relating to:

- thermal resistance, so as to prevent the passage of heat by conduction, avoiding thermal bridges (especially in the case of metal profiles) which can cause localized condensation of water vapor on the internal surfaces of the enclosures;
- resistance to oxidation, exposure to ultraviolet rays and corrosion of metal parts;
- water tightness, by providing connective solutions (through the correct application of joints) capable of waterproofing situations of accumulation (e.g. in the interstices and along the projections) and of infiltration (e.g. in the joints);
- air-tightness, providing connective solutions capable of preventing infiltration of high flow rates into the interior, combined with the need to avoid the phenomena of water vapor condensation and acoustic transmission. This is achieved through the correct application of sealants, gaskets (for closing and filling joints) and drainage methods;
- vapor-tightness by providing direct solutions for the evacuation of condensation water;
- fire resistance;
- durability and maintainability, providing solutions capable of allowing periodic cleaning, repair and replacement of technical elements [7].

The building envelope study examines the typological and productive, functional and executive articulation of:

- the curtain wall system (*stick system),* consisting of a load-bearing frame of mullions and transoms (generally made of aluminum or steel, as well as PVC) arranged according to structural, morphological and enclosure requirements using glass elements (generally double-glazed) and/or panels (such as the *panel system*, in the form of prefabricated two-dimensional elements with an outer covering surface and thermal and acoustic insulation layers). This system, both for flat or polygonal development, involves installation on site after preparation of the frame elements, which consists of cutting and machining the profiles (mainly extruded aluminum) to allow assembly procedures (using scaffolding at the perimeter of the building). These procedures concern, in particular:

- the execution of the vertical elements (mullions) to the main supporting structure;
- the execution of the horizontal elements (transoms) to the mullions;
- the application of the vertical enclosures (in double glazing or *spandrel* form, these with monolithic glazing and panels behind or with steel, aluminum, composite or stone cladding elements) (Fig. 6);

Fig. 6 GaS Architects, *Campus Milano Internazionale*, Milan (©Courtesy of Pichler Projects).

- the structural façade system (*structural sealant glazing* or *glass curtain wall*), determined by the load-bearing frame (defined as adaptation and, in general, made of aluminum or steel) assembled in the laboratory and applied on site to the load-bearing structure of the architectural organism (with the use of scaffolding at the perimeter of the construction), where the closing elements (generally, monolithic glass panes or double-glazing) are made using structural silicone (without mechanical constraints, inside the profile grooves). The structural façade system consists of:

- the *strip window* or *two-sided supported system*, also known as "two-sided glazing", where the closing elements are applied to the frame by means of the structural silicone sealant, on the two opposite vertical sides of the mullions and by means of mechanical retaining flaps on the horizontal sides;
- the *total wall* or *four-sided supported system,* also known as "four-sided glazing", in which the closing elements are applied to the frame using structural silicone sealant on all sides of the profiles (Fig. 7);

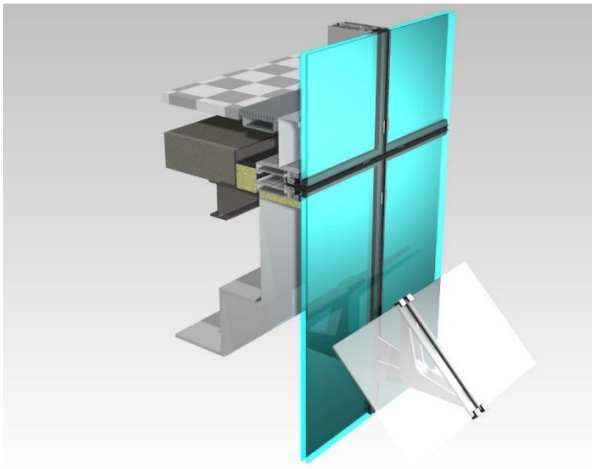


Fig. 7 Sidell Gibson Architects, *One Snow Hill* Building, Birmingham (©Courtesy of Focchi).

- the *unitized façade system*, determined by the load-bearing mullion and transom frame (generally in aluminum or steel) with a vertical extension that includes the height of the compartment and the stringcourse section, with closing elements. This system, preferred for vertical constructions

requires that the unitized façade system components are structurally independent and that they are connected to each other with telescopic or plug-in joints: these are capable of allowing, after installation, the movements for calibration on the façade level, carried out by means of the vertical and horizontal configuration profiles for performing the contiguity connections (Fig. 8);

Fig. 8 Park Associati, *U 27* Building, Assago (Milan) (©Courtesy of Focchi).

- the *double skin façade system*, determined in the transition from the continuous curtain to the multilayer type and articulating the specific performance of the layers and the relative technical elements. This involves the possibility of creating the cavity between the two walls for thermal and acoustic insulation, for ventilation and for housing functional devices as well as for plant ducts. The double-wall involves the use of a glass screen outside the vertical enclosure, with the aim of optimizing the functions of the ventilation in the cavity. The double skin system can be divided into the following types: façade with continuous cavity, without division, with discontinuous cavity and active façade (Fig. 9);

Fig. 9 5+1AA with Jean-Baptiste Pietri Architectes, *Fiera Milano* Office Building, Rho (Milan) (©Courtesy of Pichler Projects).

- the *suspended curtain wall* or *point fixed curtain wall* (or structural glass façade), determined by the use of glass sheet enclosure elements (made of vertical groups) by means of punctual constraints (in the form of mechanically supported fixing devices). These in turn are connected to steel cable frames anchored to the main load-bearing structure of the architectural organism (or to steel frames, usually tubular). Moreover, the *point-fixing façade system* is determined by the assembly of the glazing elements (with or without perforation) on metal supports (generally made of steel) (Fig. 10) [8], [9].

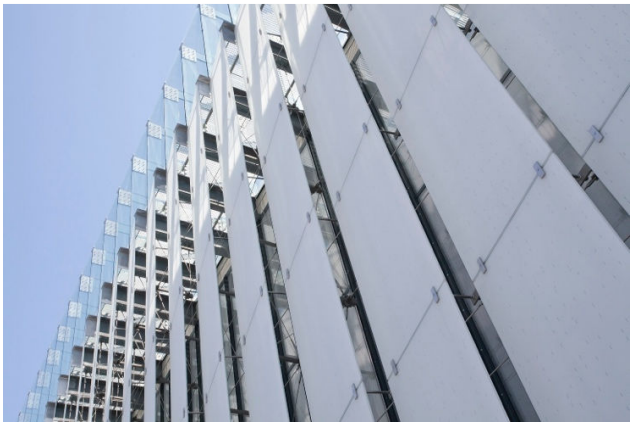


Fig. 10 Stelmach & Partners, *CSK* (*Centrum Spotkania Kultur*), Lublin (©Courtesy of Lilli Systems).

The composition of the envelope systems, constituted by the use of transparent enclosures, considers the surfaces applied to the façades in their morphological expression and interaction with the external environmental stresses, considering especially the procedures of control of the thermal and luminous transmission. The explanation concerns the productive, functional and use modalities of the techniques and materials aimed, in general, at the reduction of the loss or accumulation of heat (due to solar radiation), as well as the dynamic selection of the solar rays and the calibration of the natural light, in

observance to the types of use required. On this basis, the composition of the envelope systems is outlined with respect to the types, compatible solutions and procedures of use expressed in relation to the main references of a thermal, luminous and radiant nature. This is done through the definition of performance (at the mechanical, thermo-hygrometric, lighting and acoustic levels), execution (according to the relationships with the frames and joints) and the in-depth examination of the main factors of a physical, functional and applicative nature. In addition, the treatment detects the contributions of studies referring both to materials and devices from the experimental research sectors, and to the possibilities of being transferred from the advanced technology sectors. These contributions support, in the context of the systems under examination, the methods aimed at optimizing the physical, mechanical, chemical and thermal characteristics of the transparent surfaces, in order to calibrate the comfort conditions in the built spaces and the energy management criteria. In this case, the study exposes high-performance materials and compounds, indicated in the principles of both mediation between light transmission and thermal conduction (in order to reduce thermal dispersion without affecting transparency), and control of incident solar radiation (also with the adoption of selective coatings or shielding devices) (Fig. 11), [10].

Fig. 11 John McAslan & Partners, *Vitro* Tower, London (©by the Author).

In this respect, transparent enclosure elements applied to façade systems are established by:

- the passive use and transformation of solar radiation, using technical solutions aimed at capturing, accumulating and distributing the energy produced without the need for plant equipment. These solutions take on the functions of controlling the microclimate of the built spaces and the energy balance, according to the basic principles of solar heating and lighting;
- the use and active transformation of solar radiation, involving the use of technical solutions aimed at capturing, accumulating and distributing solar energy. These solutions involve the contribution of devices (in the form of collectors) capable of integrating the exploitation of heat, natural light or convection phenomena related to air flows.

Moreover, the area under examination addresses the visualization of alternative types for façade systems and is part of the progressive and linear developments in the construction sector and functional deepening. This is done with respect to a scenario where adaptable products and materials are used, simple, finished and small sized elements are marketed and diffused within the industrial logics characterized by high flexibility, capable of including and combining different solutions. The use of opaque enclosures expresses the façade systems with respect to:

- the tendency to rationalize and reinvent both the components and the conventional modes of application and interface, in an integrated approach to the multiplicity and variety of expressive possibilities;
- the hybridization of traditional materials, in order to legitimize the maintenance of the solid and massive presence, within the growing virtuality and ephemeral, dynamic and metamorphic configuration of the envelope (Fig. 12) [11].

Fig. 12 Maurizio Varratta, *iGuzzini Lab*, Recanati (©Courtesy of Pichler Projects).

The morphological and functional development that is outlined appears to be directed towards affirming the steady character of the unitary composition, by the expressive balance of the closing parts, as a response to the broken continuity of traditional languages and techniques, to the overlapping and integration of types and multiple fragments in the external constitution of architecture. This is achieved through:

- the application of elements capable of wrapping both the surfaces and the supporting and connecting apparatus, grafting themselves between the current processes of technical complexity and productive and linguistic articulation, often leading to the loss of the homogeneity and material congruity of the envelope;
- the assimilation of the characteristics of industrial production, aimed at overcoming the dichotomy between tradition and innovation, constituting an experimental basis both for the updating of sedimented practices and

techniques accepted as valid, and for the promotion of compatible solutions, of significant insertions characterized by the complementarity between consolidated and evolved materials and construction procedures [12].

IV. The Environmental and Energy Functioning of Envelope Systems

The envelope systems are defined as applications aimed at isolating, filtering or absorbing external climatic stresses (mainly thermal and luminous), in order to reduce the impact of technical devices. The systems are presented as instruments of mediation between the external environmental conditions and the ergonomic aspects of the built spaces, through specific performance reactions. These reactions are calibrated and variable according to the stresses and comfort requirements, highlighting the possible contributions to the reduction of energy use. In this regard, the systems are treated as interchange tools because of their ability to respond to external stresses, through the development of different functional levels (in terms of material, structure and thickness) and through the use of regulation means that allow to handle (in a natural, or passive form, and in an artificial, or active form) the interactions with the environment. For this reason, the enclosure elements behave as osmotic membranes acting according to processes of exchange of energy, light and air flows.

The typological examination introduces the identification of the closing elements that constitute the stratification of the envelope systems, according to the specialized functions developed with the precise subdivision of the performances assigned to the different materials and devices. The analysis then focuses on the perspectives of dynamic interaction between the envelope systems and the external environment. This is done by observing the criteria aimed at realizing built spaces in a steady and balanced way, with the possibility of transferring, modifying or repelling external stresses, according to the sustainable and low-energy concept of the components made up as reactive and sensitive surfaces to climatic stresses (Figs. 13, 14) [13].

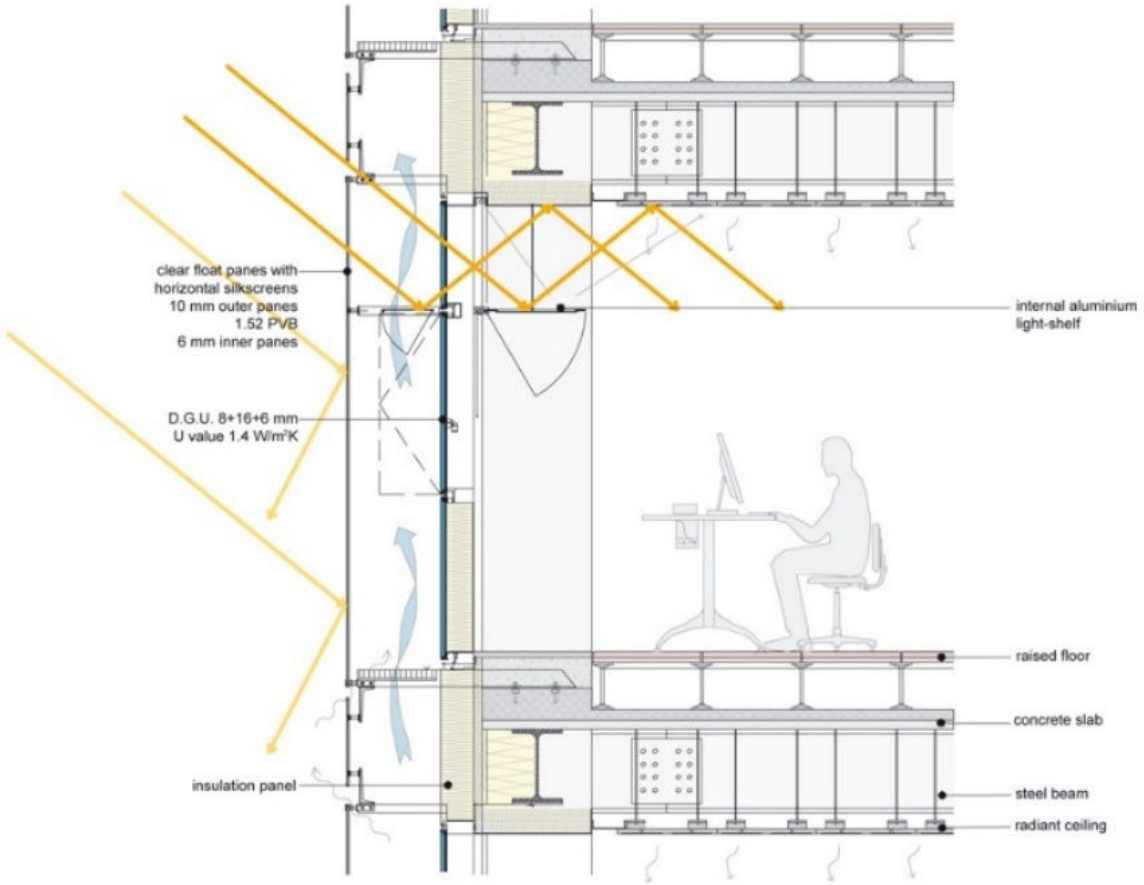


Fig. 13 Mario Cucinella Architects, *SIEEB* (*Sino-Italian Ecological and Energy Efficient Building*), Beijing (©Courtesy of Mario Cucinella Architects).

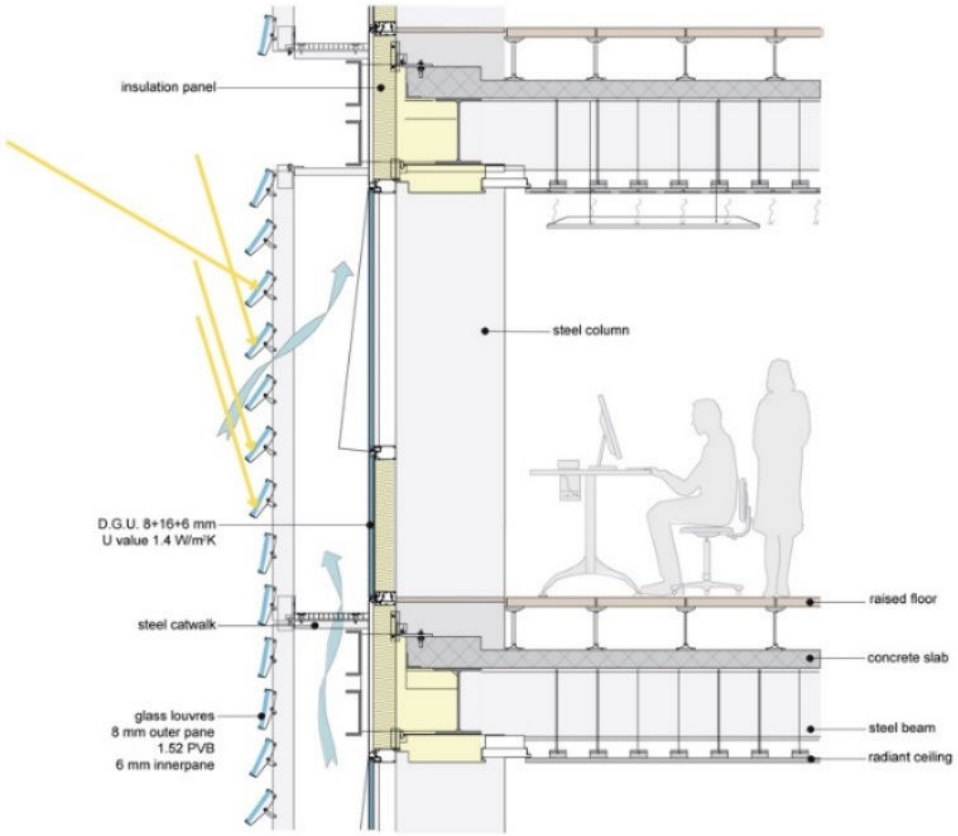


Fig. 14 Mario Cucinella Architects, *SIEEB* (*Sino-Italian Ecological and Energy Efficient Building*), Beijing (©Courtesy of Mario Cucinella Architects).

The advanced building envelope is considered as a *dynamic interface*, i.e., as a structure of exchange between the environmental stresses and the fruition needs of the internal spaces, having evolutive plasticity and adaptation qualities to the different loads coming from the environment. The experimentation here expresses the intention to integrate the climatic conditions and to convey them into the interior spaces according to established procedures and levels, to build components in the form of biomechanical prototypes in which the different parts specialize to achieve a certain function. In particular, the envelope systems are assumed as mediation and reaction apparatuses towards external stresses, considering their sensitive and organic nature (in the form of *technical skin*), their physical and material consistency, which is progressively determined towards integrated layers and performances that act, with adaptation and control capacity, according to the well-being requirements and reduction of energy consumption in built spaces. This assumption therefore focuses, on the one hand, on the aspects related to the adaptation and progressive control of thermal and energy exchanges between the internal micro-environment and the external macro-environment, identifying this area as a catalyst in the type-technological evolution of the building envelope; on the other hand, on the experimentation of technological and construction solutions aimed at optimizing these exchanges, as well as, obviously, on the expressive and communicative content that this experimentation implies. The definition of the envelope systems is thus carried out in the analysis of the dimensional consistency, which considers the building envelope as a system consisting of interacting parts, whose thickness varies according to the overall operation. Furthermore, the envelope systems are defined as external apparatuses capable of interpreting the functions and needs of their users as an *eco-efficient envelope*, as selective and multi-purpose filters, in relation to climatic conditions and ergonomic comfort requirements (of a thermo-hygrometric nature, relating to temperature control, humidity levels and ventilation. They also have a visual aspect, relating to perception towards the outside and control of lighting levels; an acoustic aspect, relating to insulation from sound stress; and an olfactory aspect, relating to control of air quality) (Fig. 15) [14].

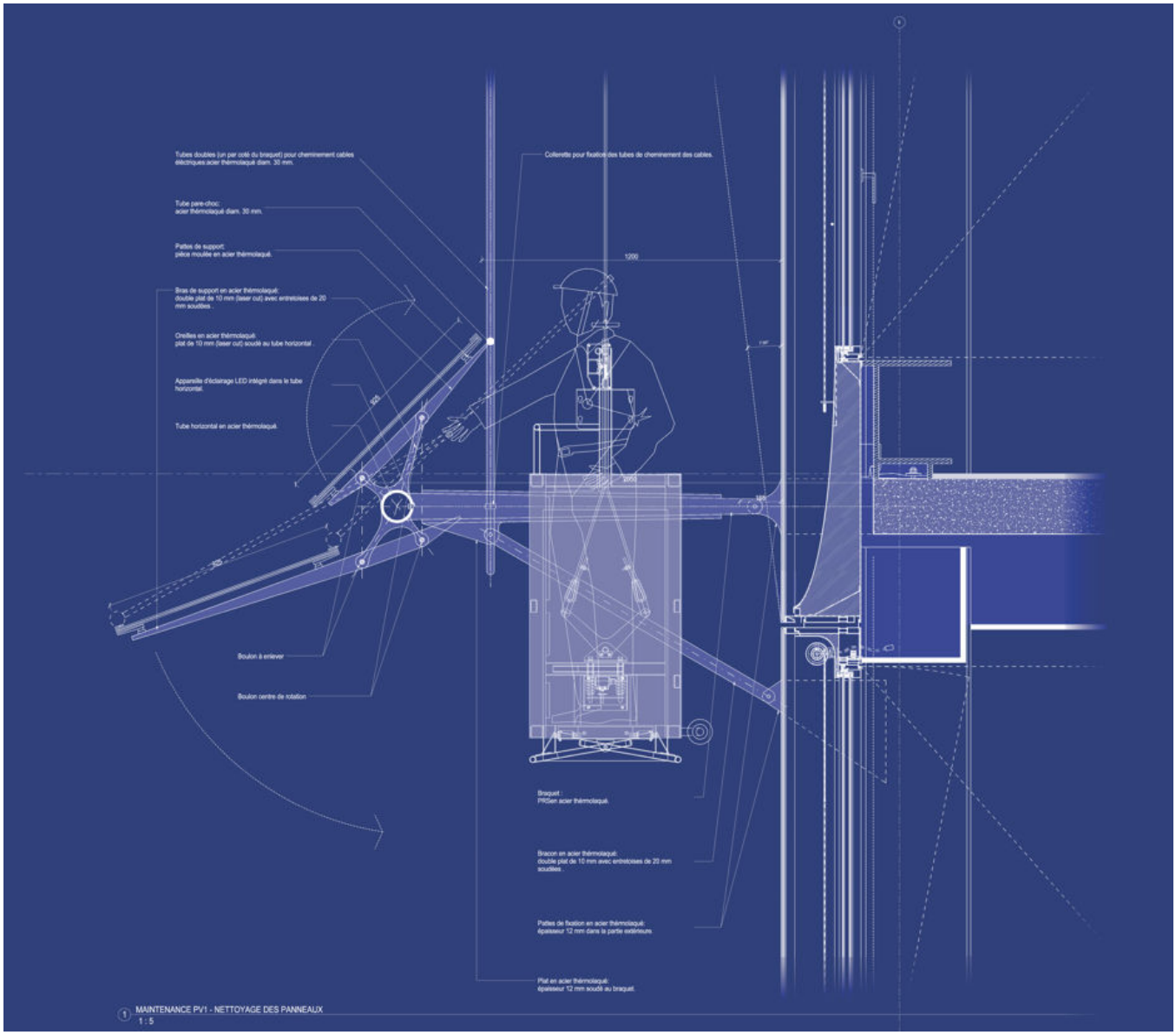


Fig. 15 Renzo Piano Building Workshop, *Courthouse*, Paris (©Courtesy of RPBW).

The analysis is then expressed through the investigation of systems as organic compounds, adaptable and adjustable as *biological skin*s and as *multifunctional skins*, i.e. as absorbing, radiating, reflecting, filtering and transferring (thermal, luminous, aerial) devices. In particular, the use of elements with dynamic and reactive behavior assumes the use of surfaces for the control of solar radiation, consisting of filtering or shielding sections able to modulate their transparency according to the level and distribution of natural light in the interior spaces. The closing elements of the envelope systems are especially investigated with respect to their character of modified physicality, which the experimental research tends to transform in thick and in *intelligent systems interface*. At the same way, also the main materials that constitute the external surfaces are analyzed with respect to their transformation processes from stable entities to designable entities according to a specific performance plan. In this case, the application of the envelope materials, in the form of designable entities, is examined with respect to the outcomes of the solutions in which the functions tend to become complex (in a controlled and managed way) and combined (in a *solid-state* form), achieving multiple performances through the correlation of different agents and layers. This is done by:

- the integration possibilities of functions related to enclosure elements, where relationships and interfaces (physical, performing) are arranged between individual parts and materials of a system or component;
- the development of custom-made enclosure elements, using materials specifically produced to perform certain functions and without having to adapt to the limits imposed by

original and predetermined properties. This is achieved by replacing composites made from multiple devices, assembled together, with elements resulting from the union of different surfaces within a single polyvalent layer;

- the systems capability to react to environmental stresses, according to passive or active regulation processes (by manual, mechanical or computerized commands) induced by electrical, thermal and lighting triggers that modify, through alterations in the physical or chemical structure, the disposition of the enclosing elements or the properties of materials (Fig. 16), [15].

Fig. 16. Norman Foster and Partners, *Greater London City Hall*, London (©by the Author).

The study of envelope systems is linked to research on sustainability in the construction industry, which considers the need to reduce energy consumption and polluting emissions with respect to the implementation of solutions that can establish high environmental performance. In this regard, the contribution of the envelope systems is related to the performance of technological systems, considering the operational procedures designed to balance the relationship between climatic conditions and well-being in built spaces. The design and application field of building envelope systems includes the evolution of conceptual and operational apparatuses referring to the foundations and requirements of environmental sustainability, addressing:

- the eco-efficiency of transformation processes, whereby the planning, production and construction process is determined, in its global form, both in the interaction with the eco-systems harmony, and in the acquisition of the appropriate levels of both environmental and urban quality;
- the paradigms of sustainability, defined by the principles expressed in a series of reports and protocols sanctioned internationally. These address the consequences of environmental impacts (caused to a large extent by management practices, especially energy management, of buildings) and are aimed both at protecting the environment and bio-ecological balances, and at preserving non-renewable (both material and energy) resources, as a basic condition for development (Fig. 17).

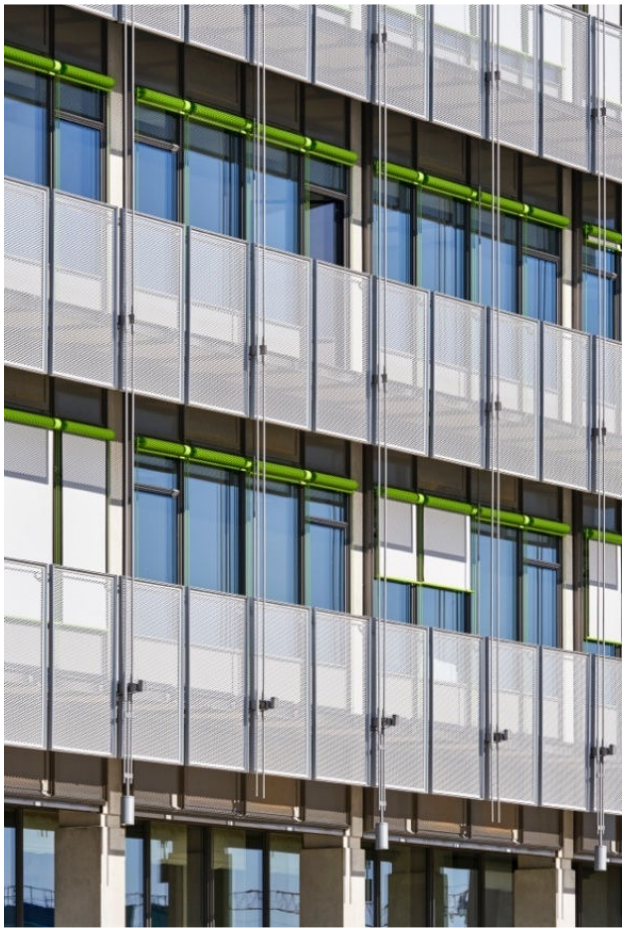


Fig. 17. Renzo Piano Building Workshop, *École Normale Supérieure*, Paris-Saclay (©Courtesy of RPBW).

In this sense, the analysis of envelope systems is linked to the way they interact with climate factors, thereby promoting the primary examination (aimed at the design, functional and application conception) of the external environmental conditions relating to:

- the (apparent) solar path and height above the horizon according to the specific context (latitude) and seasonal periods, which determine the intensity and angle of solar radiation;
- the intensity of solar radiation according to the orientation and inclination of the elevations;
- the solar radiation, either direct and/or indirect (according to the radiation being reflected from the celestial vault and the surrounding environment), according to the different wavelengths;
- the amount of energy input estimated according to the climatic conditions and the energy demand linked to the intended uses [16].

In general, building envelope systems are permeable to solar radiation, leading to the generation of energy inputs that can be accumulated, transmitted (by thermal conduction) and radiated into the interiors (depending on the thermal conductivity of the closing materials). The exposure, due to the different irradiation, affects the energy production of solar radiation,

contributing to solar gains following the heating of the built spaces and their structures. In this regard, vertical façades (facing east, west and south) have a high energy production (which is mainly derived from direct radiation) during the periods of highest exposure to sunlight, thus indicating the need for solar protection.

With regard to the requirements concerning indoor climatic conditions, in relation to the use of building envelope systems, it is considered:

the air temperature and relative humidity of the built-up spaces;

- the temperature and surface airstreams in the area around the closing elements;
- wellbeing requirements related to both the quality and quantity of light transmission and visual conditions (Figs. 18, 19).

Fig. 18. Gatermann + Schossig, *Capricorn Haus*, Düsseldorf (©Courtesy of Schüco).

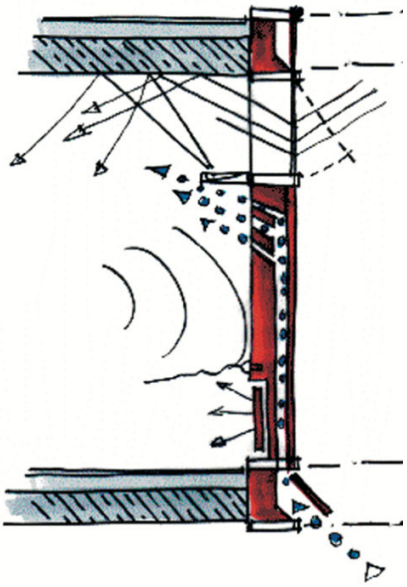


Fig. 19. Gatermann + Schossig, *Capricorn Haus*, Düsseldorf (©Courtesy of Schüco).

As far as the thermal quality of the envelope systems is concerned, which effects the possibility of energy saving through the exploitation of solar radiation, it is observed:

- the low thermal inertia, which provides sensitivity to cooling and overheating phenomena of built-up spaces due to prevailing transient thermal phenomena;
- the thermal dispersion due to infiltration losses, related to the joints between structural elements (i.e., with respect to frames) and closing elements.

At the production, design and construction level, it is necessary to define operational conditions of balance between conflicting requirements, such as the need to ensure a positive energy contribution during the winter period, reducing dispersion and providing thermal storage by transforming solar radiation into heat (according to the phenomena of thermal inertia) (Fig. 20), [17].

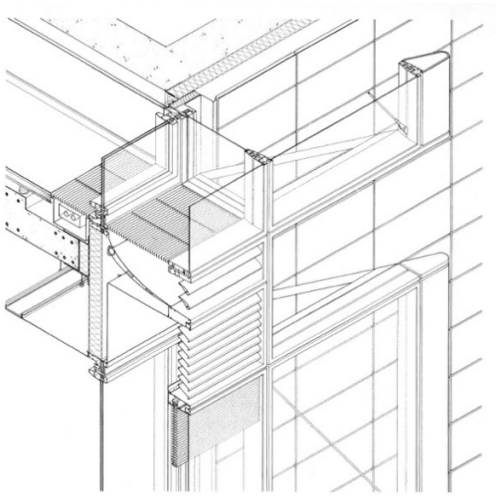


Fig. 20. Norman Foster and Partners with RKW (Rhode, Kellermann, Wawrowsky), *Arag* Tower, Düsseldorf (©Courtesy of RKW).

The combination of passive and active operating procedures regulates the creation of systems capable of self-regulation, sensitive to external climatic changes and to comfort requirements (both thermal and visual) and interiors ventilation. In this regard, the environmental and adaptive strategy develops the envelope systems according to their metabolic effectiveness and their instinctive reactive capacity, configuring themselves as *intelligent skins* endowed with automatic performance (by means of functional criteria of autonomous regulation) and membranes defined as *biological skins* (capable of reacting against external agents through the activation of sensors and protective devices). Furthermore the biological relationship considering the regulation systems (computerized building management systems, BMS) and the corresponding opportunities for opening and closing, protection, shielding and environmental incorporation, identifies the *hypothalamic function* reactive to external and internal stresses (Figs. 21, 22).

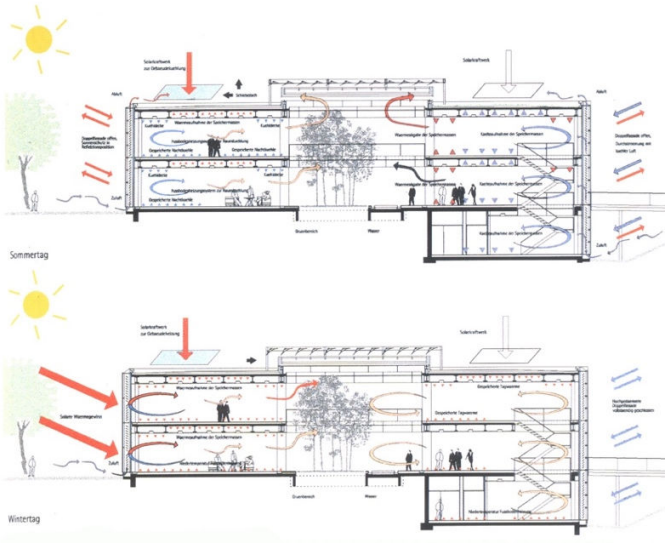


Fig. 21. Martin Webler and Garnet Geissler, *Götz* Headquarters, Würzburg (Germany) (©Courtesy of Götz).

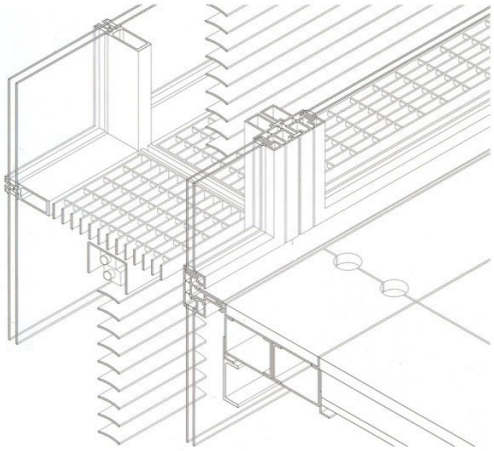


Fig. 22. Martin Webler and Garnet Geissler, *Götz* Headquarters, Würzburg (Germany) (©Courtesy of Götz).

The systems design thus exposes the *techno-organic* properties through the interpretation and assimilation of environmental conditions in a combination with the use of advanced techniques (in *organitech* form). In this manner, the study involves the investigation of artificial (or organic) systems integrated with natural systems, as tools for accumulation, channelling, protection and calibration of passive energies that can provide buildings with heating, air conditioning and ventilation. Consequently, the building envelopes are expressed by environmental diaphragms and by *neuronic façades*, built like "natural organisms", or rather like machines that aim to reproduce, manage and metabolize natural processes according to criteria of active understanding.

The design of building envelope systems is aimed at:

- securing natural lighting levels (according to reflection, absorption and diffusion criteria) in order to reduce energy consumption;
- controlling the overall transmission component of solar radiation (with particular attention to ultraviolet wave transmission), while not penalizing the transmission component within the visual field;
- controlling the perceptual-visual relationships and gradations;
- controlling air quality and humidity, especially through natural ventilation;
- reducing acoustic stress, through reflection and/or absorption of sound waves (Figs. 23, 24), [18].

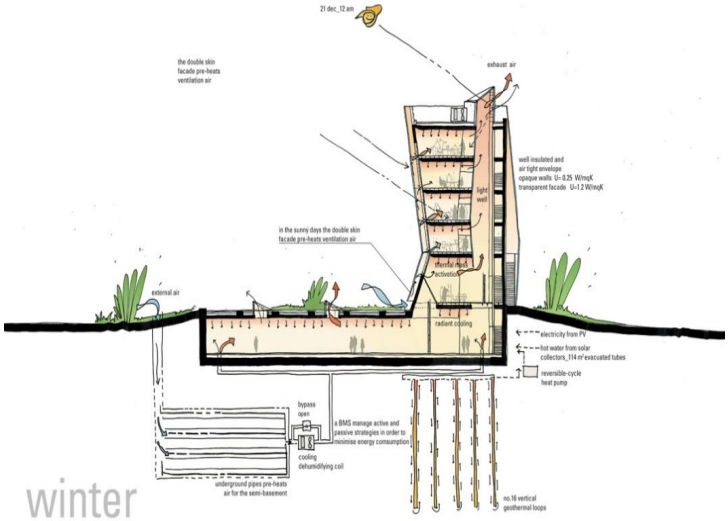


Fig. 23. Mario Cucinella Architects, *CSET* (*Centre for Sustainable Energy Technologies*), Ningbo (China) (©Courtesy of Mario Cucinella Architects).

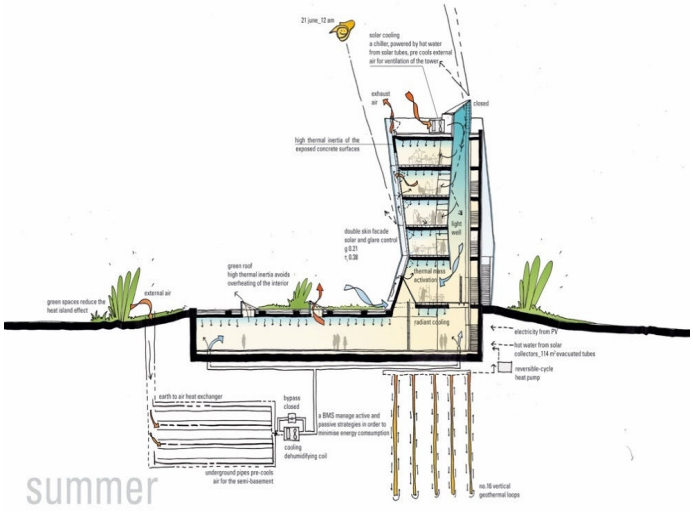


Fig. 24. Mario Cucinella Architects, *CSET* (*Centre for Sustainable Energy Technologies*), Ningbo (China) (©Courtesy of Mario Cucinella Architects).

V. Interactive Processing of Envelope Systems

The interactive design of the envelope systems analyses the constitution of façade surfaces as an information support and as a communicative tool, according to technological and multimedia business, along with the progressive dematerialization and immaterial transformation of architecture. In this respect, the composition of the building envelope is expressed through the acquisition of new visual and virtual potentials, which transcend the material aspects, which aim at the metamorphosis of curtains and which stand out as communicative devices. Within this scenario, the composition of envelope systems is configured both according to the loss of perspective stability and according to the emphasis on membranes and programmable surfaces (Fig. 25).

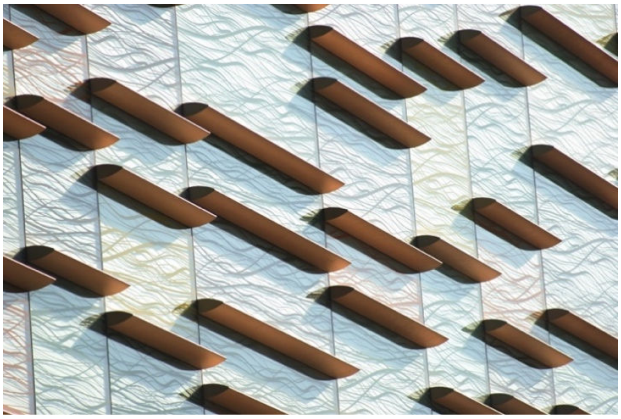


Fig. 25. Irene Sabato (Polis Engineering), Ivo Pellegri (DAGA Studio), *Guna Headquarters*, Milan (©Courtesy of Sipam).

The combination of expressive and performance opportunities, of working methods and morphological experiments, on which the composition of the interactive envelope is based, supports the evolution of stylistic features, so that the architectural works take on increasingly enigmatic aspects: the absolute essentiality of the interiors is matched by an exasperated sophistication of design and effect on the outward appearance, to the point of making us imagine the future landscape of our cities only as a continuous series of thin, changing envelopes. Such experimentation leads to alienation from the context in which the building envelopes are placed, in accordance with the desire to affirm that each work of architecture establishes itself through its own conceptual contents and its own principle of independent identity (Fig. 26), [19].

Fig. 26. Gianandrea Barreca and Giovanni La Varra, *"B5" Building, RCS - Media Group Headquarters*, Milan (©Courtesy of Focchi).

Specifically, the envelope systems are considered according to:

- the formulation of scenographic and catalyzing mechanisms open to multiple expressive and functional solutions, such as accumulators of pictures and as urban transmitters, assigning to the façade curtains the function of configuring themselves as an autonomous and communicative support;
- the elaboration of conceptual installations, through which the temporary, ephemeral and suggestive content of visual engagement is revealed, whose surfaces take in the suggestions of the media culture, thus asserting themselves as media façades, or hypersurfaces (as supports for the expressive potential of media);
- the modes of interaction and merging between the architectural work and the context, through the development of surfaces marked by discontinuities, foldings or stratifications, according to fluid and dynamic morphologies. The surfaces are examined as membranes, active and structured, connected to the concept of continuous movement and modulation. They are not identified by the necessity to produce thickness in the wrapping of the perimeter curtains, but with regard to the paradigms of immediacy and instantaneousness (Fig. 27), [20].

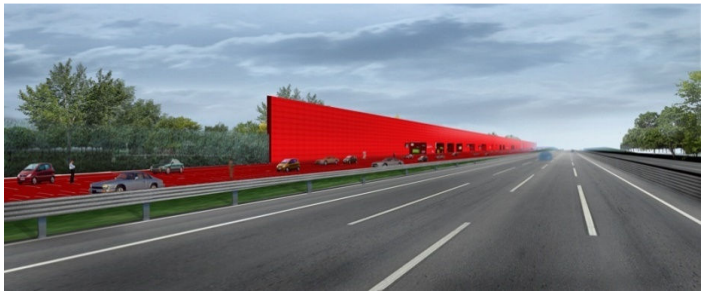


Fig. 27. Atelier Jean Nouvel + Studio Blast, Technological Scientific Park *Kilometro Rosso*, Stezzano (Bergamo) (©Courtesy of Studio Blast).

The experimental elaboration of the envelope, then, examines the application of the enclosures on the basis of the

environmental and spatial fading of boundaries (intended as the loss of their role as a clear-cut boundary between content and exterior). This is achieved through the dematerialization of the facades (aimed at a free conceptual, spatial, perceptive and evocative flow), operating on filters, diaphragms and transparencies. The interactive processing of envelope systems concerns, then, the dematerialization of containers, whereby surfaces reveal themselves as a media epidermis, as sensors capable of connecting reality phenomena and information demands.

The physical and material characters of surfaces are examined in relation to the loss of their tectonic consistency, proceeding to the expression of the permeability conditions, both functional and fruitive, and towards the aleatory articulation in conceptual and visual transitions, through:

- the contribution of digital processing, which allows to represent the organic, dynamic and metamorphic aspects of the envelope virtualization; the development of hypermedia perception standards, aimed at the deep layers of intellectual, emotional and sensorial reactivity;
- the plastic tension of the enclosure, taken to the extreme of its functions, whereby the barriers of closure are bypassed by the inclusion and dilution of visual passages (Figs. 28, 29), [21].

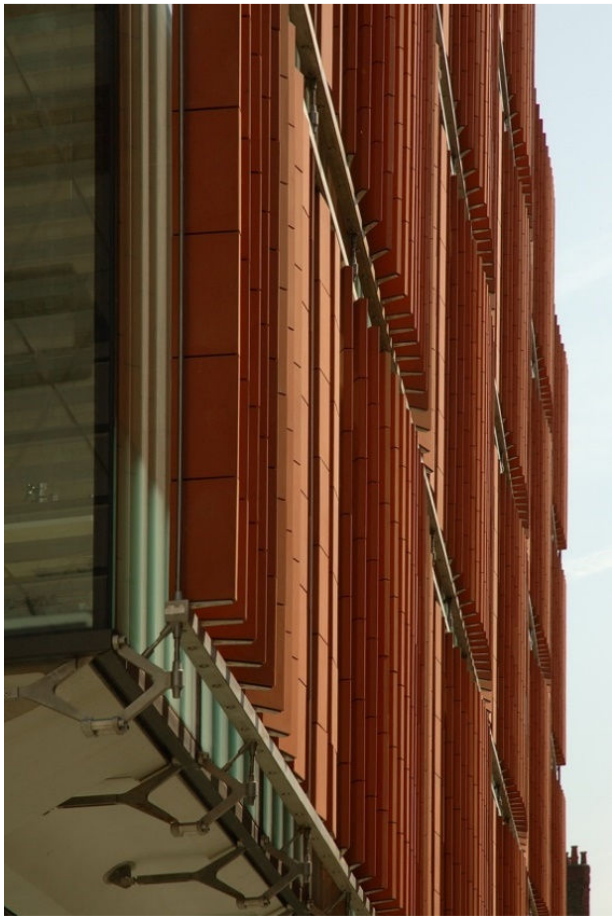


Fig. 28. Nicholas Grimshaw and Partners, *UCL Cancer Institute*, London (©Courtesy of Sannini Project).

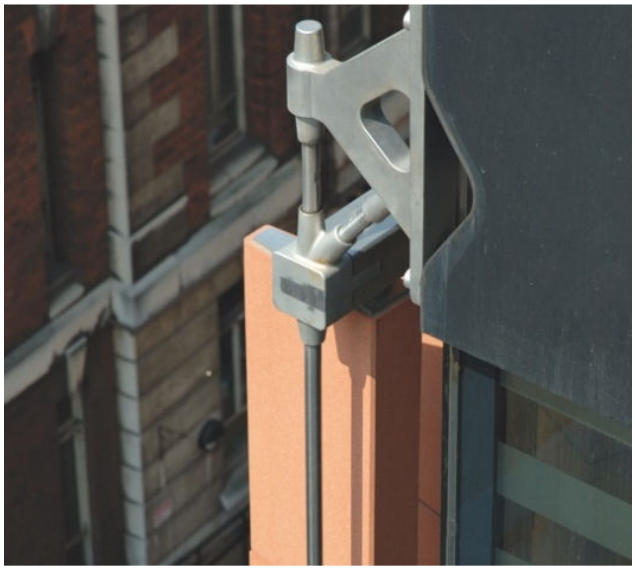


Fig. 29. Nicholas Grimshaw and Partners, *UCL Cancer Institute*, London (©Courtesy of Sannini Project).

Therefore, the enclosures in fulfilling the purpose of wrapping and delimiting, are conceived as removable slabs, as almost immaterial and mobile presences, in order to generate dialectical relationships between built spaces and the external context and to emphasize the interactive and organic logic of architecture. In particular, the wrapping of the building envelopes is determined through differing or calibrated densities, according to the principle of porosity and reticulation. This is achieved by using cuts, pixels and openings, engraved and interposed on the perimeter curtains. The enclosures are thus examined, in general, according to:

- the procedures of dematerialization and interconnection, both spatial and visual, considering the surfaces in their light, transparent constitution, aimed at specifying a symbiotic and engaging relationship with the external space. The envelope systems are thus conceived through flexible and reactive, metamorphic and unstable configurations, in relation to their transparency and opacity (Fig. 30);

Fig. 30. Valentina Bonato, Dario Cagol and Helmuth Niedermayr, *Alperia Tower*, Bolzano (©Courtesy of Oskar DaRiz).

- the organic deformation procedures, considering the curtains as textures that flow through space due to their porous constitution, perceptible and intelligible according to the temporality of movement. The envelope systems are then intended as a vibrant and changeable configuration, sensitive and interactive, adjustable to the urban and immaterial environment (Fig. 31).

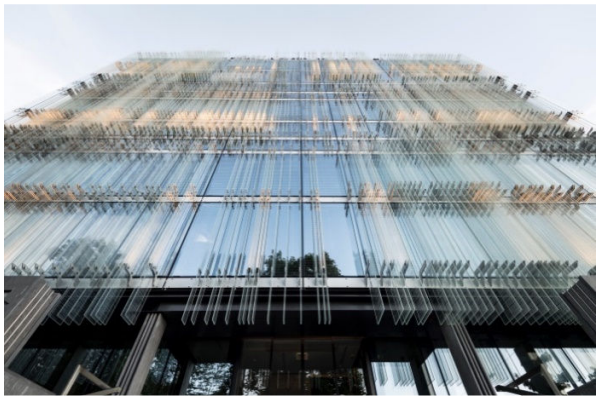


Fig. 31. Giovanni Vaccarini Architetti, *Société Privée de Gérance* (*SPG*), Geneva (©Courtesy of Alex Filz).

The envelope systems are therefore articulated through a number of observations on part of contemporary architecture that displays, through its own external sensitive films, through its own skin, a desire for effective communication, relating to a position of discontinuity with respect to the urban context: these are design researches that seek to introduce new image accumulators and place identities, according to complex aggregation devices and hierarchies, no longer capable of being identified and ordered according to the usual categories of urban analysis (Fig. 32), [22], [23].

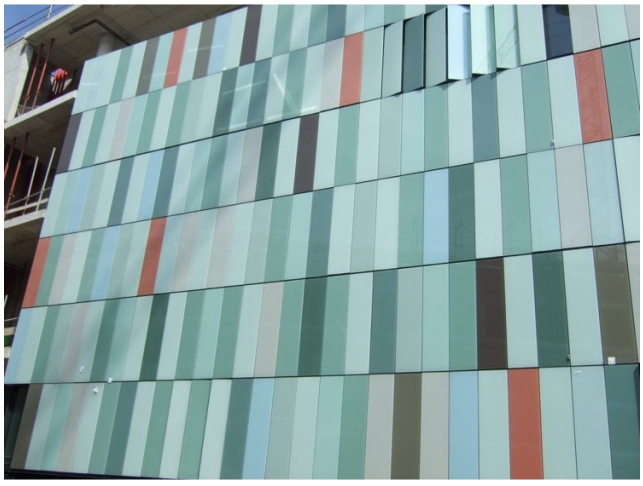


Fig. 32. Matthias Sauerbruch and Louisa Hutton, Area "Mac567", *Maciachini Center*, Milan (©Courtesy of AluK).

REFERENCES

[1] T. Herzog, R. Krippner, W. Lang, *Fassaden Atlas*, Institut für Internationale Architektur - Documentation, Münich, 2004.

[2] S. Deswarte, B. Lemoine, *Advanced Building Systems. A Technical Guide for Architects and Engineers*, Birkhäuser, Basel, 2003.

[3] S. Murray, *Contemporary Curtain Wall Architecture*, Princeton Architectural Press, New York, 2009.

[4] M. Nastri, *Future Façade Systems*, in I. Paoletti, M. Nastri, Editors, *Material Balance. A Design Equation*, Springer, Cham, 2021, pp. 83-103.

[5] A. Watts, *Modern Construction Envelopes*, Springer, Wien-New York, 2005.

[6] C. Schittich, *Building Skins*, Birkhäuser, Basel-Boston-Berlin, 2001.

[7] U. Knaack *et alii*, *Façades. Principles of Construction*, Birkhäuser, Basel, 2007.

[8] T. Herzog, R. Krippner, W. Lang, *Facade Construction Manual*, Birkhäuser, Basel-Münich, 2008.

[9] E. Oesterle *et alii*, *Double-Skin Facades*, Prestel, Münich, 2001.

[10] M. Nastri, *Involucro e architettura*, Maggioli, Santarcangelo di Romagna, 2008.

[11] K. Boswell, *Exterior Building Enclosures. Design Process and Composition for Innovative Facades*, Wiley & Sons, Hoboken, NJ., 2013.

[12] J. Lovell, *Building Envelopes. An Integrated Approach*, Princeton Architectural Press, New York, 2010.

[13] C. Schittich, *Solar Architecture. Strategies, Visions, Concepts*, Birkhäuser, Basel-Boston-Berlin, 2003.

[14] A. Aksamija, *Sustainable Facades. Design Methods for High-Performance Building Envelopes*, Wiley & Sons, Hoboken, NJ., 2013.

[15] A. A. Hamid, *Design and Retrofit of Building Envelope*, Bookbaby, Pennsauken, NJ., 2012 4th Ed.

[16] A. Syed, *Advanced Building Technologies for Sustainability*, Wiley & Sons, Hoboken, NJ., 2012.

[17] K. Konis, S. Selkowitz, *Effective Daylighting with High-Performance Facades. Emerging Design Practices*, Springer, Wien-New York, 2017.

[18] G. Hausladen, M. de Saldanha, P. Liedl, *ClimateSkin. Building-skin Concepts that Can Do More with Less Energy*, Birkhäuser, Basel, 2008.

[19] M. H. Haeusler, *Media Facades. History, Technology, Content*, AV Edition, Ludwigsburg, 2009.

[20] S. Murray, *Translucent Building Skins. Material innovations in modern and contemporary architecture*, Routledge, New York, 2013.

[21] R. Fortmeyer, C. D. Linn, *Kinetic Architecture. Designs for Active Envelopes*, The Images Publishing Group, Mulgrave, 2014.

[22] J. Moloney, *Designing Kinetics for Architectural Facades. State Change*, Routledge, New York, 2011.

[23] M. H. Haeusler, G. Tscherteu, M. Tomitsch, *New Media Facades. A global survey*, AV Edition, Ludwigsburg, 2012.

Social Capital and Affordable Housing for Vulnerable Housing

Patricia Lopez

***Abstract*—**Vulnerable households often do not have the resources to meet their housing needs and are forced to seek help from others for basic services or insecure housing. The government has generally covered the housing needs of vulnerable households with rules and regulations designed to identify and help low-income people (Jencks, 1992). This support generally requires low-income households to meet eligibility criteria and follow the rules of the government that supports them. This approach has not solved the problems of secure housing for vulnerable households and can generate dependency on the support system by reproducing the housing insecurity of vulnerable households (Jencks, 1992; Schiller, 2001) mainly in developing countries like Mexico.

Programs that promote community participation and generate decision-making processes in which community members participate can provide more sustainable positive outcomes for housing for vulnerable households (Jencks, 1992; Mancini & Marek, 2004; Schiller, 2001). The work on social capital explains how residents of vulnerable households can affect secure housing outcomes.

Housing problems for vulnerable households such as financial pressure, funds for construction, limited government operational capacity, and cultural isolation represent limits to access to safe housing for vulnerable households. An alternative to overcome the problems of the government construction model is the social capital approach, it represents the mechanism that involves the community in the building construction of affordable housing (Friere, 1994; Medoff and Sklar, 1994). This paper describes housing access alternatives for vulnerable households in México, analyzes the role and comparative advantages of social capital in access to secure housing, and compares possible risks and alternatives to vulnerable populations. The results describe how households have successfully mobilized their resources and used their collective power to influence the direction and decisions that affect their neighborhoods or communities.

***Keywords*—**social capital and community, affordable housing, comparative analysis, government intervention and poverty trap.

Patricia López is with the Ibero-American University, Mexico (e-mail: patyloro@gmail.com).

Prodrug-Based Intracellular Delivery of Zerumbone for Breast Cancer Therapy

Eltayeb E.M. Eid

Department of Pharmaceutical Chemistry and Pharmacognosy, Unaizah College of Pharmacy, P.O.BOX 5888 Unaizah 51911, Qassim University, Saudi Arabia
Email: eem.eid@qu.edu.sa

The idea

The idea relates to breast cancer therapy in an animal model. More specifically is for breast cancer targeting therapy by releasing zerumbone (ZER) as a toxic agent intracellularly through a prodrug strategy using cell-penetrating peptide as a carrier.

Brief description of the idea

A prodrug for treating breast cancer or inhibiting metastasis using an animal model is proposed. The prodrug is prepared by conjugating iRGD with ZER. iRGD is a cell-penetrating peptide that has amino acids sequence (CRGDK/RGPD/EC), whereas ZER **(Fig 1.),** is a sesquiterpene compound isolated from tropical ginger. The conjugation will be through an acid-labile or pH-responsive linker. Examples of acid-labile linkers include imine group, hydrazone, cis-aconityl, ketal bond, acetal bond, carboxylic hydrazone bond, and trityl bond. The conjugated delivery system, as a prodrug, will be internalized into the cancer cell via the attachment of iRGD into the αβ integrin receptor on the cancer cell surface. The internalization of the designed prodrug is investigated by the entry mechanism and intracellular process of iRGD-ZER into breast cancer tumor cells to determine whether iRGD carrier peptide could disturb the cancer cell biochemistry by which ZER is endocytosed. Endocytosis is a well-reported entry mechanism for different nanoscale drug carriers. Both iRGD and ZER will be labeled with two different fluorescent dyes for tracking the cleavage of iRGD-zerumbone bond. The intracellular release of ZER as a toxic agent will be in the late endosome process. In this process, the pHi, being 4.0-6.5, enables the cleavage of the linker between iRGD and ZER. When the zerumbone is inside the cancer cell, it will disturb the cell cycle through the elevation of caspase three or/and also inhibiting the activation of NF-κB, resulting in cancer cell apoptosis and stopping the proliferation, respectively. The iRGD will be degraded and digested in the lysosome process. The illustration in **Fig. 2**. explains the clear delivery system pathway.

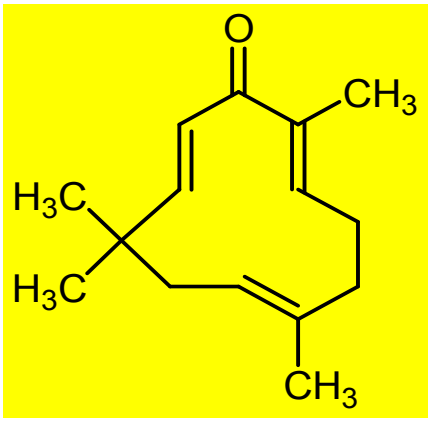


Fig. 1: The structure of zerumbone (ZER)

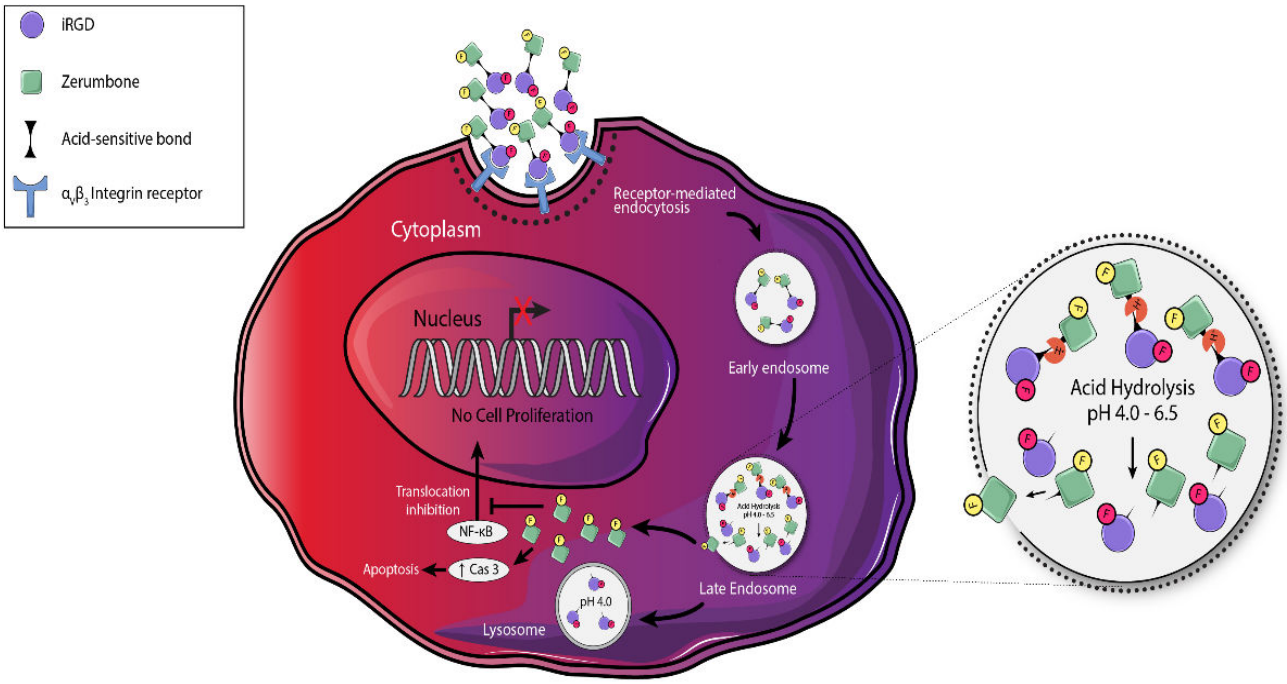


Fig. 2: Zerumbone (ZER) (cargo, indicated as a green square) will be conjugated with iRGD, a cell-penetrating peptide (carrier, indicated as a purple ball). The conjugation will be through an acid-labile linker or pH-responsive cleavable bond (indicated as a black line). Both iRGD and ZER should be labeled precisely by two different fluorescent dyes (indicated as a circle F) to track the cleavages and their sites in the intracellular regions and to validate the traceless ZER in the extracellular compartments. Also, to avoid an altered cytotoxic activity as the result of ZER traces, the acid-labile linkers that will be used for conjugation should be stable at pHe range 6.5-7.4. The internalization of the conjugated delivery system onto the cancerous cell will be through receptor-mediated endocytosis, in which the iRGD peptide (carrying ZER) binds to αvβ3

Integrin receptors in cancer cell surface that is usually overexpressed in the cancer cells more than the normal cells. The release of ZER as a toxic agent will take place in the late endosome process by exploiting the pHi value range 4.0 - 6.5, that considered being as a unique physical property for acid liable linker's cleavage. Then, ultimately, ZER will stop cell proliferation, or in a more explicit scientific term, it will disturb the cancer cell biochemistry.

Deep learning with Noisy Labels : Learning True Labels as Discrete Latent Variable

Azeddine El-Hassouny, Chandrashekhar Meshram, Geraldin Nanfack

Azeddine El-Hassouny, Geraldin Nanfack are with ENSIAS, Mohammed V University in Rabat, BP 713, Agdal Rabat, Morocco

Chandrashekhar Meshram is Post Graduate Studies and Research in Mathematics Department, Jayawanti Haksar Government Post Graduation College, College of Chhindwara University, Betul, M.P., India 460001

Abstract

In recent years, learning from data with noisy labels (Label Noise) has been a major concern in supervised learning. This problem has become even more worrying in Deep Learning, where the generalization capabilities have been questioned lately. Indeed, deep learning requires a large amount of data that is generally collected by search engines, which frequently return data with unreliable labels. In this paper, we investigate the Label Noise in Deep Learning using variational inference. Our contributions are : (1) exploiting Label Noise concept where the true labels are learnt using reparameterization variational inference, while observed labels are learnt discriminatively. (2) the noise transition matrix is learnt during the training without any particular process, neither heuristic nor preliminary phases. The theoretical results shows how true label distribution can be learned by variational inference in any discriminate neural network, and the effectiveness of our approach is proved in several target datasets, such as MNIST and CIFAR32.

© 2015 Published by Elsevier Ltd.

KEYWORDS:
Label Noise, Deep learning, variational inference, MNIST, CIFAR32.

1. Introduction

The growing of large amounts of data in recent years, and the development of computational capabilities, such as GPUs have propelled the era of Deep Learning [1], which is another facet of machine learning based essentially on artificial neural networks.

Deep Learning showed impressive results, especially in the case of supervised classification. However, it requires large amounts of data that are usually collected and then annotated.

Data collection and annotation are carried out either by search engines or by experts. Search engines often return data with Label Noise, the experts themselves may make errors while annotating the data.

As in machine learning, Label Noise considerably reduces the performance of Deep Learning, of which these generalization capacities have an enormous sensitivity [2].

Although some approaches have been proposed in the literature, almost no method provides a sufficiently robust solution. We, therefore, propose a reclassification of these methods, showing how the noisy label can be corrected, by learning the true label using variational inference and discrete latent variable.

Due to the difficulty to re-annotate collected dataset by engines, because of time-consuming, and even if there is an expert annotation, experts can produce errors, deep learning is facing a crucial problem of performance in the case of noisy labels.

In the literature, many research works handling label noise in deep learning techniques have been proposed. However, some latest research focus on specific domains such as medical image analysis, plants diseases, paraphrase identification, etc. Karimi et al. [3] conducted a state of the art on handling label noise in deep learning for medical image. In their study, they exploit three types of label noise and use many

*Azeddine Elhassouny (Corresponding author) (email:azeddine.elhassouny@ensias.um5.ac.ma).
[1]Chandrashekhar Meshram (email:csmeshram84pdf@gmail.com).
[2]Geraldin Nanfack (email:nanfackg123@gmail.com).

approaches to reduce its negative impact. Li et al. [4] developed a deep paraphrase identification model to robust deep learning base model on dataset with corrupted labels. By exploiting the selection of features and the low-rank matrix recovery in the learning procedure, Wang et al. [5] proposed multi-label classifiers based on low-rank constraints (ML-LRC) approach, which handle label noise in multi-label learning. It proved that the impact of label noise depends on type and design of learning model [3]. To decrease the impact of label noise, Zhang et al. [6] proposed a loss function called Rectified Cross-Entropy, which let learnimng model focus on true label and ignoring noisy one.

Many works of literature have attempted to solve this problem, but many of them have proposed heuristic approaches or descriptive statistics based on noise estimation. Furthermore, several variational inference techniques [7] were proposed to learn probability distributions, In addition, variational inference with discrete latent variables [8] is used to generate images (Variational AutoEncoders with discrete latent variables). Besides Bayesian statistics have shown its success in uncertainty modeling.

In this work, we study the label noise for the deep CNNs with the aim of enhancing the learning ability from noisy labeled data. The proposed approach is generally consists on Bayesian approach.

The idea is to improve learning by using variational inference on the latent variables which represent the true labels. That data is presumed to have label noise, and we use any neural network architecture. The directed acyclic graph, as simplest graphical probabilistic model (Figure 3), shows how modeling the true labels, the inputs, the labels observed in a Bayesian Network.

1.1. Motivation

Even if many researches treat label noise problem in deep learning, but many of them have proposed heuristic approaches or descriptive statistics based on noise estimation. These approaches have sometimes been successful, it remains that the majority of them are very heuristic, this doesn't fake their reusability in much more complex tasks. However, the variational inference is a serious track to learn the probability distributions of the real labels [2]. And Bayesian statistics have shown its success in uncertainty modeling. The problem posed by variational inference is that the derivative of discrete latent variables with respect to parameters does not make sense, because we cannot think about doing on derivative discrete variables since they are not at least continuous.

The contributions of this article are summarized as follow. We introduce TLasDLV (Learning Discrete Latent Variables by Reparametrization) approach, that shows how to learn the true label, as a discrete hidden variable using variational inference, without any generative model. As well as, methods to deal with stochastic gradients of Evidence Lower Bound (ELBo). Then, we have proposed ELBo of TLasDLV, as determinist and efficient algorithm to deal with label noise.

The strengths of the procedure are the following: Our proposed solution is not the best when label noises are very low, but when they become high, our method outperforms previous works. The results are quite impressive, the higher the noise rate, the stronger the regularization of accuracy imposed by our methods, to robustify the learning.

The weakness is our proposed method is not the best when label noises are very low.

The remainder of this paper is organized in the following manner. Section 2, 3 and 4 introduces notions of Label Noise in Deep Learning followed by a state of art of Label Noise in Deep Learning, and this section mainly discusses the methodology to deal with label noise along with related concepts and the proposed approach.

Section 5 dedicates to the algorithm experiments; multiple experiments are conducted and the experimental results are evaluated as well as comparative analysis. Finally, the paper is summarized in Section 6.

2. Fundamentals and Related Work

2.1. Label Noise

Benoît and Michel [9] have made a state of the art of Label Noise, summarizing proposed techniques in the literature. According to them, Label Noise is defined as a stochastic process that pollutes labels of a training sample before being presented to the learning algorithm [10]. The Label Noise can also be defined as anything that obscures the relationship between feature instances and their class.

The Label Noise can come from several sources [9]:

- *Insufficient information to the expert to properly label [11]:* in the process of labeling by the expert, some very important variables may be missing. But the expert continues the annotation without taking it into account; this may result in the Label Noise.
- *Errors due to data collection [12]:* when annotating, the expert often considers the values of variables; however, these variables are very often tainted with uncertainty. In addition, errors of frequent search engines are often found in this category.
- *Subjective annotation of data [13]:* Because the annotation of the data is very expensive at the current time, the expert may be prone to tiredness and unconfident and so he can annotate the data in a very subjective way.

- *Encoding and communication [14]:* for example, the data can be sometimes compressed or encoded with errors, and the expert is based on it to annotate.

2.2. Deep Neural Networks

Although Deep Neural networks's success has been recent, they are a powerful tool of artificial intelligence [15], [16]. The simplest learning model of artificial neural networks is Perceptron as shown in figure 1.

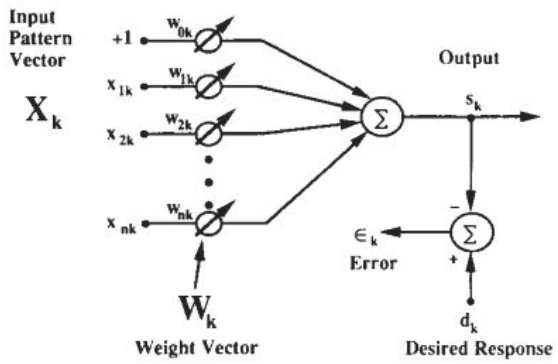


Fig. 1: The Perceptron Model described in [17]

It takes as input an instance of data $\mathbf{x} \in \mathbb{R}^D$, it is parameterized by weights $\mathbf{w} \in \mathbb{R}^D$ and a bias w_0, to produce an output $y(\mathbf{x}; \mathbf{w})$ using a real-valued activation function f. So the output of the perceptron is :

$$y(\mathbf{x}; \mathbf{w}) = f(\mathbf{w}^T\mathbf{x} + w_0)$$

The real $z = \mathbf{w}^T\mathbf{x} + w_0$ is usually called preactivation. The activation function is usually very discriminative as the sigmoidal function defined by $f(z) = \frac{1}{1+\exp(-z)}$. That was generally for a binary classification.
For a multi-classification, we use a generalized form of perceptron returning outputs like $\mathbf{y}(\mathbf{x}; \mathbf{w}) \in \mathbb{R}^C$ where C is the number of classes, and the model is given by:

$$\mathbf{y}(\mathbf{x}; \mathbf{w}) = f(W^T\mathbf{x} + \mathbf{w_0})$$

Where $W \in \mathbb{R}^{D\times C}$, $\mathbf{w_0} \in \mathbb{R}^C$, f is a vectorized form of the real function (so the real function is applied on each real value of the input vector or matrix). Generally, the activation function, in the context of multi classification, is a softmax given by

$$f(\mathbf{z}_k) = \frac{\exp(\mathbf{z}_k)}{\sum_{l=1}^{C}\exp(\mathbf{z}_l)}$$

, with $\mathbf{z} = W^T\mathbf{x} + \mathbf{w_0}$ being the preactivation vector.

Deep Learning takes effect when a large number of layers of neurons are applied to predict probabilities. To this end, there are also layers of neural networks that are very successful due to their high applicability. This is the case with convolutional layers, pooling, etc. Combination of these layers have given birth to several types of neural networks. In our study, it is not a question of naming all these layers or making a deep theory of them. But beyond that, Label Noise is an issue for any neural network classification architecture.

2.3. Label Noise in Deep Learning

The robustness of the Deep Neural Networks algorithm has been an interesting research topic due to the popularity of previous works in the field. Indeed, as we noticed, Benoît and Michel [9] have made an interesting state of the art about the subject. In that paper, the Label Noise is well explained and defined, a grouping of proposed techniques in the literature is also presented. According to them, Label Noise is defined as a stochastic process that pollutes labels of a training sample before being presented to the learning algorithm [10].

2.3.1. Imperfectly as Semi-Supervised Learning

Most of the successful methods have focused on so-called tolerant methods to robustifying neural networks, such as imperfectly as semi-supervised approaches and self-training approaches [9]. The aim of this type of method is to make a robust classifier based on training data, where clean labeled instances are distinct from noisy label ones. They assume the existence of a clean label instance, and a common approach of using the noisy sample as an unlabeled instance.

One of the first works was made by Yves and Bengio [18], in the case of semi-supervised learning, they proposed an estimation principle for any probabilistic classifier, which wants to benefit from the unlabeled instances during training.

Another study is made by Xiao et al. [19], where they proposed a general framework to train Deep Converts, with only a limited number of instances, with clean labels in the dataset. The obtained results showed that it works well on their own collected dataset Clothes 1M. The authors also found that forming a CNN from scratch with limited clean labels and massive noisy labels is better than tuning it on clean labels. Their paper is a generalization of the work of Sainbayar Sukhbaata [20], which introduced a linear layer to model the noise, but considere, given the true label, that the observed label is independent of instance data. So they proposed a model of noise distribution, which can depend on the instance, by introducing a latent variable that can take 3 different values (a binary random variable: noise-free, random noise, confusing noise), expressed as a graphical model, the training algorithm uses Expectation–maximization(EM).

Giorgio et al. [21] proposed two procedures to correct the loss of deep neural networks, by adapting a noise estimation in the case of multiclass probabilities predictions. These two procedures consist of estimating a multiclass extension of noise estimation called "backward" and "forward".

Besides, Vahdat et al. [22] presented a framework based on an undirected graphical model using a semi-supervised manner. This methods consist on conditional random fields (CRF), which modeled the relation dependence between clean labels and noisy labels.

2.3.2. Imperfectly as Self-Training

The aims of this kind of method is to build an algorithm either by correcting the loss or with some noise estimation models. That method correct the noisy labels and conserve the clean labels during the learning, this approach doesn't assume any existence of clean labels in the dataset (even though they exist).

Volodymyr and Hinton [23] proposed robust loss functions to label noise, and then applied them to label Aeria Images, or more generally to perform Semantic Segmentation. in this case, images are often incomplete and mislabeled pixels. The proposed model consists of a contamination model by asymmetric noise of the true observed class. These parameters are fixed and are modelled as bernoulli contamination model.

In 2015, Sainbayar et al. [20] investigated the performance of discriminative Deep Convnets, trained on noisy data. In order to match the noisy label distribution, they introduced an extra noise layer into the network. They consider multiclass classification, and assume more realistic asymmetric label noise. They don't suppose the availability of any clean label, as in the case of semi-supervised learning, where some data has clean labels and the rest are either unlabeled or have unreliable labels. The model of noise is done by a stochastic matrix, in order to map the true probabilistic distribution, learned by a neural network, to the noisy probabilistic distribution.

Inspired by bootstrapping, Reed et al. [24] proposed a generic way to deal with noisy and incomplete labeling by increasing the prediction objective with a notion of consistency.Learning takes place in two phases: start by creating a classifier, using examples with the reliable label, and then learn how to iteratively classify unlabeled instances. This method is relying on using a reconstruction error, called consistency objective, which dynamically updates the prediction objective, and explicitly models the noise distribution, as a matrix mapping model predictions to trainiKuang-Hueing labels. It also uses a convex combination of training labels and the prediction.

In another work, Daiki et al. [25] proposed a joint optimization framework, which can correct labels by updating labels during the training. They also found that Deep Learning doesn't memorize the noisy labels and maintains high performance for clean data under a high learning rate.

In another similar work, Lee et al.[26] proposed a framework CleanNet, which can be coupled with a deep CNN and performs classification by learning on noisy labels. This method relies on the introduction of a variable, which can take 3 values depending on whether the image is relevant to noisy labels, or whether the image is mislabeled, or whether the verification is not available. In that approach, the probability of flipping label is also fixed, and they used the features learned from images according to their distribution on being in a particular class to predict the true label.

2.4. Summary of successful Deep Learning methods for Label Noise

We summarize the main previous works of the label noise in deep learning in the following table 1.

Methods and core intuition	Type	Tasks	Dataset of experiments
Robust Lost functions with fixed noise rates [23]	Self-Training	Semantic Segmentation	Aerial Image dataset [27]
Extra Noise Layer [20]	Self-Training	Classification for discriminative Neural Networks	SVHN [28], CIFAR-10 [29], Imagenet [30]
Probabilistic Graphical Model with one latent variable trained with EM [19]	Imperfectly as Semi-Supervised Learning	Classification for discriminative Neural Networks	CIFAR-10 [29], clothes1M [19]
Bootstrapping: introducing consistency, reconstruction loss [24]	Self-Training	Detection and Classification for discriminative Neural Networks	Toronto Faces [31], MNIST [32], Imagenet [30]
Loss correction with 2 procedures and estimation of transition matrix [21]	Self-Training	Classification for discriminative Neural Networks	IMDB [33], MNIST [32], CIFAR-10 and CIFAR-100 [29]
Deep Neural Networks with CRF [22]	Imperfectly as Semi-Supervised Learning	Detection and Classification for discriminative Neural Networks	Microsoft COCO [34], CIFAR-10 [29]
CleanNet [35]	self-Training	Classification for discriminative Neural Networks	clothes1M [19], WebVision [36], Food-101 [37]
Joint optimization (network and labels) [25]	self-Training	Classification for discriminative Neural Networks	clothes1M [19], CIFAR-10 [29]

Table 1: Summary of successful deep learning methods for Label Noise

3. Requirements

This section describes the prerequisites for our proposed method and their components, a general description of the variational inference, and then the variational inference on discrete latent variables.

3.1. Variational Bayesian Inference [38], [39] and [40]

Even if, these approaches have sometimes been successful, it remains that the majority of them are very heuristic, this doesn't fake their reusability in much more complex tasks. However, the variational inference that becomes more and more utilisable is a serious track to learn the probability distributions of the real labels.

Variational inference is the core of modern Bayesian methods and graphical probabilistic models. It is very popular and finds its applicability in many inference problems, such as natural language processing, speech processing, computer vision, etc. Its basic idea : **turn probabilistic inference problem into an optimization problem.**

Let assume some observations $\mathbf{x} = x_{1:D}$, and we suppose that some $\mathbf{z} = z_{1:m}$ are hidden variables that can

generate those observations. The inference problem is to estimate the probability $p(\mathbf{z}|\mathbf{x}) = \frac{p(\mathbf{x},\mathbf{z})}{p(\mathbf{x})}$. Unfortunately, generally this posterior is intractable because of the denominator $p(\mathbf{x}) = \int p(\mathbf{x}, \mathbf{z})d\mathbf{z}$. Then, the idea is to suppose a **variational family** over the latent variables : $q(\mathbf{z}; \nu)$.

The goal would be to fit **variational parameters** ν to be close to the exact posterior $p(\mathbf{z}|\mathbf{x})$ in term of Kullback Leibler divergence.

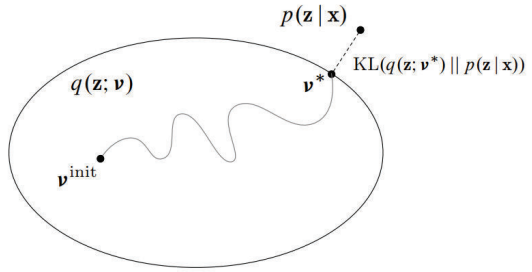


Fig. 2: Variational Inference Intuition [39]

3.2. The Evidence Lower Bound: ELBo

It can be simply written as an optimization problem:

$$\begin{aligned} q(\mathbf{z}; \nu^*) &= \arg\min_{q(\mathbf{z};\nu)} KL(q(\mathbf{z}; \nu)||p(\mathbf{z}|\mathbf{x})) \\ &= \arg\min_{q(\mathbf{z};\nu)} \int q(\mathbf{z}; \nu) \log\left(\frac{q(\mathbf{z};\nu))}{p(\mathbf{z}|\mathbf{x})}\right)d\mathbf{z} \end{aligned} \tag{1}$$

Unfortunately, this objective function is not computable, due to the fact that the second term depends also on the intractable distribution $p(\mathbf{z}|\mathbf{x})$.

On the other hand, it is easy to show that :

$$\begin{aligned} &\log(p(\mathbf{x})) = \log\left(\int p(\mathbf{z}, \mathbf{x})d\mathbf{z}\right) \\ &= \log\left(\int \frac{q(\mathbf{z};\nu)p(\mathbf{z},\mathbf{x})}{q(\mathbf{z};\nu)}d\mathbf{z}\right) \\ &\geq \int q(\mathbf{z}; \nu) \log\left(\frac{p(\mathbf{z},\mathbf{x})}{q(\mathbf{z};\nu)}\right)d\mathbf{z} \text{ by the Jensen Inequality} \\ &= E_{q(\mathbf{z};\nu)}\Big(\log(p(\mathbf{z}, \mathbf{x})) - \log(q(\mathbf{z}; \nu))\Big) \\ &= ELBo(\nu) \end{aligned} \tag{2}$$

The right term is called **Evidence Lower Bound (ELBo)**. It depends on the parameter ν and the distribution $p(\mathbf{z}, \mathbf{x})$.

We can easily show that, minimize Eq.1 is equivalent to maximizing the ELBo by :

$$\begin{aligned} &\log(p(\mathbf{x})) = \int q(\mathbf{z}; \nu) \log(p(\mathbf{x}))d\mathbf{z} \\ &= \int q(\mathbf{z}; \nu) \log\left(\frac{q(\mathbf{z};\nu)p(\mathbf{z},\mathbf{x})p(\mathbf{x})}{q(\mathbf{z};\nu)p(\mathbf{z},\mathbf{x})}\right)d\mathbf{z} \\ &= \int q(\mathbf{z}; \nu) \log\left(\frac{q(\mathbf{z};\nu)p(\mathbf{z},\mathbf{x})}{q(\mathbf{z};\nu)p(\mathbf{z}|\mathbf{x})}\right)d\mathbf{z} \\ &= \int q(\mathbf{z}; \nu) \log\left(\frac{p(\mathbf{z},\mathbf{x})}{q(\mathbf{z};\nu)}\right)d\mathbf{z} + \int q(\mathbf{z}; \nu) \log\left(\frac{q(\mathbf{z};\nu))}{p(\mathbf{z}|\mathbf{x})}\right)d\mathbf{z} \\ &= E_{q(\mathbf{z};\nu)}\Big(\log(p(\mathbf{z}, \mathbf{x})) - \log(q(\mathbf{z}; \nu))\Big) + KL(q(\mathbf{z}; \nu)||p(\mathbf{z}|\mathbf{x})) \\ &= ELBo(\nu) + KL(q(\mathbf{z}; \nu)||p(\mathbf{z}|\mathbf{x})) \end{aligned} \tag{3}$$

Because the Kullback Leibler(KL) divergence is always positive, minimizing KL is equivalent to make it null and thus maximizing the ELBo.

ELBo has the great advantage of being easily computable, but poses a serious problem for gradients. Several algorithms exist to approximate the gradient.

3.2.1. Methods to deal with stochastic gradients of ELBo

Let us recall the ELBo:

$$\mathfrak{L}(\nu) = ELBo(\nu) = E_{q(\mathbf{z};\nu)}\Big(\log(p(\mathbf{z}, \mathbf{x})) - \log(q(\mathbf{z}; \nu))\Big) \tag{4}$$

If we define $g(\mathbf{z}, \nu) = \log(p(\mathbf{z}, \mathbf{x})) - \log(q(\mathbf{z}; \nu))$, then

$$\begin{aligned} &\nabla_\nu \mathfrak{L}(\nu) = \nabla_\nu \int q(\mathbf{z}; \nu)g(\mathbf{z}, \nu)d\mathbf{z} \\ &= \int \Big(g(\mathbf{z}, \nu)\nabla_\nu q(\mathbf{z}; \nu) + q(\mathbf{z}; \nu)\nabla_\nu g(\mathbf{z}, \nu)\Big)d\mathbf{z} \\ &= \int \Big(g(\mathbf{z}, \nu)q(\mathbf{z}; \nu)\nabla_\nu \log(q(\mathbf{z}; \nu)) + q(\mathbf{z}; \nu)\nabla_\nu g(\mathbf{z}, \nu)\Big)d\mathbf{z} \\ &\text{because } \nabla_\nu \log(q) = \frac{\nabla_\nu q}{q} \\ &= E_{q(\mathbf{z};\nu)}\Big(g(\mathbf{z}, \nu)\nabla_\nu \log(q(\mathbf{z}; \nu)) + \nabla_\nu g(\mathbf{z}, \nu)\Big) \end{aligned} \tag{5}$$

There are some algorithms to approximate this gradient. However, reparametrization, as defined in [41], is a simplest tricks to do so.

Pathwise Gradients of ELBo:

It uses a reparametrization trick. Its idea is to suppose that : $\mathbf{z} = t(\epsilon, \nu)$ with $\epsilon \sim s(\epsilon)$ and then $\mathbf{z} \sim q(\mathbf{z}; \nu)$ Which means :

$$\begin{aligned} &\epsilon \sim \text{Normal}(0, 1) \\ &z = \epsilon\sigma + \mu \\ &\text{and then } z \sim \text{Normal}(\sigma, \mu) \end{aligned} \tag{6}$$

In this case, $\log(p(\mathbf{z}, \mathbf{x}))$ and $\log(q(\mathbf{z}))$ will be differentiable with respect to $\mathbf{z}$. Rewriting Eq.5 by replacing $\mathbf{z} = t(\epsilon, \nu)$, which gives:

$$\nabla_\nu \mathfrak{L}(\nu) = E_{s(\epsilon)}\Big(g(\mathbf{t}(\epsilon, \nu), \nu)\nabla_\nu \log(s(\epsilon)) + \nabla_\nu g(\mathbf{t}(\epsilon, \nu), \nu)\Big) \tag{7}$$

Finally, the reparametrization trick can be written as :

$$\begin{aligned} &\nabla_\nu \mathfrak{L}(\nu) = E_{s(\epsilon)}\Big(g(\mathbf{t}(\epsilon, \nu), \nu)\nabla_\nu \log(s(\epsilon)) + \nabla_\nu g(\mathbf{t}(\epsilon, \nu), \nu)\Big)\nabla\mathfrak{L}(\nu) \\ &= E_{s(\epsilon)}\Big(\nabla_\nu g(\mathbf{t}(\epsilon, \nu), \nu)\Big) \\ &= E_{s(\epsilon)}\Big(\nabla_\mathbf{z}(\log(p(\mathbf{z}, \mathbf{x})) - \log(q(\mathbf{z}, \nu)))\nabla_\nu t(\epsilon, \nu) - \nabla_\nu \log(q(\mathbf{z}, \nu))\Big) \\ &= E_{s(\epsilon)}\Big(\nabla_\mathbf{z}(\log(p(\mathbf{z}, \mathbf{x})) - \log(q(\mathbf{z}, \nu)))\nabla_\nu t(\epsilon, \nu)\Big) \end{aligned} \tag{8}$$

This last formula is easily computable and produces low variance gradients.

4. Learning Discrete Latent Variable by Reparametrization

There are not enough studies on Learning Discrete Latent Variables by Reparametrization in variational inference. In facts, it just recently that [8] and [42] showed how learning hidden discrete variables by introducing a **Relax-One-Hot-Categorical distribution**, as well as categorical reparameterization with Gumbel-Softmax.

The problem posed by variational inference is that the derivative of discrete latent variables with respect to parameters doesn't make sense, because it impossible to calculate derivatives with respect of derivative discrete variables since they are not at least continuous. So [8] and [42] proposed a way to introduce a continuous distribution, that can map a discrete distribution by softmax activation.

Let be $\mathbf{z} \sim \text{Cat}(\pi_\nu)$ a categorical variable (a one-hot representation) parameterized by $\pi_\nu \in \mathbf{\Delta_{C-1}}$(the simplex) $\subset \mathbb{R}^C$.

And we want to sample this distribution by a reparameterization trick.Furthermore, a random variable G is said to have a standard Gumbel distribution if :

$$G = -\log(-\log(U)) \text{ whith } U \sim \text{Unif}(0,1) \tag{9}$$

Then, by the property that if $\epsilon_k \sim \text{Gumbel}(0,1) = -\log(-\log(u_k))$ whith $u_k \sim \text{Unif}(0,1), k = 0, ..., C-1$ are random variables, we got:

$$\mathbf{z} = \mathbf{one\_hot}\Big(\arg\max_k (\epsilon_k + \log(\pi_{\nu k}))\Big) := g(\epsilon, \nu) \tag{10}$$

However, due to the fact that **argmax** is not differentiable with respect to ν, the previous authors introduced **Gumbel-Softmax trick**. So instead of having a discrete valued random variable $\mathbf{z}$, they proposed to consider a softmax-map

$$f_\tau(\mathbf{z})_k = \frac{\exp(z_k/\tau)}{\sum_{k=0}^{C-1}\exp(x_k/\tau)}$$

So instead of obtaining $\mathbf{z}$ after sampling ϵ, one obtain :

$$y_k = \frac{\exp\Big((\epsilon_k + \log(\pi_{\nu k}))/\tau\Big)}{\sum_{k=0}^{C-1}\exp\Big((\epsilon_k + \log(\pi_{\nu k}))/\tau\Big)} \text{ for } k = 0, ..., C-1 \tag{11}$$

with τ is the **temperature parameter**.

For low temperatures (around 0) the Gumbel-Softmax distribution becomes identical to the distribution of $\mathbf{z}$

Finally, we can write

$$y = g(\nu, \epsilon)$$

with g is differentiable. And

$$p_{\nu,\tau}(y_0, y_1, ..., y_{C-1}) = \Gamma(C)\tau^{C-1}\Big(\sum_{k=0}^{C-1}\pi_{\nu_k}/y_k^{-\tau}\Big)^{-C}\prod_{k=O}^{C-1}(\pi_{\nu_k}/y_k^{\tau+1}) \tag{12}$$

For non-zero temparatures, a Gumbel-softmax variable $\mathbf{y}$ does not follow z the one-hot-categorical distribution. If in the forward pass, $\mathbf{z}$ is keept, there is no problem, but in order to backpropagate the gradient, there is the need to use $\mathbf{y}$ which is continuous, in the backward pass, by approximating $\nabla_\theta \mathbf{z} \simeq \nabla_\theta \mathbf{y}$: this is called **Straight-Through Gumbel Softmax Estimator**.

4.1. Learning True Labels as Discrete Latent Variable

This section shows how with variational inference, one can learn the true label, as a discrete hidden variable without any generative model. We call the approach **TLasDLV**.

4.2. ELBo of TLasDLV

Variables of a machine learning model or algorithm can be represented by a probabilistic graphical model, as shown in figure 3.

We assume that the data $(\mathbf{x}_n, \tilde{\mathbf{y}}_n), n = 1...N$ is generated through a distribution $p(\mathbf{x}, \tilde{\mathbf{y}})$. But our goal is to learn the distribution $p(\tilde{\mathbf{y}}|\mathbf{x}; \theta)$ discriminatively, and also learn the non observable variable, which is for $\mathbf{y}$ supposed generated by a distribution $p(\mathbf{y}|\mathbf{x}; \phi)$.

More clearly, the figure 3a shows that given $\mathbf{x}$, we can get the true label $\mathbf{y}$ by the parameter ϕ, and by just having the label we can obtain the observed label $\tilde{\mathbf{y}}$ with a parameter θ. This is a simple case where the probability of making an error by an expert depends only on the true label. The figure 3b shows that given $\mathbf{x}$, we can get the true label $\mathbf{y}$ by the parameter ϕ, and but we need sometimes $\mathbf{x}$ and $\mathbf{y}$ to get the observed label $\tilde{\mathbf{y}}$ with a parameter θ.

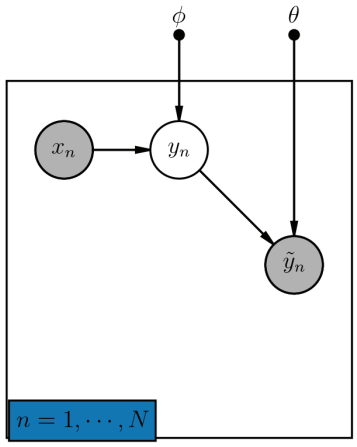
Anyway, if we consider $\mathbf{y}$ as a hidden variable (like $\mathbf{z}$) in 1, we can show by replacing on Eq. 3 that :

$$\begin{aligned}&\log(p(\tilde{\mathbf{y}}|\mathbf{x}))\\ &\geq E_{q(\mathbf{y}|\mathbf{x};\phi)}\Big(\log(\textstyle\sum_{y=1}^{C} p_\theta(\tilde{\mathbf{y}}|\mathbf{y}, \mathbf{x}))) - KL(q(\mathbf{y}|\mathbf{x};\phi)||p(\mathbf{y}|\mathbf{x})\Big)\\ &= ELBo(\phi)\end{aligned} \tag{13}$$

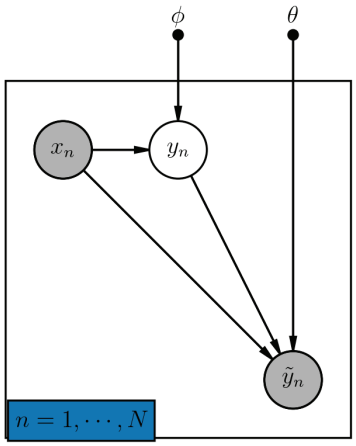
The first term $p_\theta(\tilde{\mathbf{y}}|\mathbf{y}, \mathbf{x}))$ is generally the **reconstruction probability** as (his equivalent in the basic variational inference is $p(\mathbf{x}|\mathbf{z})$) : it learns to get the observed label $\tilde{\mathbf{y}}$ given $\mathbf{x}$ and $\mathbf{y}$. It is like the ***likelihood***.

The second term distribution is $q(\mathbf{y}|\mathbf{x}; \phi)$ which is able to learn the exact intractable posterior r$p(\mathbf{y}|\mathbf{x}, \tilde{\mathbf{y}})$ (its equivalent in basic variational inference is $q(\mathbf{z}|\mathbf{x}; \phi)$).

The last distribution is $p(\mathbf{y}|\mathbf{x})$, which is a fixed distribution, called **the prior** (Its equivalent in variational inference is $p(\mathbf{z})$).



(a) Noise At Random Probabilistic Graphical Model



(b) Noise Not At Random Probabilistic Graphical Model

Fig. 3: The Probabilistic Graphical Model

4.3. The discrete latent distribution

In this section, we discuss how we can compute those probability distributions.

The ELBo of Eq. 13 shows 3 probabilities.

- The first is $p(\tilde{\mathbf{y}}|\mathbf{y}, \mathbf{x})$, and it can be computed as :
  - Either by considering that given $\mathbf{y}, \tilde{\mathbf{y}}$ doesn't depends on $\mathbf{x}$ (naive approach), and then this probability can be viewed as a stochastic matrix $T_\theta(\tilde{\mathbf{y}}|\mathbf{y})$, one has to learn. However, learning that directly isn't easy. So instead of computing $log(\sum_{y=1}^{C} p(\tilde{\mathbf{y}}|\mathbf{y}, \mathbf{x})))$ with that form, we compute the cross-entropy using the following Eq..

$$log(\sum_{y=1}^{C} p(\tilde{\mathbf{y}}|\mathbf{y}, \mathbf{x}))) = \sum_{k=1}^{C} t_k \log\left(\left[\mathbf{z}T_\theta(\tilde{y}, \mathbf{z})\right]_k\right) \tag{14}$$

    where $\mathbf{t}$ is the observed one-hot label of the current training example, $\mathbf{z}T_\theta(\tilde{y}, \mathbf{z})$ is the matrix product between the C-column vector $\mathbf{z}$ and the $C \times C$ matrix $T_\theta(\tilde{y}, \mathbf{z})$, with $\mathbf{z}$ following a Gumbel-Softmax distribution;
  - Either by using another matrix like a muti-dense layer: this manner is more difficult practically.
  - The second distribution is $q(\mathbf{y}|\mathbf{x}; \phi)$, which is the **true label** distribution, that can be computed by Gumbel-Softmax (Eq. 15) distribution as mentioned previously.

$$\begin{aligned} &q(\mathbf{z}|\mathbf{x}; \phi) \\ &= \Gamma(C)\tau^{C-1}\left(\sum_{k=1}^{C} \pi_{\phi_k}/z_k^{-\tau}\right)^{-C} \prod_{k=1}^{C} (\pi_{\phi_k}/z_k^{\tau+1}) \end{aligned} \tag{15}$$

    ϕ is the parameter of the discriminative neural network.

- Finally, the prior $p(\mathbf{y}|\mathbf{x})$ can be set to be just the Uniform one-hot-categorical distribution.

5. Proposed Algorithm of TLasDLV

The TLasDLV algorithm consists of the following phases as in the pseudo-code below :

Algorithm 1 TLasDLV

Inputs $\{(\mathbf{x_n}, \tilde{\mathbf{y}}_\mathbf{n}); n = 1, ..., N\}$: the dataset with the noisy labels, ϕ: the parameters of the discriminative neural network architecture, to be optimized

Output ϕ : the parameter learned, θ : the parameter of the stochastic matrix learned.

1: Initialize θ of $T_\theta(\tilde{\mathbf{y}}, \mathbf{y})$ with stochastic matrix ($C \times C$)
2: **while** $i \leq T$ **do**
3: $y =$ *Sample from Gumbel-Softmax with parameters ϕ (using Eq.15 or Eq.12)*
4: Normalize to probabilities the values of the stochastic matrix $T_\theta(\tilde{\mathbf{y}}, \mathbf{y})$
5: $\tilde{\mathbf{y}} = \mathbf{y}T_\theta(\tilde{\mathbf{y}}, \mathbf{y})$
6: Update ϕ and θ by stochastic gradient descent on the Negative of ELBo (Eq.13 and Eq.14). =0

The flowchart of the algorithm 1 is shown in figure 4.

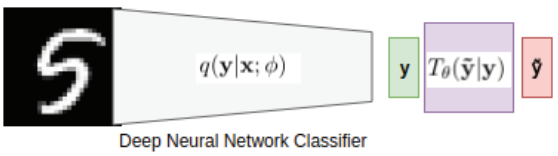


Fig. 4: Visualization of TLasDLV algorithm

6. Results and Discussion

The proposed algorithm TLasDLV is implemented using the Tensorflow framework. We first exploit the algorithm TLasDLV on the commonly used dataset. The results can be found as follow.

6.1. Datasets and Deep Neural Networks architectures

We have tested TLasDLV with MNIST [32] dataset for the classification of handwritten digits. The adopted architecture is 2 fully connected of 128, with dropouts of 0.5. We added noise to an asymmetric percentage with the disturbances of [21]. Those disturbances are : $2 \rightarrow 7; 3 \rightarrow 8; 5 \leftrightarrow 6; 7 \rightarrow 1$.

6.2. Results and Discussion

Learning without label noise is normally done without memorization using cross-entropy as a loss. The test and training precision curves, as shown in figure 5 are almost identical: it is the ability of neuron networks to learn from the data.

Once we apply a noise on the labels, as in the previous section, learning with cross-entropy becomes very sensitive to noise as in figure 6a. Even worse, the neural network tends to memorize labels. But with TLasDLV as seen in figure 6b, it is clear that the learning becomes more robust and our method is obviously not sensitive to these noises. Another advantage from this graph is the ability to regularize learning with TLasDLV.

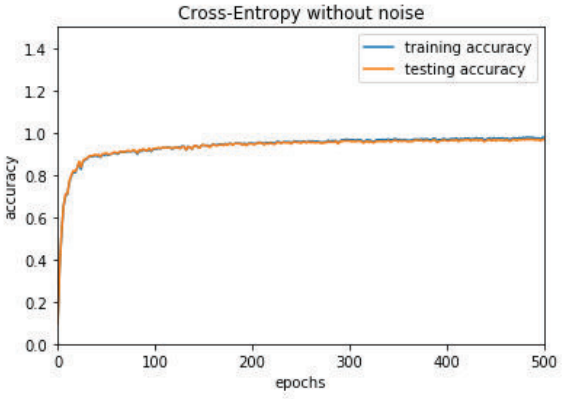
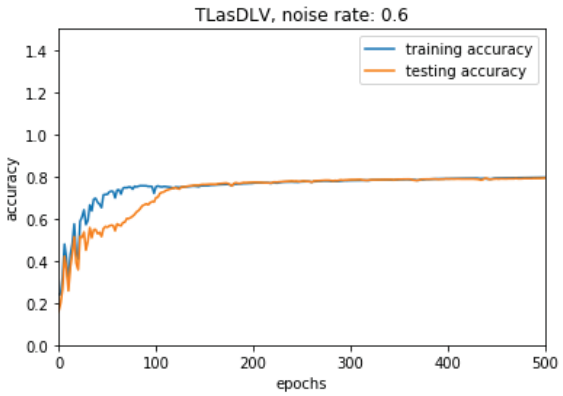


Fig. 5: Cross-Entropy with noise rate : 0.0

We also wanted to test if our method TLasDLV can work well in the case where there is no label noise. The curve is then depicted in the figure 7. In this figure 7a, the testing and training accuracy curves are almost identical in the case where there is no label noise. But more surprising, when there is label noise with a percentage of 0.4 figure 7b, our method TLasDLV is able to well robustify the learning by imposing a strong regularization of accuracy. The results are quite impressive.

(a) Cross Entropy with noise rate: 0.6



(b) True Label as Discrete Latent Variable with noise rate: 0.6

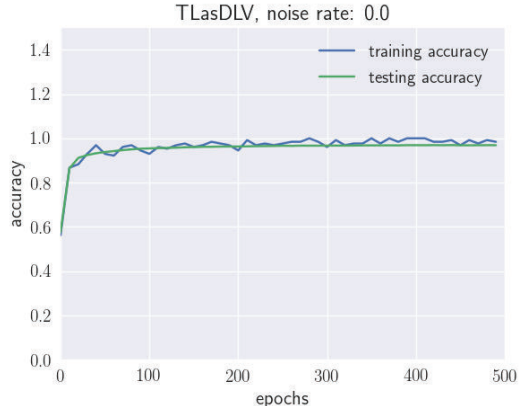
Fig. 6: 2-Fully-Connected-128: Cross-entropy Vs TLasDLV, noise rate: 0.6

By the end, we tried to visualize the stochastic transition matrix learned according to the noise level or noise rate. The resemblances whatever the level of noise are: firstly, in the first iterations, the matrix seems to privilege symmetrical noise, and yet no heuristic has been provided to it, nor constraint; secondly this transition matrix becomes more and more a rather sparse matrix converging towards a fixed transition matrix.

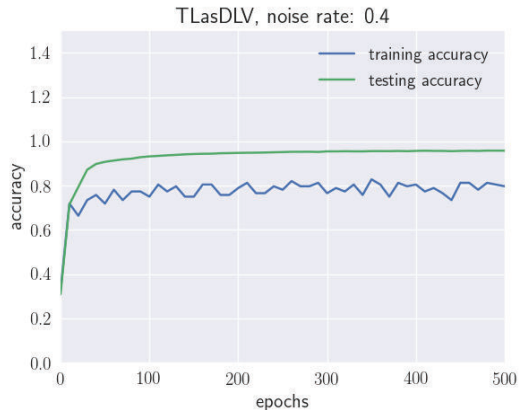
- In the case where the noise rate is zero (no label noise), the sub-figures of the figure 8 show us that the learned stochastic matrix tends to converge towards the identity matrix which is the real solution.

- In the case where the noise rate is 0.2, figure 9, the stochastic matrix tends slowly (comparing figure 8c and figure 9c) towards the identity matrix which is not the solution; we understand in fact that it is because the label noise is very small.

- When the noise of label increases, as it is the case here (0.4), figure 11, the stochastic transition matrix begins to model very early the conditional



(a) TLasDLV Noise rate: 0.0



(b) TLasDLV Noise rate: 0.4

Fig. 7: Learning with TLasDLV

probabilities, becoming less sparse at the beginning of learning 10b, it tries somehow to capture the noisy classes;

- Then finally, when the percentages of the noise of label are rather important (0.6) figure 11, the stochastic transition matrix captures the noisy classes very early (figure 11b and figure 11c) and converges probably towards the true matrix in spite of its great sparsicity.

6.3. Benchmarking

One of the problems with the Noise Label in Deep Learning in terms of research topic is that, not all previous works have a common standard for comparing results. However, we compared several works that did their tests on the same architecture as ours, with the same parameters. The following table 2 summarizes the results obtained with comparable methods.

As we can see from these results, our proposed solution is not the best when label noises are very low; when they become high, our method outperforms previous works.

Methods	No Noise	Noise rate: 0.2	Noise rate: 0.4	Noise rate: 0.6
cross-entropy	0.979	0.969	NA	0.53 ± 0.006
bootstrap soft	0.979	0.969	NA	0.53 ± 0.004
bootstrap hard	0.979	0.968	NA	0.55 ± 0.013
backward	0.979	0.969	NA	0.674 ± 0.015
forward	0.979	0.969	NA	0.649 ± 0.044
TLasDLV(ours)	0.971	0.969	0.9579 ± 0.003	0.6911 ± 0.014

Table 2: Results and Comparison

7. Conclusion

Thus, the robust neural networks consist of latent variables and Bayesian networks. The former is employed as a true labels modeler while the latter deals with label noise as latent variables and is responsible for classification.

Experimental results demonstrated the model with state-of- the-art performance on popular public datasets such as MNIST and CIFAR10. It achieves a validation accuracy of 69.111.4% on the public datasets, which outperforms the state-of-the-art performance. Even when label noises are very low, the average accuracy reaches 97.1%, which is comparable to previous works.

References

[1] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (7553) (2015) 436.

[2] C. Zhang, S. Bengio, M. Hardt, B. Recht, O. Vinyals, Understanding deep learning requires rethinking generalization, arXiv preprint arXiv:1611.03530.

[3] D. Karimi, H. Dou, S. K. Warfield, A. Gholipour, Deep learning with noisy labels: Exploring techniques and remedies in medical image analysis, Medical Image Analysis 65 (2020) 101759.

[4] B. Li, T. Liu, B. Wang, L. Wang, Label noise robust curriculum for deep paraphrase identification, in: 2020 International Joint Conference on Neural Networks (IJCNN), IEEE, 2020, pp. 1–8.

[5] X. Wang, J. Xie, L. Yu, X. Tao, Ml-lrc: Low-rank-constraint-based multi-label learning with label noise, in: 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Vol. 1, IEEE, 2020, pp. 129–136.

[6] Z. Zhang, Y. Li, Y. Qin, et al., Making deep neural networks robust to label noise: A reweighting loss and data filtration, in: 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP), IEEE, 2019, pp. 289–293.

[7] M. D. Hoffman, D. M. Blei, C. Wang, J. Paisley, Stochastic variational inference, The Journal of Machine Learning Research 14 (1) (2013) 1303–1347.

[8] E. Jang, S. Gu, B. Poole, Categorical reparameterization with gumbel-softmax, arXiv preprint arXiv:1611.01144.

[9] B. Frénay, M. Verleysen, Classification in the presence of label noise: a survey, IEEE transactions on neural networks and learning systems 25 (5) (2014) 845–869.

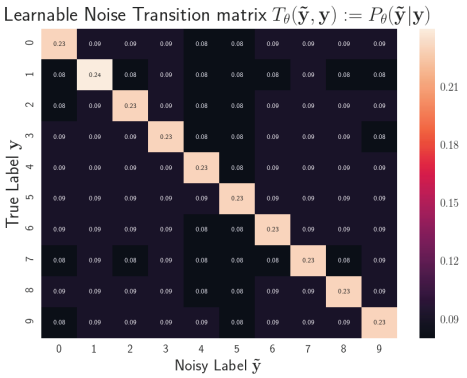
[10] D. Angluin, P. Laird, Learning from noisy examples, Machine Learning 2 (4) (1988) 343–370.

[11] C. E. Brodley, M. A. Friedl, Identifying mislabeled training data, Journal of artificial intelligence research 11 (1999) 131–167.

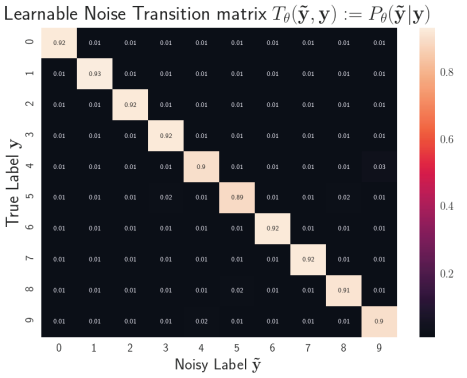
[12] R. J. Hickey, Noise modelling and evaluating learning from examples, Artificial Intelligence 82 (1-2) (1996) 157–179.

[13] M. F. McNitt-Gray, S. G. Armato, C. R. Meyer, A. P. Reeves, G. McLennan, R. C. Pais, J. Freymann, M. S. Brown, R. M. Engelmann, P. H. Bland, et al., The lung image database consortium (lidc) data collection process for nodule detection and annotation, Academic radiology 14 (12) (2007) 1464–1474.

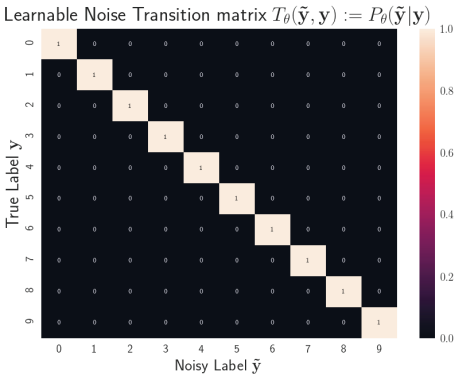
[14] M. B. Propp, D. L. Propp, Method and apparatus for data encoding and communication over noisy media, uS Patent 5,944,842 (Aug. 31 1999).
[15] C. Bishop, C. M. Bishop, et al., Neural networks for pattern recognition, Oxford university press, 1995.
[16] H.-I. Suk, An introduction to neural networks and deep learning, in: Deep Learning for Medical Image Analysis, Elsevier, 2017, pp. 3–24.
[17] B. Widrow, M. A. Lehr, 30 years of adaptive neural networks: perceptron, madaline, and backpropagation, Proceedings of the IEEE 78 (9) (1990) 1415–1442.
[18] Y. Grandvalet, Y. Bengio, Semi-supervised learning by entropy minimization, in: Advances in neural information processing systems, 2005, pp. 529–536.
[19] T. Xiao, T. Xia, Y. Yang, C. Huang, X. Wang, Learning from massive noisy labeled data for image classification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2691–2699.
[20] S. Sukhbaatar, J. Bruna, M. Paluri, L. Bourdev, R. Fergus, Training convolutional networks with noisy labels, arXiv preprint arXiv:1406.2080.
[21] G. Patrini, A. Rozza, A. K. Menon, R. Nock, L. Qu, Making deep neural networks robust to label noise: a loss correction approach, stat 1050 (2017) 22.
[22] A. Vahdat, Toward robustness against label noise in training deep discriminative neural networks, in: Advances in Neural Information Processing Systems, 2017, pp. 5601–5610.
[23] V. Mnih, G. E. Hinton, Learning to label aerial images from noisy data, in: Proceedings of the 29th International conference on machine learning (ICML-12), 2012, pp. 567–574.
[24] S. E. Reed, H. Lee, D. Anguelov, C. Szegedy, D. Erhan, A. Rabinovich, Training deep neural networks on noisy labels with bootstrapping.
[25] D. Tanaka, D. Ikami, T. Yamasaki, K. Aizawa, Joint optimization framework for learning with noisy labels, arXiv preprint arXiv:1803.11364.
[26] K.-H. Lee, X. He, L. Zhang, L. Yang, Cleannet: Transfer learning for scalable image classifier training with label noise, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 5447–5456.
[27] V. Mnih, G. E. Hinton, Learning to detect roads in high-resolution aerial images, in: European Conference on Computer Vision, Springer, 2010, pp. 210–223.
[28] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, A. Y. Ng, Reading digits in natural images with unsupervised feature learning, in: NIPS workshop on deep learning and unsupervised feature learning, Vol. 2011, 2011, p. 5.
[29] A. Krizhevsky, G. Hinton, Learning multiple layers of features from tiny images.
[30] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, ImageNet: A Large-Scale Hierarchical Image Database, in: CVPR09, 2009.
[31] A. Coates, A. Y. Ng, The importance of encoding versus training with sparse coding and vector quantization, in: Proceedings of the 28th International Conference on Machine Learning (ICML-11), 2011, pp. 921–928.
[32] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE 86 (11) (1998) 2278–2324.
[33] A. L. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, C. Potts, Learning word vectors for sentiment analysis, in: Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies-volume 1, Association for Computational Linguistics, 2011, pp. 142–150.
[34] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C. L. Zitnick, Microsoft coco: Common objects in context, in: European conference on computer vision, Springer, 2014, pp. 740–755.
[35] K.-H. Lee, X. He, L. Zhang, L. Yang, Cleannet: Transfer learning for scalable image classifier training with label noise, arXiv preprint arXiv:1711.07131.
[36] W. Li, L. Wang, W. Li, E. Agustsson, L. Van Gool, Webvision database: Visual learning and understanding from web data, arXiv preprint arXiv:1708.02862.
[37] L. Bossard, M. Guillaumin, L. Van Gool, Food-101–mining discriminative components with random forests, in: European Conference on Computer Vision, Springer, 2014, pp. 446–461.
[38] C. W. Fox, S. J. Roberts, A tutorial on variational bayesian inference, Artificial intelligence review 38 (2) (2012) 85–95.
[39] D. Blei, R. Ranganath, S. Mohamed, Variational inference: Foundations and modern methods.
[40] D. M. Blei, A. Kucukelbir, J. D. McAuliffe, Variational inference: A review for statisticians, Journal of the American Statistical Association 112 (518) (2017) 859–877.
[41] D. P. Kingma, M. Welling, Auto-encoding variational bayes.
[42] C. J. Maddison, A. Mnih, Y. W. Teh, The concrete distribution: Acontinuous relaxation of discrete random variables, arXiv preprint arXiv:1611.00712.



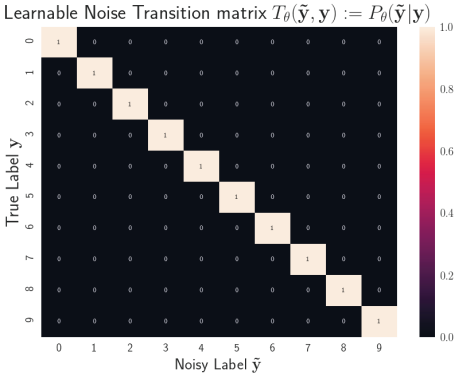
(a) Learnable Transition Matrix after 1 iteration



(b) Learnable Transition Matrix after 50 iterations

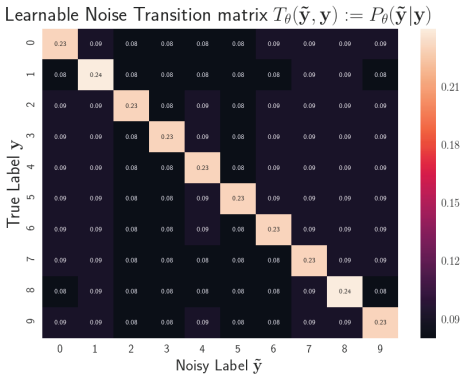


(c) Learnable Transition Matrix after 100 iterations

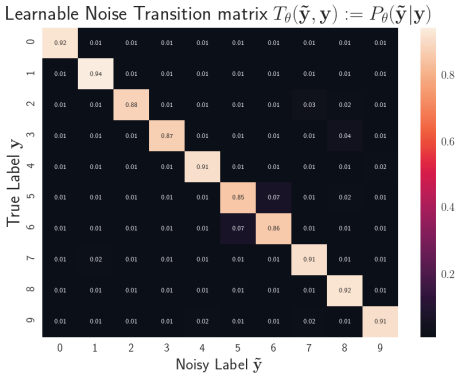


(d) Learnable Transition Matrix after 400 iterations

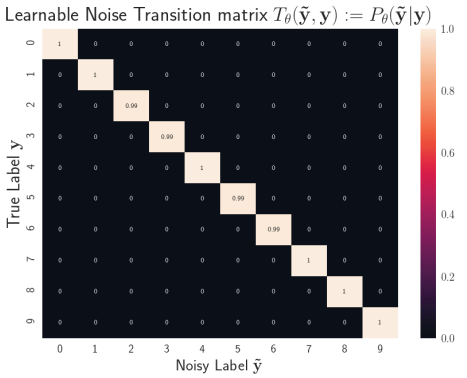
Fig. 8: Learnable Transition Matrix during training with TLasDLV, noise rate: 0.0



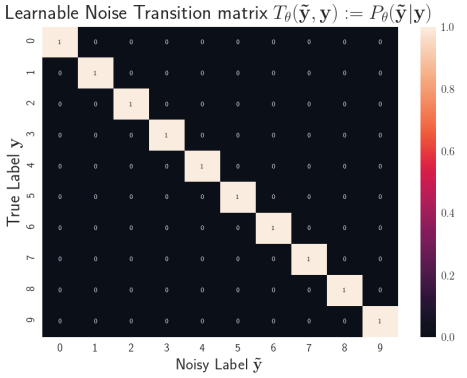
(a) Learnable Transition Matrix after 1 iteration



(b) Learnable Transition Matrix after 50 iterations

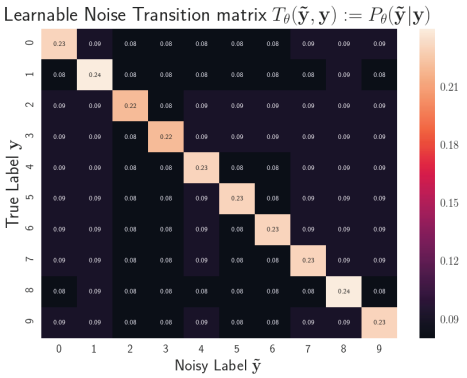


(c) Learnable Transition Matrix after 100 iterations

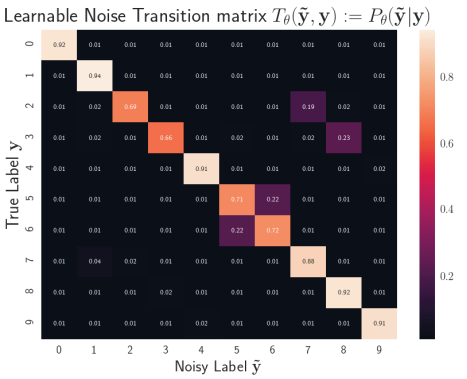


(d) Learnable Transition Matrix after 400 iterations

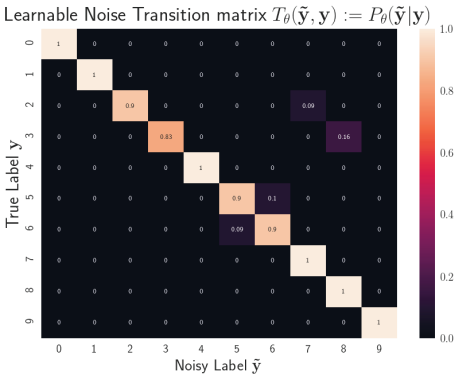
Fig. 9: Learnable Transition Matrix during training with TLasDLV, noise rate: 0.2



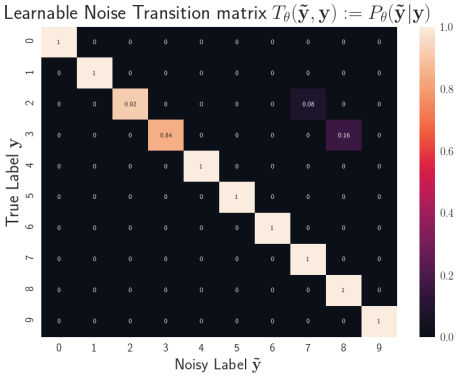
(a) Learnable Transition Matrix after 1 iteration



(b) Learnable Transition Matrix after 50 iterations

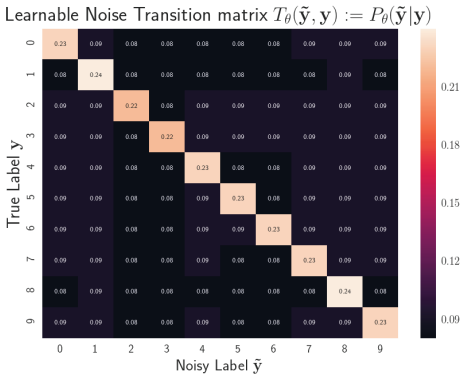


(c) Learnable Transition Matrix after 100 iterations

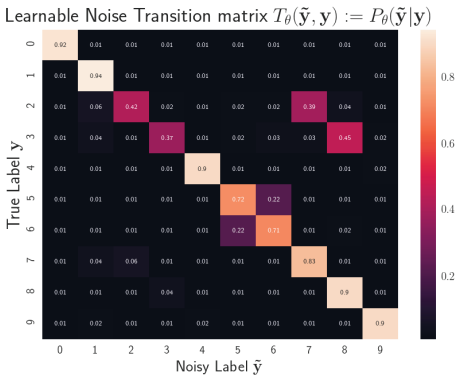


(d) Learnable Transition Matrix after 400 iterations

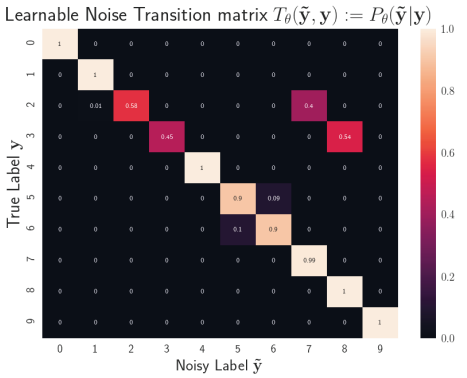
Fig. 10: Learnable Transition Matrix during training with TLasDLV, noise rate: 0.4



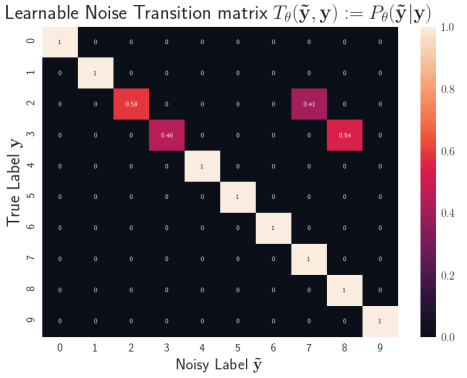
(a) Learnable Transition Matrix after 1 iteration



(b) Learnable Transition Matrix after 50 iterations



(c) Learnable Transition Matrix after 100 iterations



(d) Learnable Transition Matrix after 400 iterations

Fig. 11: Learnable Transition Matrix during training with TLasDLV, noise rate: 0.6

Analysis and optimization of Fault-tolerant Behaviour of Motors in Electric Vehicular Systems

Ms. Shivani Jitendra Khare[1], Mr. Shubham Singh[1], Mr. Siddharth Roy[1], Prof. Yogesh B. Mandake[1], Dr. Deepak S. Bankar[1] (Hod, Department Of Electrical Engineering)

[1]Bharati Vidyapeeth (Deemed to be) University, College of Engineering, Pune

* Correspondence: shivani.khare-coep@bvucoep.edu.in

Abstract

The advent of electrification in vehicular drives has proven to be challenging over time. An increase in electronics in drive systems has demonstrated escalation in performance efficiency but has, in turn, escorted extensive concerns regarding internal electric faults and safety. Failure of the drive caused by the abnormal conditions will ultimately cause the failure of the entire electric vehicle (EV) system. Thus, the fault tolerance capability of a motor drive is of utmost importance for enhancing motor speed limit range and safety. Several methods of fault detection have already been worked upon. The scope of this research includes analyses of torque ripple values, State of Charge (SOC) decline rate and minimum-maximum torque acquired by Switch Reluctance Motor (SRM) and Permanent Magnet Synchronous Motor (PMSM) drives under varied fault conditions. Furthermore, a solution to increase the fault tolerance capability of the PMSM motor drive will be discussed.

Keywords: PMSM, SRM, electric vehicle system, EV drives, fault-tolerance, state of charge, efficiency, torque

1. INTRODUCTION

Owing to the drawbacks of utilisation of fossil fuels combined with the greenhouse emissions, it has become imperative to develop techniques that would minimise the unfavourable outcomes associated with the use of these sources of energy generation. A dire need for advancement leading to reduced emissions aroused which in turn lead to extensive scientific research and development of methodologies that would contract emissions and conserve the environment. Evolution of battery systems was one such major step that led towards the beginning of electric vehicular systems [1]. With their ability to produce zero greenhouse gas emissions, low overall carbon footprint and higher efficiency, adoption of electric vehicular systems began at a fast pace [2], [3]. Ability of integration of electric vehicle to the power grid and the alternative of application of renewable energy systems for charging has compensated a lot of issues regarding charging of electric vehicle in the past.

The EV market grew to a net percentage of 28% of the net total of vehicles on road from year 1897 to 1900. However, certain drawbacks led to the constant decrease in momentum of electric vehicles (EVs) in the market. Simultaneously, the internal fuel combustion vehicles took over with an enhanced momentum. The major reason of these vehicles to gain popularity originated from the low oil prices during a certain period [4]. In 1996, with the evolution of concept of EV1 by General Motors, the electric vehicular systems came into the picture once again and began gaining popularity. As the evolution of the vehicular market escalated, company's such as Tata, Ford, Honda, Toyota, Tesla, Chevrolet, Nissan Leaf began manufacturing electric vehicles of their own. The first successful commercial hybrid EV, Toyota Prius, launched in 1997 in the country of Japan, gained immense popularity and was extensively sold with 18,000 units in the very first year of its launch in the market. The major reasons of this popularity were the benefits of zero exhaust gas emissions, enhanced efficiency, narrowed greenhouse emissions and low carbon footprint [5], [6]. However, a net of 7.2 million EVs were available in the market by the year 2019. This approximately translates to 1 % of the global car inventory sector, proving that EVs still have to travel a long way in the domain of development and implementation. In accordance with the latest 2021 EV statistics, the EV global share market grew to the percentage of 8.3% consisting of battery-based EVs of the global inventory sector. A lot of countries have commented on pursuing 100% electric vehicular system in their automotives by the year 2050.

2. OVERVIEW OF MOTORS IN ELECTRIC VEHICULAR SYSTEMS

The major concern here are the problems that are being faced in the electric vehicular systems, right from the manufacturing stage to the implementation on road [7]. We examine the EV market statistical data from the year 2010 to 2020. Observations show that the major motors that are employed in the electric vehicular systems are that of Permanent Magnet Synchronous Motors (PMSM), Switched Reluctance Motor (SRM) and Induction Motor (IM). Among these motors, PMSM has proven to be the efficient and is rated as the first choice among the manufacturers.

Table 1 – Statistics for motor usage in market

EV model	Power (kW)	Motor	Year
Mahindra e2o Plus	19 – 30	IM	2016
Renault Kangoo ZE	44	PMSM	2011
Mitsubishi I –MiEV	47	PM	2010
Volkswagen E-up	60	PMSM	2019
Renault Zoe	65	PMSM	2012
LandRover	70	SRM	2013
Renault Fluence Z.E.	70	PMSM	2012
Nissan Leaf	80	PMSM	2010
BJEV EC5	80	PMSM	2019
Hyundai Ioniq Electric	88	PMSM	2016
Hyundai Kona	88-150	PMSM	2018
BYD E6	90	PMSM	2014
BMW i3	125	PMSM	2013
Xpeng G3	139	PMSM	2018
Mercedes-Benz EQC	150*2	IM	2019
BJEV EU5	160	PMSM	2018
Tesla Model X	193-375	IM	2015
Tesla Model 3	211-340	PMSM	2020
Tesla Model S	235-568	IM	2012
NIO EC6	320	PMSM	2020
NIO ES6	320	PMSM	2020

PMSM are high energy density motors. Permanent magnet used in PMSM is of non-renewable nature and is of Neodymium Fe Boron (NdFeB) and Samarium Cobalt (SmCo) magnets. These motors are available in two variations – Internal PMSM and Surface mount PMSM. With advantages, comes drawbacks [8], [9]. The high cost of permanent magnets in PMSM and the low fault tolerance capability of the motor are two of the major concerns of the motor.

For achieving the same output power, the cost of SRM is half as compared to PMSM and is 80% less than that of IM [10]. Although SRM is cost beneficial and has enhanced speed range which is a result of improved high temperature resistance characteristics, the poor capacity of torque constraints the energy recovery effect of the motor [11], [12]. Also, presence of large amounts of harmonics and noise in the torque characteristics of the motor is another drawback. The low efficiency of IM and the moderate cost and torque density characteristics of the motor has led to decreased application of IM in the current electric vehicular system.

Hence, even though the cost of PMSM is high, PMSM is still preferred [13]. This is because, the regenerative braking is more effective in the constant power region near the base speed during the motor operation [14]. This enables the motor to recover the energy of braking more efficiently as compared to SRM and IM [15]. The economics spent on improving the characteristics of torque and eliminating harmonic content and noise content in SRM motor is subsided by the cost of PMSM and its characteristics. Hence, PMSM is highly preferred for electric vehicular operation.

3. EV MODELLING

Fig. 1 below shows the EV modelling flow diagram.

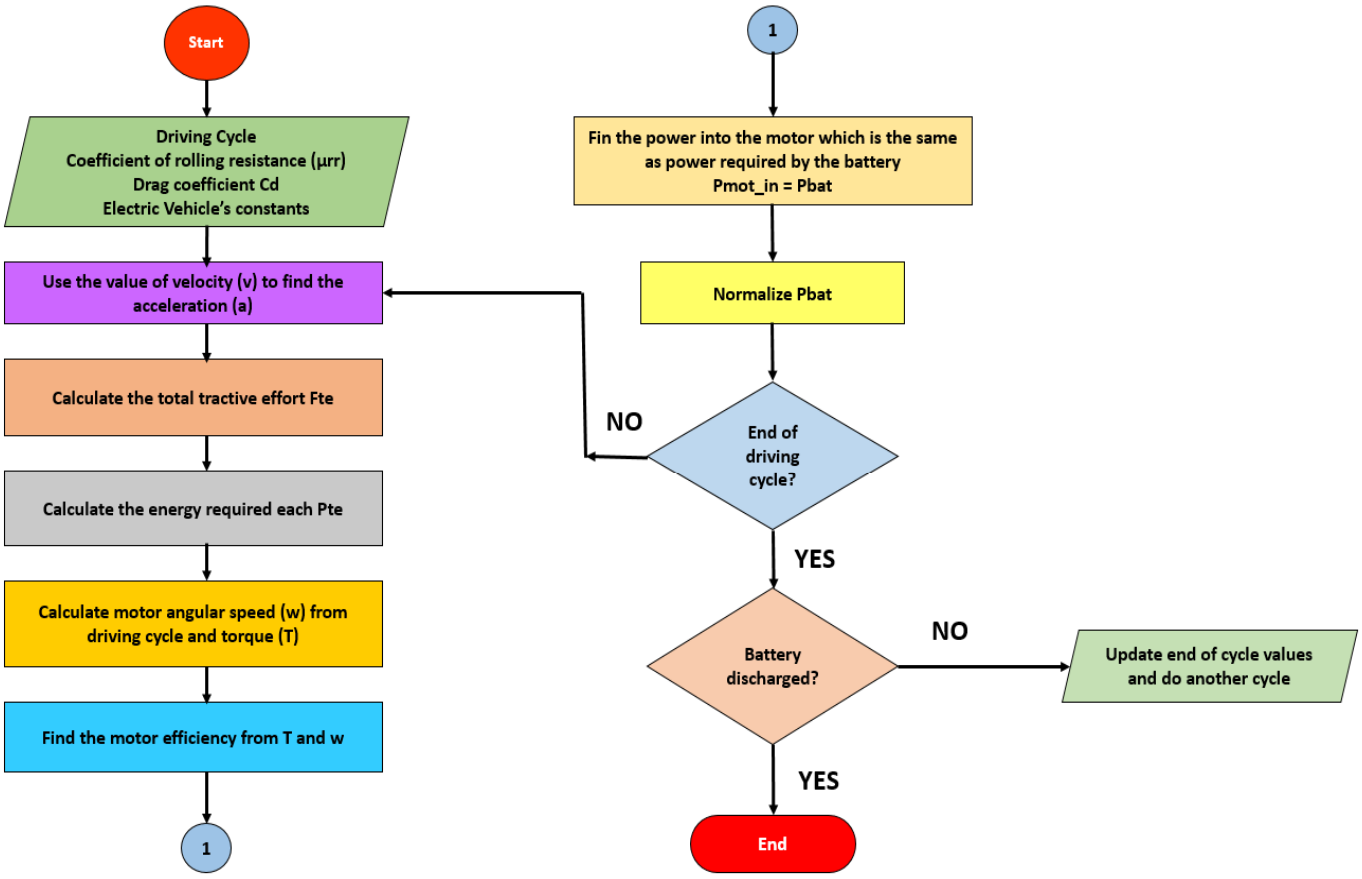


Fig 1: EV modelling flow diagram

4. EXPLORATORY DATA ANALYSIS OF PMSM CHARACTERISTICS USING REAL – TIME DATA

4.1. About Dataset

The dataset considered for the data analysis has comprises of several data collected by sensors placed on PMSM. Test bench measurements were gathered by the LEA department at Paderborn University. The data recordings were sampled at 2 Hz frequency. This data set is mildly anonymized.

4.1.1. *Key points about the dataset*

- The data set measures.csv contains data of 41 unique profile id's and 13 columns. Each row represents the collected data for the particular profile_id. Following columns are present in the dataset –

```
u_q   coolant  stator_winding      u_d  stator_tooth  motor_speed      i_d      i_q       pm  stator_yoke   ambient   torque  profile_id
```

- Each column represents observations for the following parameters shown in the Table 2.

Table 2 – Dataset Parameters

Serial Number	Parameter	Parameter definition
1	u_q	Voltage q-component measurement in dq-coordinates (in V).
2	coolant	Coolant temperature ($°C$).
3	stator_windings	Stator winding temperature ($°C$) measured with thermocouples.
4	u_d	Voltage d-component measurement in dq-coordinates.

5	Stator_tooth	Stator tooth temperature ($°C$) measured with thermocouples.
6	motor_speed	Motor speed (RPM).
7	i_d	Current d-component measurement in dq-coordinates.
8	pm	Permanent magnet temperature ($°C$) measured with thermocouples and transmitted wirelessly via a thermography unit.
9	stator_yoke	Stator yoke temperature ($°C$) measured with thermocouples.
10	ambient	Ambient Temperature ($°C$)
11	profile_id	Distinguishing feature of multiple measurement sessions.

- Following code snippet in Fig. 2 below shows detailed information about data type of the parameters gathered from PMSM sensors.
- No missing values were found in the dataset for both the generation data and the weather data. Hence, there is no requirement of dataset cleaning.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1330816 entries, 0 to 1330815
Data columns (total 13 columns):
u_q               1330816 non-null float64
coolant           1330816 non-null float64
stator_winding    1330816 non-null float64
u_d               1330816 non-null float64
stator_tooth      1330816 non-null float64
motor_speed       1330816 non-null float64
i_d               1330816 non-null float64
i_q               1330816 non-null float64
pm                1330816 non-null float64
stator_yoke       1330816 non-null float64
ambient           1330816 non-null float64
torque            1330816 non-null float64
profile_id        1330816 non-null int64
dtypes: float64(12), int64(1)
memory usage: 132.0 MB
```

Fig 2: Data type description

4.2. Determining Correlations

Determining the data distribution and correlation among the dataset parameters, we have considered constructing a heat map for generating suitable conclusions about the degree of correlation present in the dataset parameters. ***Seaborn library*** has been made use of for generating the heatmaps.

4.2.1. *Generating heatmap*

Fig. 3 below shows the heat map for data correlation.

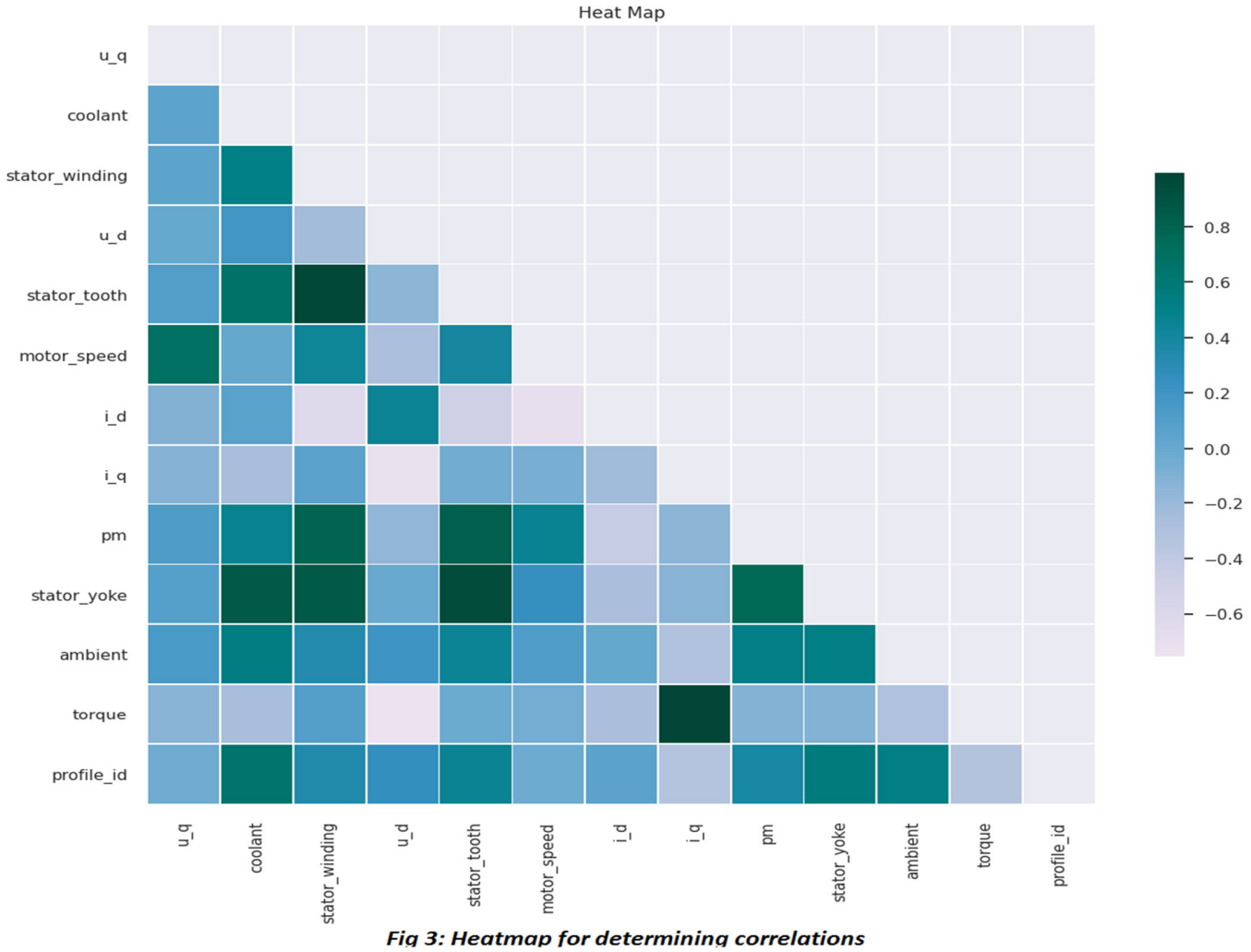


Fig 3: Heatmap for determining correlations

- ***Observations***

From the heatmap we observe the following –

- High correlation is observed between **i_q** and **torque.**
- High correlation is observed between **stator_tooth** and **stator_winding.**
- High correlation is observed between **stator_yoke** and **stator_tooth.**
- Negative Linear correlation is observed between **u_d** and **i_q.**
- Negative Linear correlation is observed between **u_d** and **torque.**
- Negative Linear correlation is observed between **i_d** and **stator_winding.**

Through the above analysis we have observed that high correlation between i_q and torque exist. Therefore, in this research paper, PMSM and SRM motors are being analysed for fault tolerance behaviour on the basis of electromagnetic torque characteristics. The EV model consisting of both the motors are being subjected to similar power ratings and fault conditions, thereby generating results and conducting an effective comparative study. Thereafter, an effective method for improvement in fault tolerance of PMSM has been simulated in the paper.

5. MODELLING OF THE DYNAMICS OF ELECTRIC VEHICLE

The dynamics of an EV is modelled to determine the total tractive effort required to maintain the motion of the vehicle in the forward direction. Fig. 4 below shows the vehicle dynamics. The total tractive force required is given by the sum of the following forces:

1. **Force Rolling Resistance**, $F_{rr} = \mu_{rr}.m.g$

where μ_{rr} is coefficient of rolling resistance, m is the mass of vehicle and g is the acceleration due to gravity constant.

2. **Force Hill Climb**, $F_{hc} = m.g.\sin(\Psi)$
   where Ψ is the slope angle.
3. **Force Aerodynamic**, $F_{ad} = \frac{1}{2}.\rho.A.C_d.V^2$
   where ρ is the air density, A is the frontal area, C_d is the drag coefficient and V is the velocity of the vehicle.
4. **Force Linear Acceleration**, $F_{la} = m.a$
   where a is the acceleration of the vehicle.

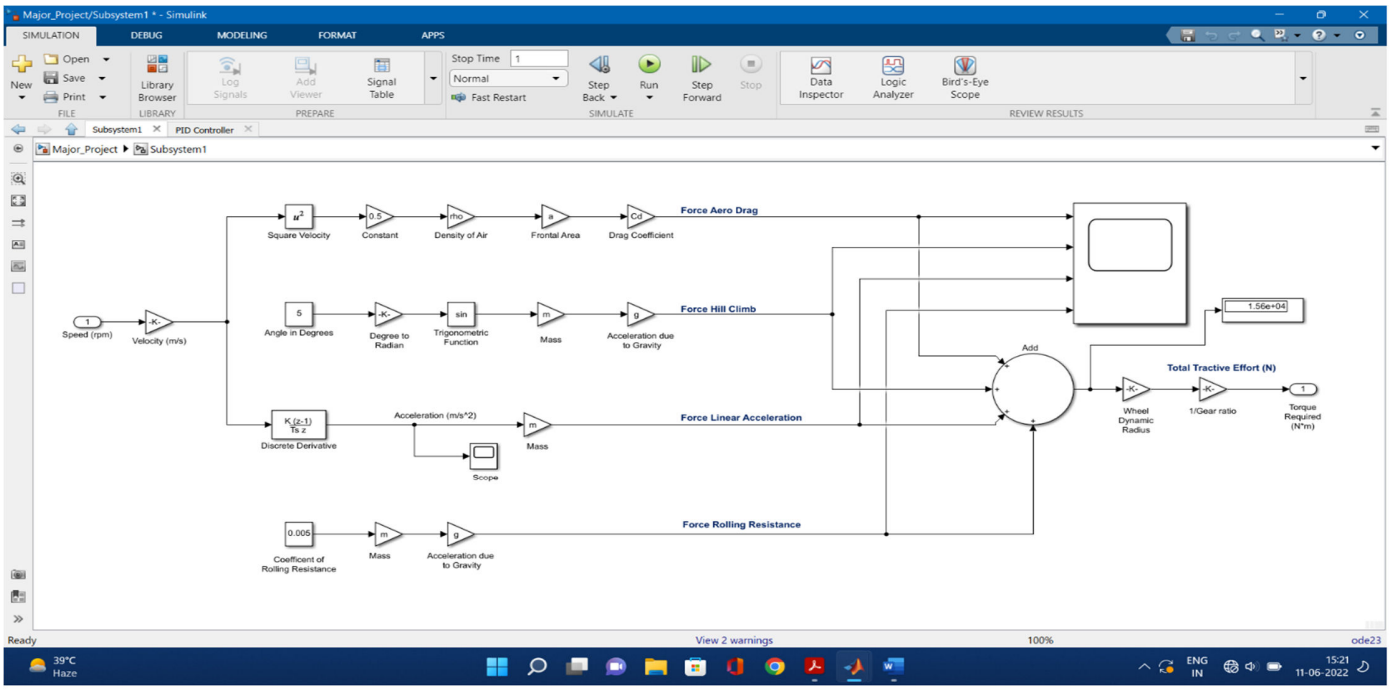


Fig 4: Simulink model for vehicle dynamics

The parameters used in the Simulink model of the electric vehicle dynamics are listed in the Table 3 below.

TABLE 3 – Electric Vehicle Dynamics

S. No.	Modelling Parameters	Values
1.	Vehicle Mass [m] (kg)	1000
2.	Frontal area [A] (m^2)	1.9992
3.	Coefficient of Rolling Resistance [μ_{rr}]	0.005
4.	Drag Coefficient [C_d]	0.6
5.	Acceleration Due to Gravity [g] (m/s^2)	9.81
6.	Density of Air [ρ] (kg/m^3)	1.25
7.	Inclination Angle [Ψ] (deg.)	5
8.	Wheel radius (m)	0.316

6. FAULT ANALYSIS

Simulation and fault analysis of an electric vehicle is necessary, as it helps to determine the vehicle performance under the abnormal conditions and ensures the reliability of the electric drive system.

For the purpose of analysis in this paper, the SRM and PMSM driven electric vehicle model were first simulated in the normal condition and then with the various abnormal conditions.

Various types of faults considered in this paper are:

1. Single-phase short circuit fault
2. Phase-to-phase short circuit fault
3. Three-phase ground fault

6.1. Normal Operation

- SRM driven electric vehicle model was simulated in the normal condition and the outputs were obtained in terms of electromagnetic torque, speed variation, motor current and flux as shown below in Fig. 5, Fig. 6, and Fig. 7, respectively.

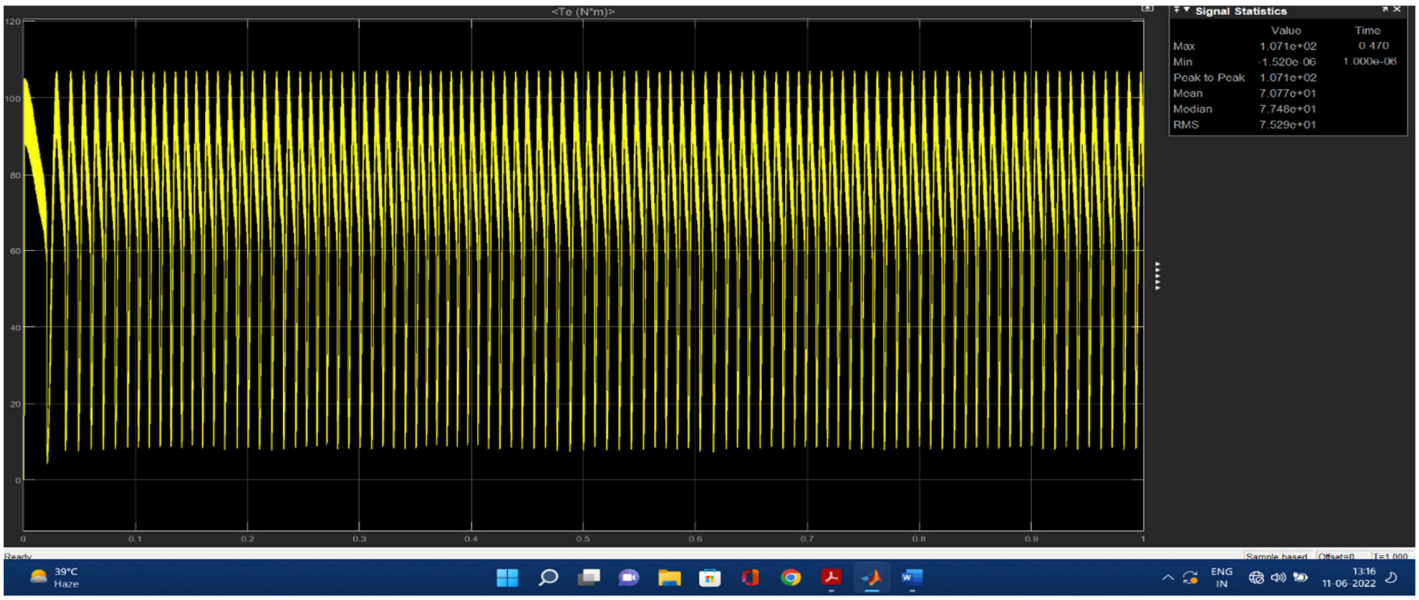


Fig 5: SRM Torque for normal operation

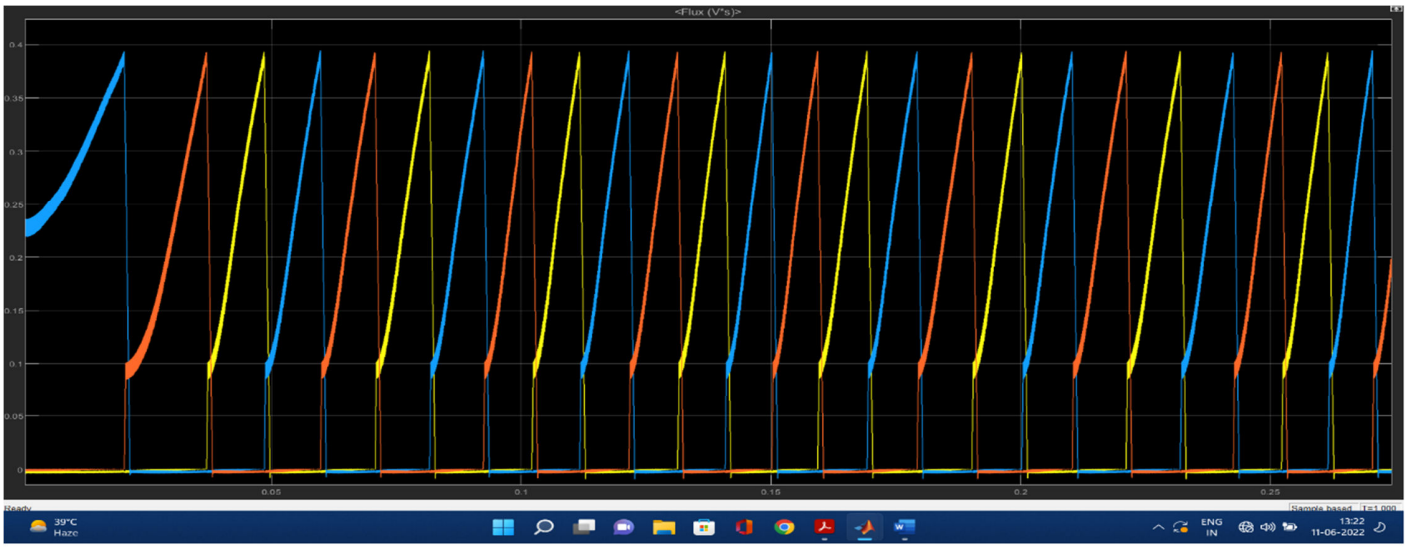


Fig 6: SRM Flux for normal operation

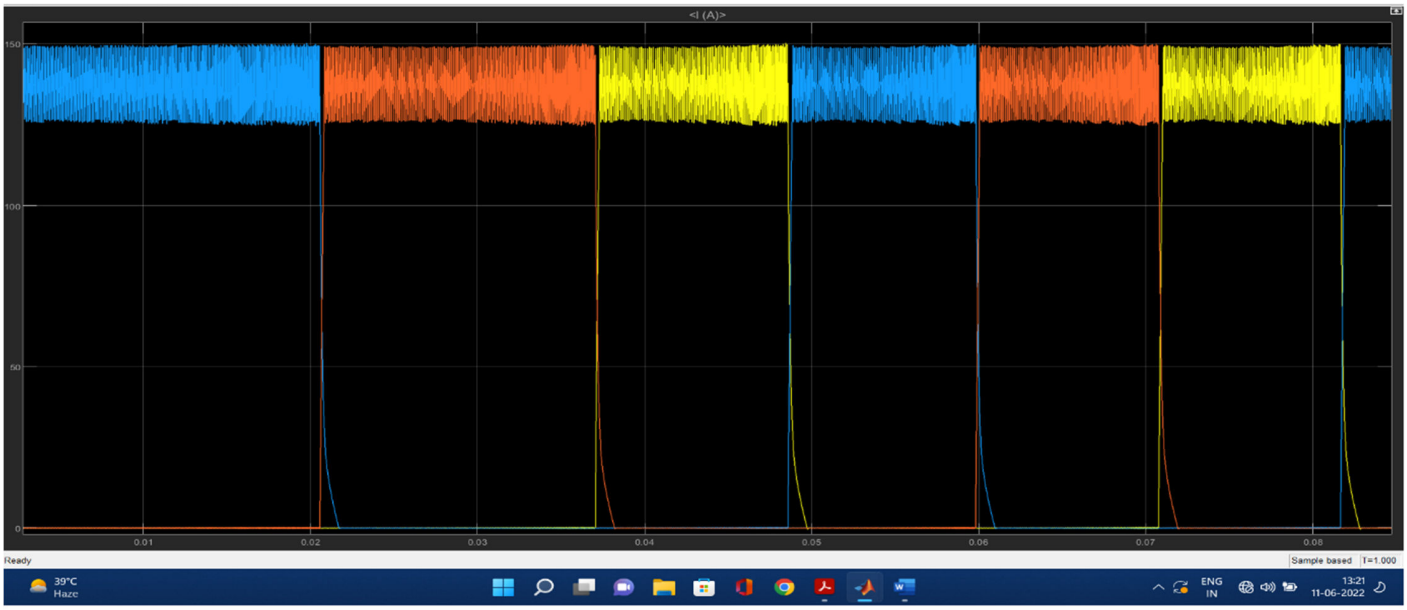


Fig 7: SRM Current for normal operation

- PMSM driven electric vehicle model was simulated in the normal condition and the outputs were obtained in terms of electromagnetic torque, speed variation, motor current and flux as shown below in Fig. 8 and Fig. 9.

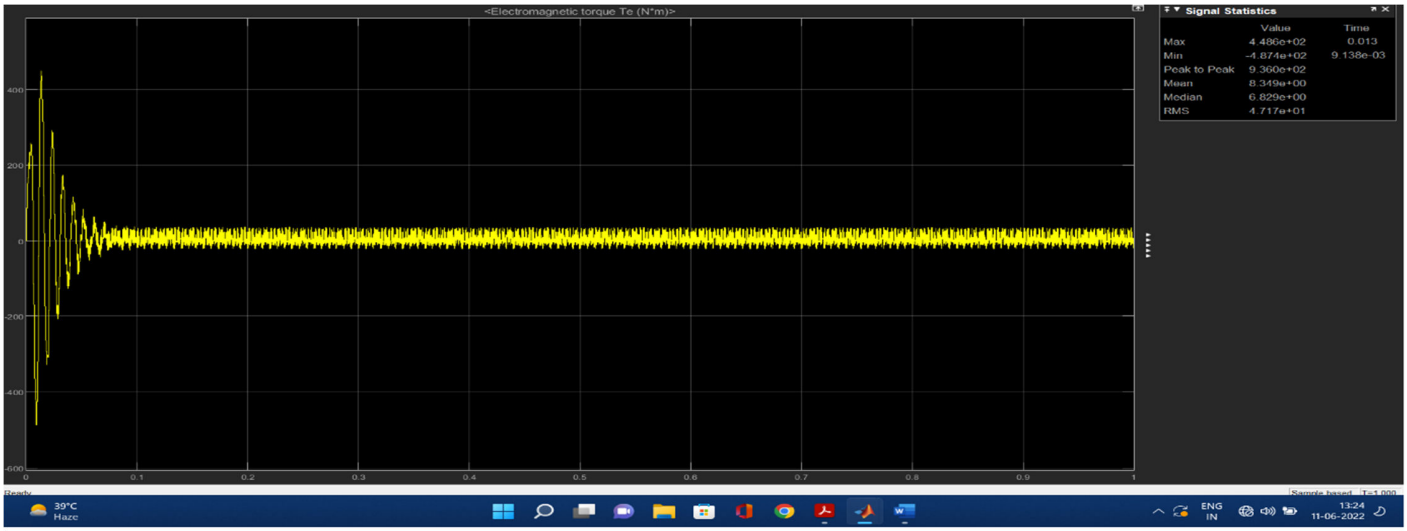


Fig 8: PMSM Torque for normal operation

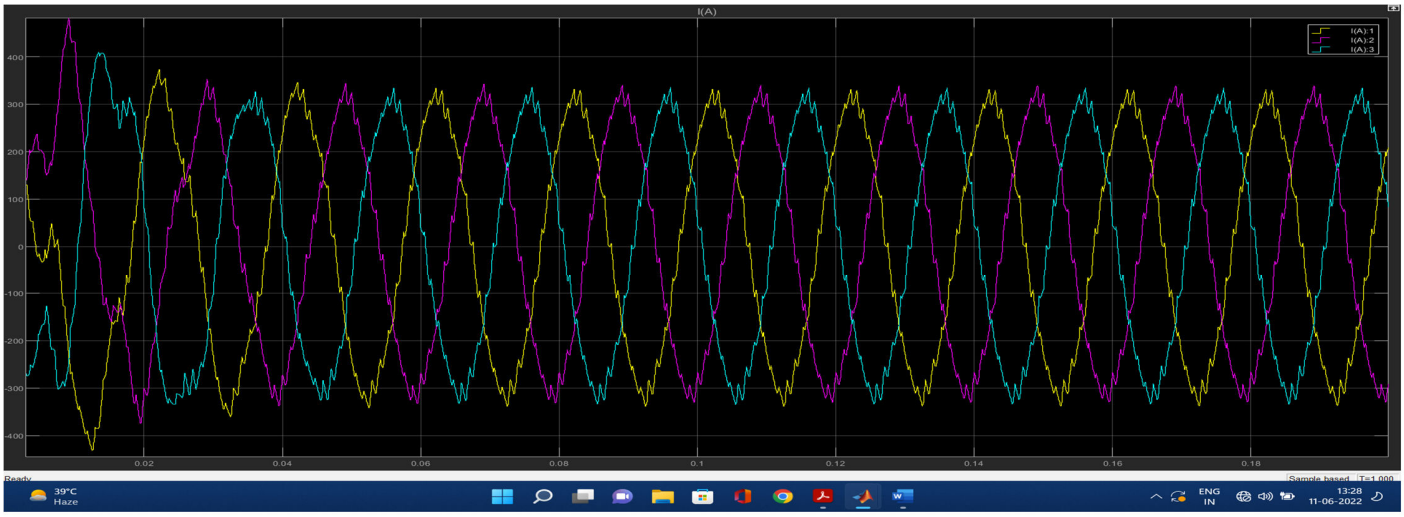


Fig 9: PMSM Flux for normal operation

6.2. Abnormal Operation

6.2.1. *Single-phase short circuit fault*

The single-phase short circuit fault was created using the 3-ph fault block (in the Simulink library) and was assumed to occur at 0.25 seconds. The fault was cleared at 0.67 seconds by entering value for the switching time in the block. The results of simulation for SRM and PMSM (3-ph) motor are shown in Table 4. Fig. 10 and Fig. 11 below show the fault characteristics of SRM and PMSM.

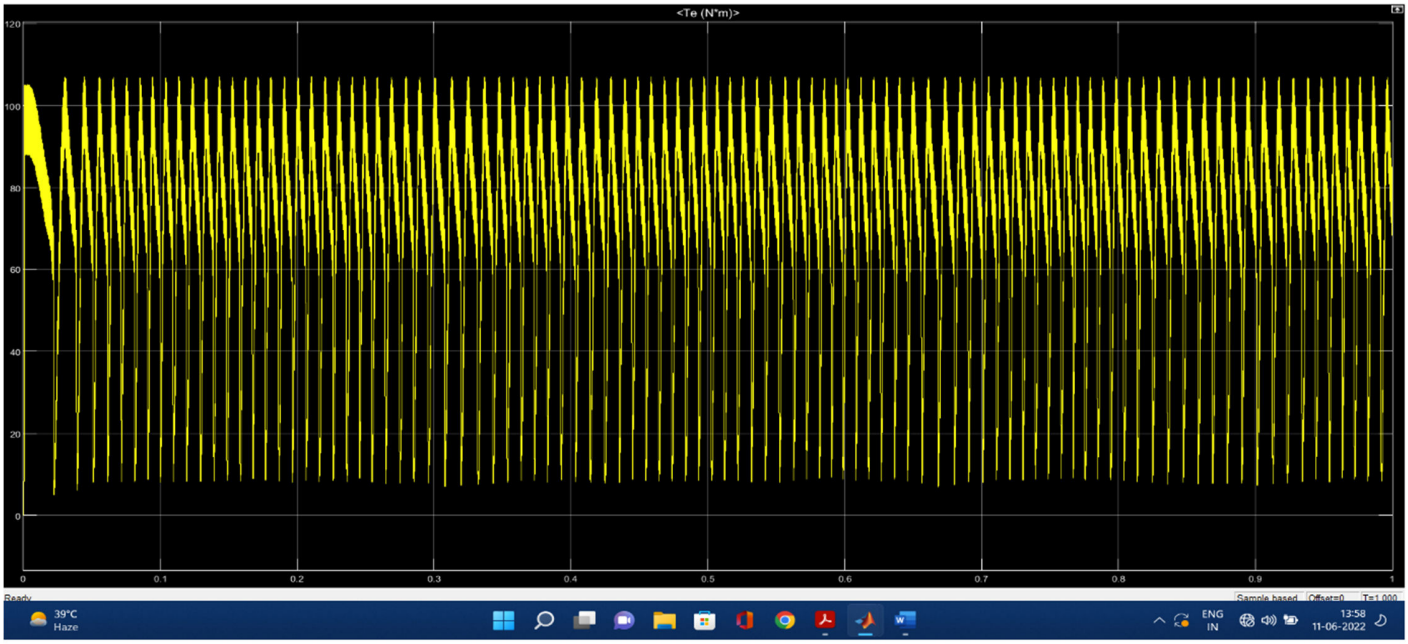


Fig 10: SRM Torque for single phase fault

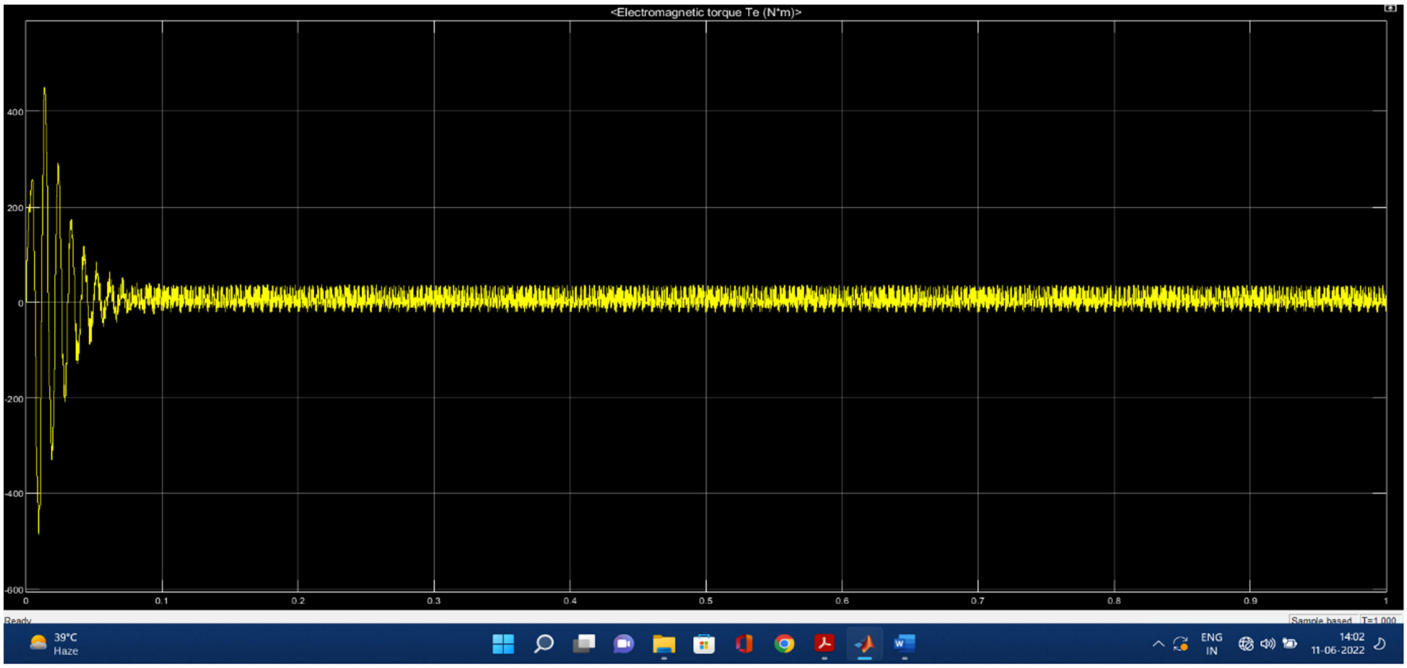


Fig 11: PMSM Torque for single phase fault

6.2.2. *Phase-to-phase short circuit fault*

The phase-to-phase short circuit fault was created using the 3-ph fault block (in the Simulink library) and was assumed to occur at 0.25 seconds. The fault was cleared at 0.67 seconds by entering value for the switching time in the block. The results of simulation for SRM and PMSM (3-ph) motor are shown in Table 4. Fig. 12 and Fig. 13 below show the fault characteristics of SRM and PMSM.

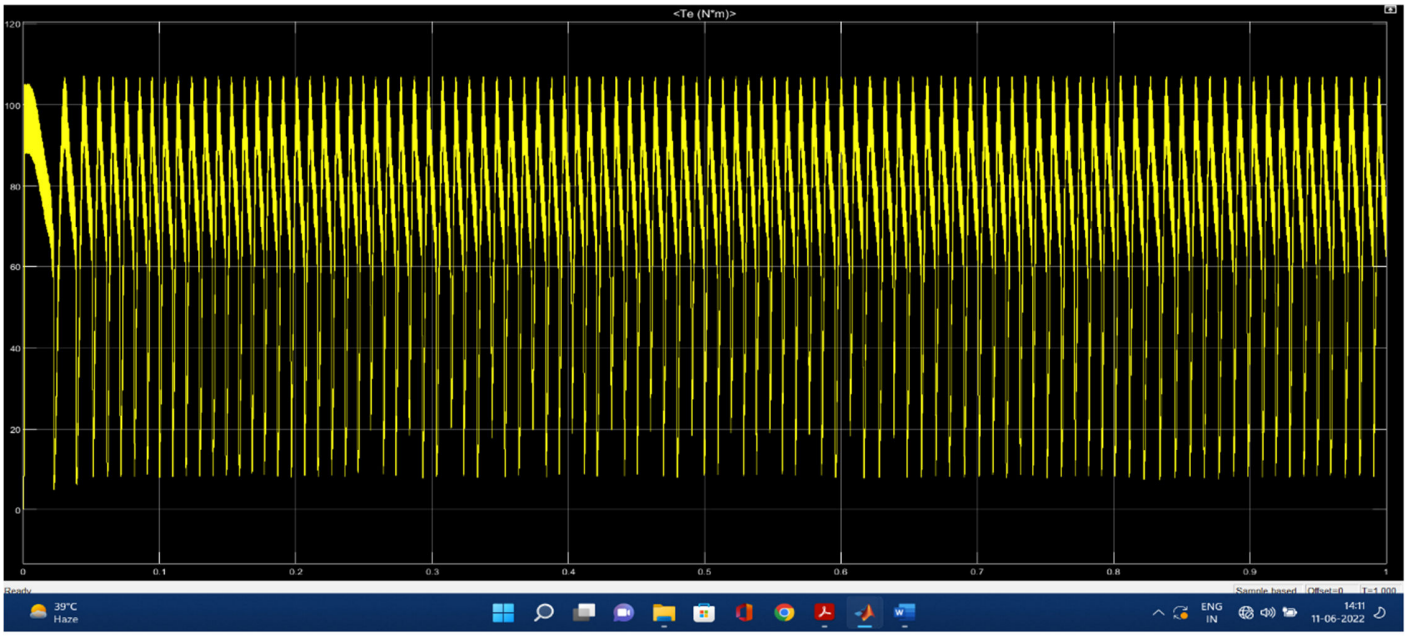


Fig 12: SRM Torque for two phase

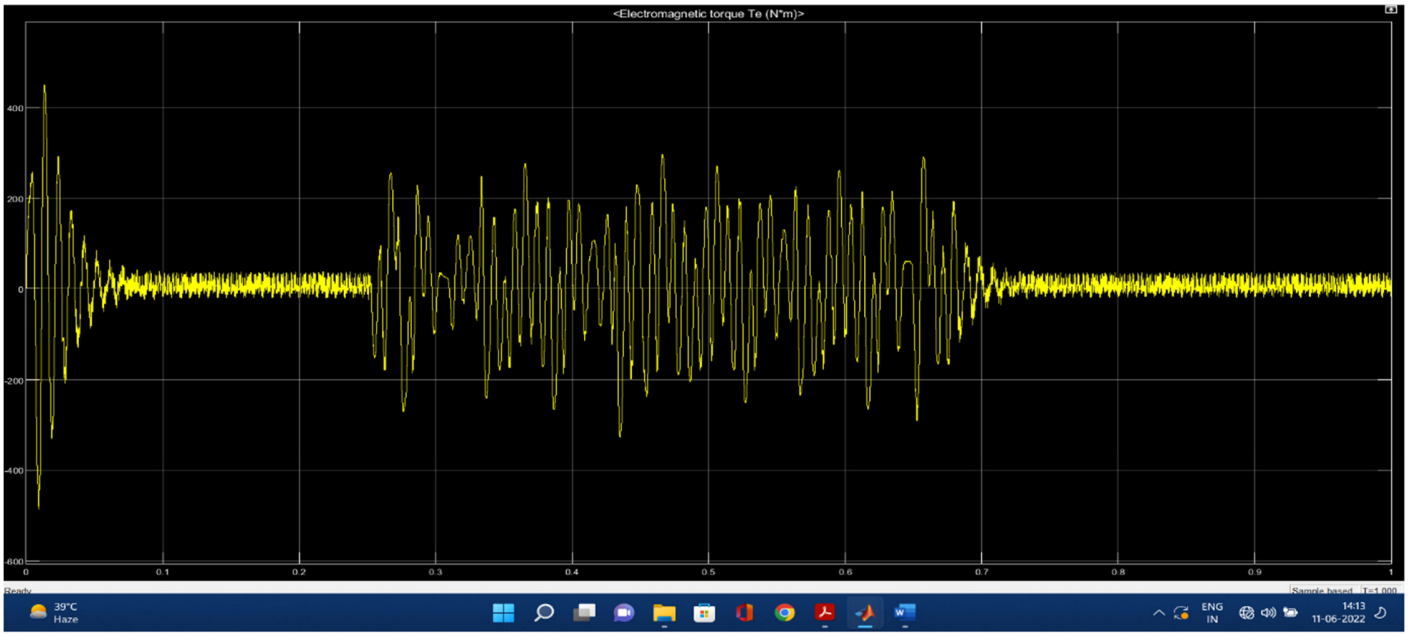


Fig 13: PMSM Torque for two phase fault

6.2.3. *Three-phase ground fault*

The three-phase short circuit fault was created using the 3-ph fault block (in the Simulink library) and was assumed to occur at 0.25 seconds. The fault was cleared at 0.67 seconds by entering value for the switching time in the block. The results of simulation for SRM and PMSM (3-ph) motor are shown in Table 4. Fig. 14 and Fig. 15 below show the fault characteristics of SRM and PMSM.

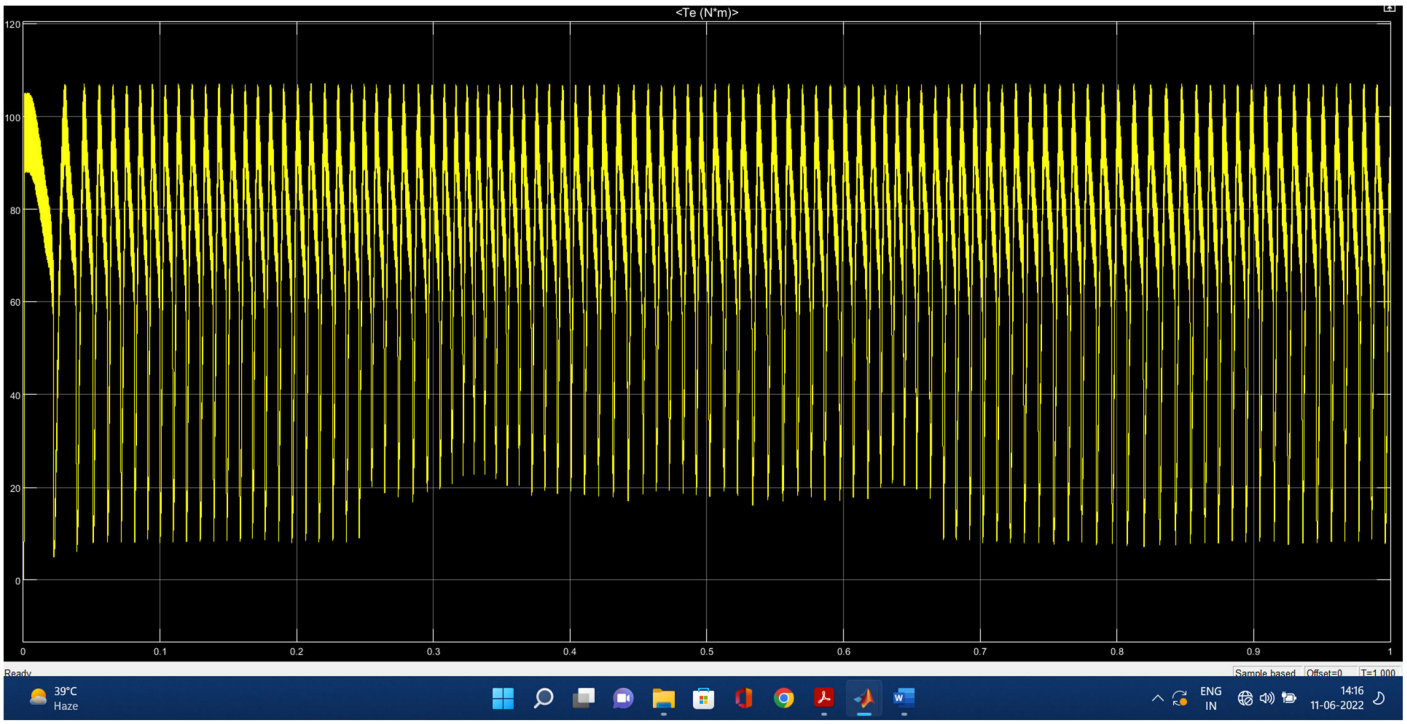


Fig 14: SRM Torque for three phase short circuit condition

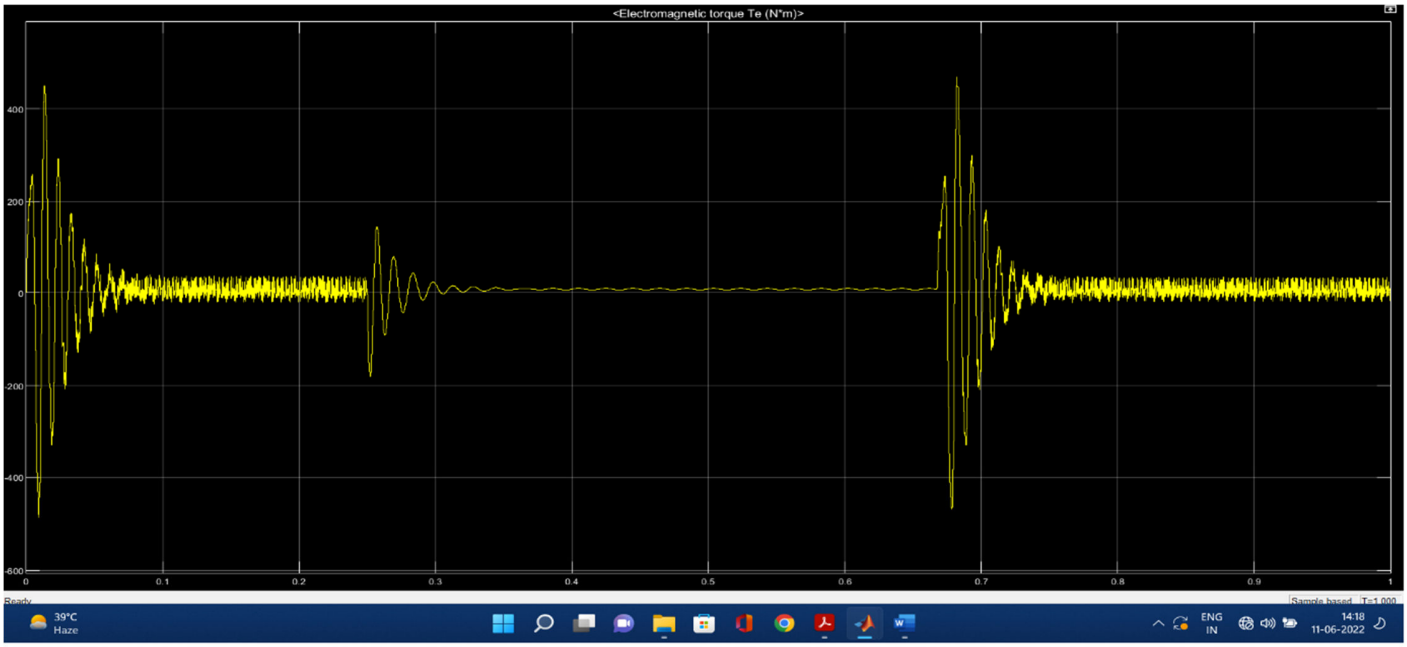


Fig 15: PMSM Torque for three phase short circuit condition

- ***OBSERVATION TABLE***

TABLE 4 – Comparison of different operating states

Type of Operation	SRM			PMSM (3-ph)		
	Max. Torque	Min. Torque	Torque Ripple	Max. Torque	Min. Torque	Torque Ripple
Normal Operation	107.1	-0.15e-05	1.442	448.64	-487.4	19.84

1-ph Short-Circuit Fault	107.1	-0.15e-05	1.422	448.64	-487.4	19.79
2-ph Short-Circuit Fault	107.1	-0.15e-05	1.421	448.64	-487.4	9.39
3-ph Short-Circuit Fault	107.1	-0.15e-05	1.416	469.0	-487.4	14.12

7. OPTIMISATION TECHNIQUE OF PMSM FOR ELECTRIC VEHICULAR APPLICATION

7.1. Normal Operation

Five-phase PMSM driven EV model was simulated in the normal condition and the electromagnetic torque output of the motor is shown below in Fig. 16.

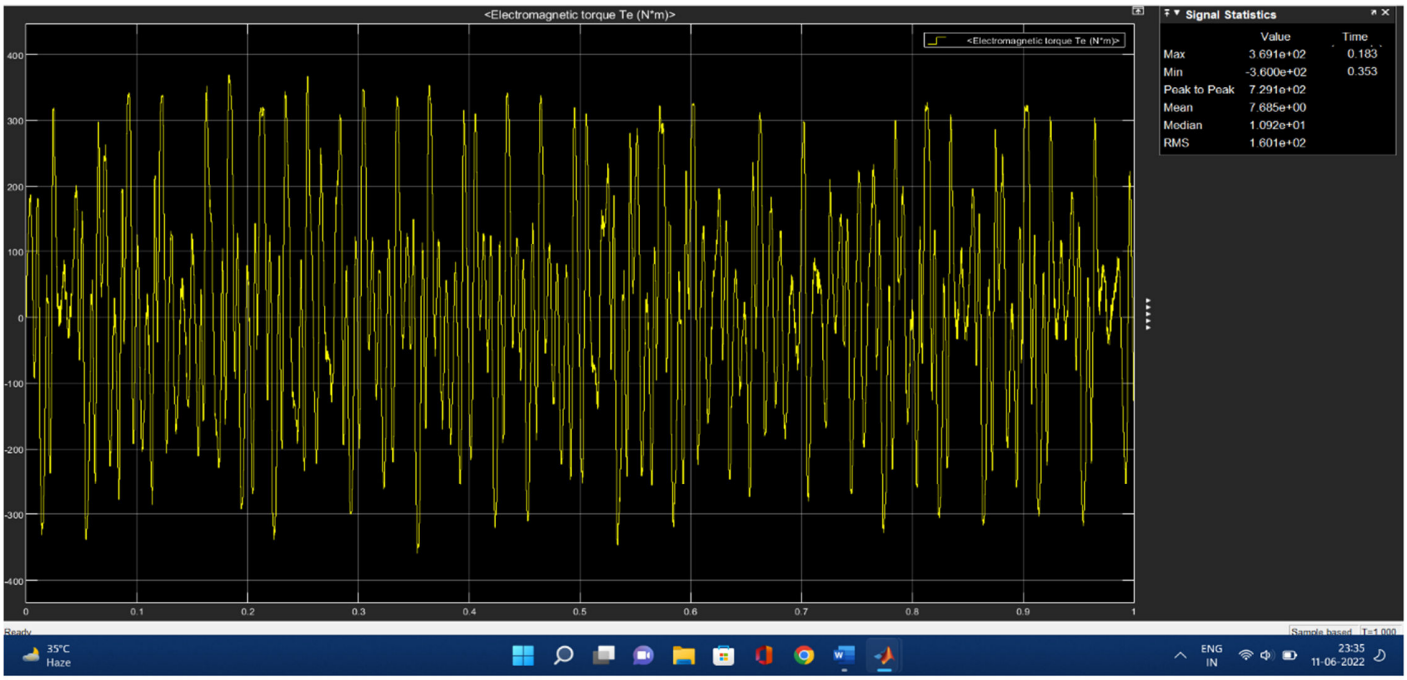


Fig 16: PMSM (5-phase) Torque for normal operation

7.2. Abnormal Operation

7.2.1. Single-phase short circuit fault

The single-phase short circuit fault was created using the 3-ph fault block (in the Simulink library) and was assumed to occur at 0.25 seconds. The fault was cleared at 0.67 seconds by entering value for the switching time in the block. The simulation result of the 5-ph PMSM motor is shown in Table 5. Fig. 17 below shows the fault characteristics of PMSM.

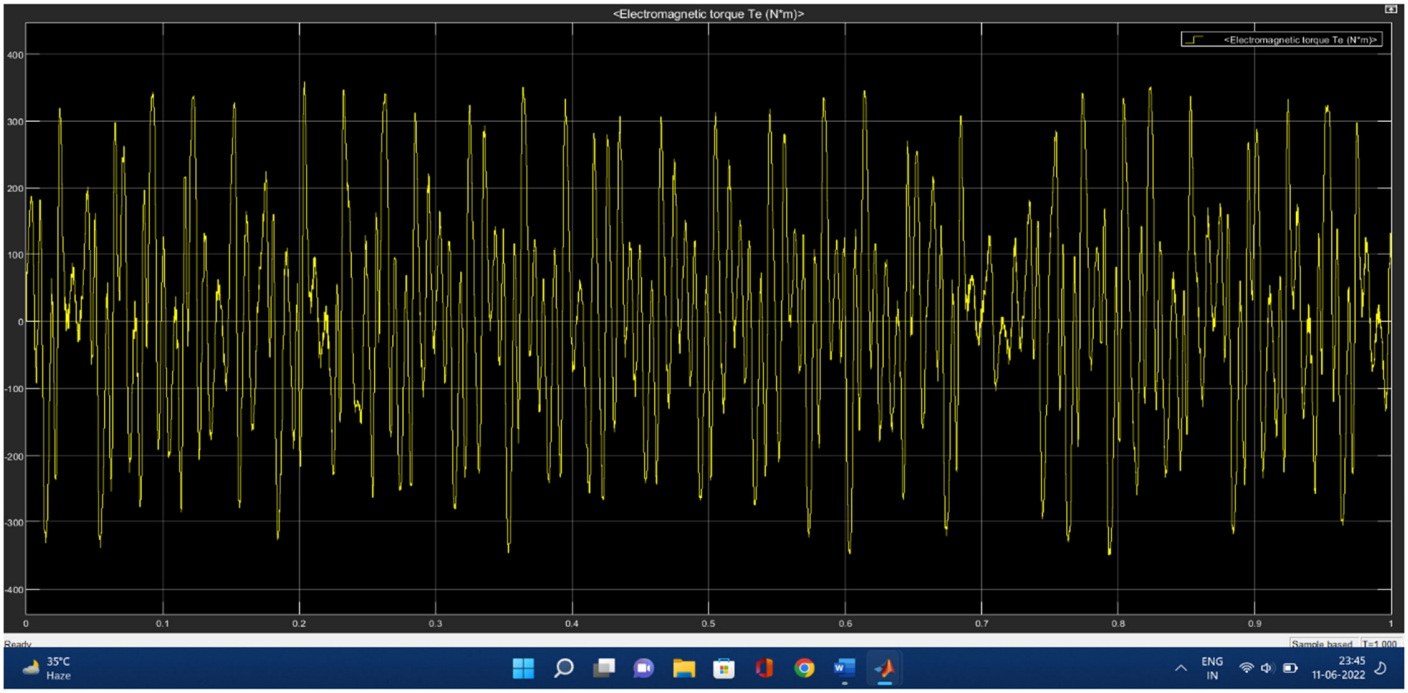


Fig 17: PMSM (5-phase) Torque for single phase fault

7.2.2. Phase-to-phase short circuit fault

The phase-to-phase short circuit fault was created using the 3-ph fault block (in the Simulink library) and was assumed to occur at 0.25 seconds. The fault was cleared at 0.67 seconds by entering value for the switching time in the block. The simulation result of the 5-ph PMSM motor is shown in Table 5. Fig. 18 below shows the fault characteristics of PMSM.

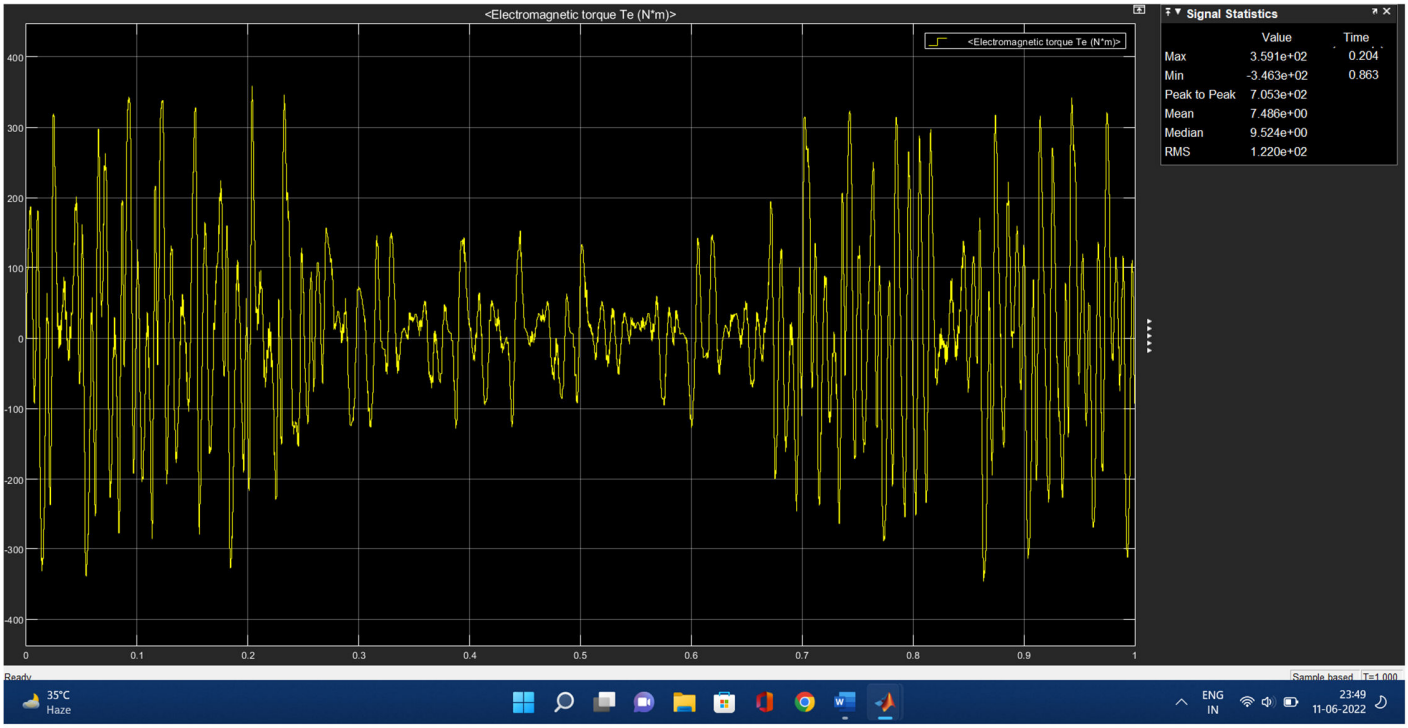


Fig 18: PMSM (5-phase) Torque for two phase fault

7.2.3. Three-phase short circuit fault

The three-phase short circuit fault was created using the 3-ph fault block (in the Simulink library) and was assumed to occur at 0.25 seconds. The fault was cleared at 0.67 seconds by entering value for the switching time in the block. The simulation result of the 5-ph PMSM motor is shown in Table 5. Fig. 19 below shows the fault characteristics of PMSM.

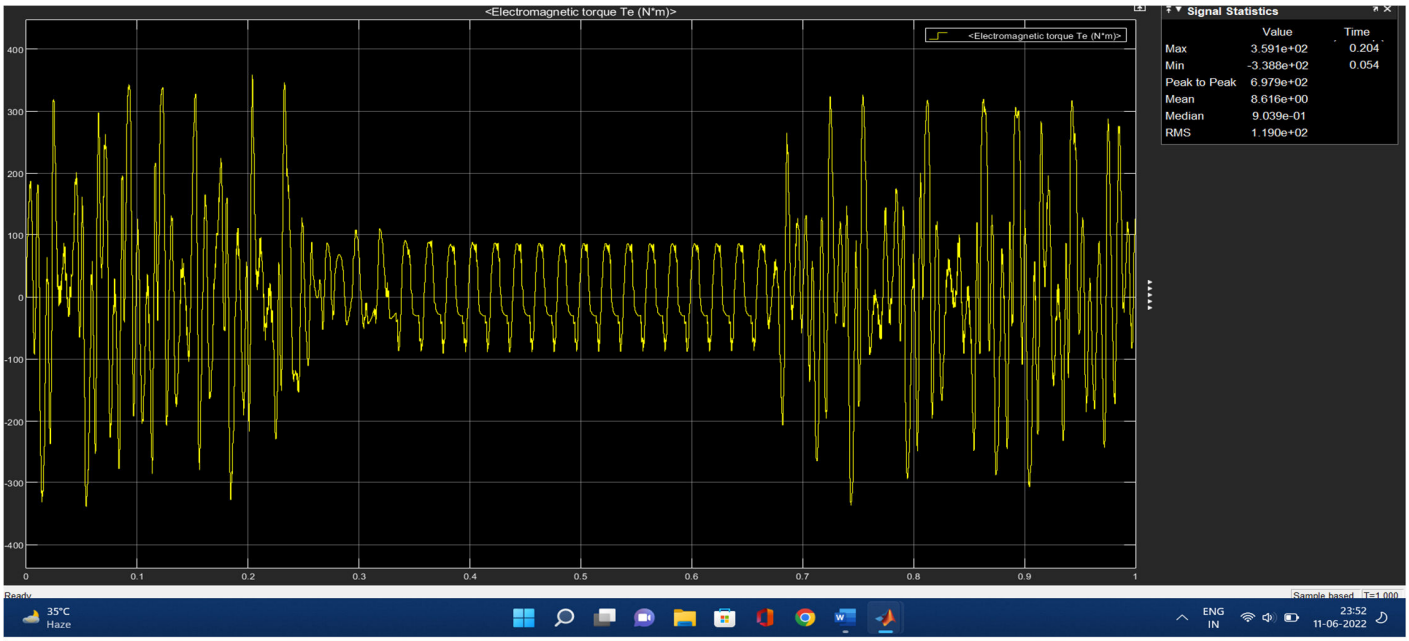


Fig 19: PMSM (5-phase) Torque for three phase fault

- **OBSERVATION TABLE**

TABLE 5 – Comparison of different operating states of 5-ph PMSM motor

Type of Operation	PMSM (5-ph)		
	Max. Torque	**Min. Torque**	**Torque Ripple**
Normal Operation	369.1	-360.0	4.554
1-ph Short-Circuit Fault	359.1	-349.5	4.586
2-ph Short-Circuit Fault	359.1	-346.3	5.781
3-ph Short-Circuit Fault	359.1	-338.8	5.864

8. CONCLUSION

The following conclusions can be drawn from the research conducted with respect to EV systems –

i) PMSM and SRM motors are the two most motors that are extensively preferred for electric vehicular systems.

ii) Both of these motors have their respective advantages and disadvantages. However, through market research, we conclude that PMSM is preferred over SRM due to its excellent torque density benefits.

iii) Through heatmaps drawn from exploratory data analysis conducted on a real-time data set gathered by the LEA department at Paderborn University, we conclude that temperature rise conditions and thereby faults are highly correlated with torque and quadrature current characteristics.

iv) Fault tolerance capability is examined by the torque ripple parameter as calculated and discussed in Table 4. We conclude that the fault tolerance of 3-Phase SRM is better than that of PMSM subjected to normal and short circuit faults under similar conditions of motor and battery ratings.

v) The optimisation technique proposed and simulated in section 7 demonstrates and proves to improve the fault tolerant capability of PMSM motor subjected to similar conditions as compared to the earlier results.

REFERENCES

[1] S. Enyedi, "Electric cars — Challenges and trends," 2018 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), 2018, pp. 1-8, doi: 10.1109/AQTR.2018.8402776.

[2] P. Rodge and K. Joshi, "Electric Vehicles in India: Current Status, Future Trend and Environmental Impact," 2018 International Conference on Smart Electric Drives and Power System (ICSEDPS), 2018, pp. 22-27, doi: 10.1109/ICSEDPS.2018.8536034.

[3] Snehalika and S. Das, "Emerging trends in Electric Vehicle in Indian market," 2019 3rd International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE), 2019, pp. 514-518, doi: 10.1109/RDCAPE47089.2019.8979012.

[4] M. Ehsani, K. V. Singh, H. O. Bansal and R. T. Mehrjardi, "State of the Art and Trends in Electric and Hybrid Electric Vehicles," in Proceedings of the IEEE, vol. 109, no. 6, pp. 967-984, June 2021, doi: 10.1109/JPROC.2021.3072788.

[5] C. C. Chan, "Present status and future trends of electric vehicles," 1993 2nd International Conference on Advances in Power System Control, Operation and Management, APSCOM-93., 1993, pp. 456-469 vol.1.

[6] X. Zhou et al., "The current research on electric vehicle," 2016 Chinese Control and Decision Conference (CCDC), 2016, pp. 5190-5194, doi: 10.1109/CCDC.2016.7531925.

[7] O. Stopka, M. Stopková, J. Ližbetin, J. Soviar and J. Caban, "Development Trends of Electric Vehicles in the Context of Road Passenger and Freight Transport," 2020 XII International Science-Technical Conference AUTOMOTIVE SAFETY, 2020, pp. 1-8, doi: 10.1109/AUTOMOTIVESAFETY47494.2020.9293526.

[8] B. Frieske, M. Kloetzke and F. Mauser, "Trends in vehicle concept and key technology development for hybrid and battery electric vehicles," 2013 World Electric Vehicle Symposium and Exhibition (EVS27), 2013, pp. 1-12, doi: 10.1109/EVS.2013.6914783.

[9] M. Marinelli et al., "Electric Vehicles Demonstration Projects - An Overview Across Europe," 2020 55th International Universities Power Engineering Conference (UPEC), 2020, pp. 1-6, doi: 10.1109/UPEC49904.2020.9209862.

[10] I. Husain et al., "Electric Drive Technology Trends, Challenges, and Opportunities for Future Electric Vehicles," in Proceedings of the IEEE, vol. 109, no. 6, pp. 1039-1059, June 2021, doi: 10.1109/JPROC.2020.3046112.

[11] X. Meng, H. Gao, W. Zhang and H. Liang, "Research on Fault Diagnosis of Electric Vehicle Power Battery Based on Attribute Recognition," 2020 5th Asia Conference on Power and Electrical Engineering (ACPEE), 2020, pp. 1321-1325, doi: 10.1109/ACPEE48638.2020.9136559.

[12] L. Wang, L. Wang, W. Liu and Y. Zhang, "Research on fault diagnosis system of electric vehicle power battery based on OBD technology," 2017 International Conference on Circuits, Devices and Systems (ICCDS), 2017, pp. 95-99, doi: 10.1109/ICCDS.2017.8120458.

[13] J. F. Ramahaleomiarantsoa, N. Héraud, O. Bennouna, E. J. R. Sambatra and J. M. Razafimahenina, "Modeling & fault diagnosis system for electric vehicles," IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society, 2012, pp. 4127-4132, doi: 10.1109/IECON.2012.6389229.

[14] D. Li, Z. Zhang, P. Liu, Z. Wang and L. Zhang, "Battery Fault Diagnosis for Electric Vehicles Based on Voltage Abnormality by Combining the Long Short-Term Memory Neural Network and the Equivalent Circuit Model," in IEEE Transactions on Power Electronics, vol. 36, no. 2, pp. 1303-1315, Feb. 2021, doi: 10.1109/TPEL.2020.3008194.

[15] R. Xiong, Q. Yu and W. Shen, "Review on sensors fault diagnosis and fault-tolerant techniques for lithium ion batteries in electric vehicles," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2018, pp. 406-410, doi: 10.1109/ICIEA.2018.8397751.

A Key Parameter in Ocean Thermal Energy Conversion Plant Design and Operation

Yongjian Gu, Ph.D., P.E.

The United State Merchant Marine Academy, 300 Steamboat Rd
Kings Point, New York, 11024. USA
Tel: (526)726-5719, Email: Guy@usmma.edu

***Abstract*—** Ocean thermal energy is one of the ocean energy sources. It is a renewable, sustainable, and green energy source. Ocean thermal energy conversion (OTEC) applies the ocean temperature gradient between the warmer surface seawater and the cooler deep seawater to run a heat engine and produce useful power output. Unfortunately, the ocean temperature gradient is not big. Even in the tropical and equatorial regions, the surface water temperature can only reach up to 28ºC and the deep water temperature can be as low as 4ºC. The thermal efficiency of the OTEC plants, therefore, is low. In order to improve the plant thermal efficiency by using the limited ocean temperature gradient, some OTEC plants use the method of adding more equipment for the better heat recovery, such as heat exchangers, pumps, etc. Obviously, the method will increase the plant complexity and cost. The more important impact of the method is the additional equipment needs to consume power too which may have an adverse effect on the plant net power output, in turn the plant thermal efficiency. In the paper, the author first describes varied OTEC plants and the practice using the method of adding more equipment for improving the plant thermal efficiency. Then the author proposes a parameter, plant back work ratio ϕ, for measuring if the added equipment is appropriate for the plant thermal efficiency improvement. Finally in the paper, the author presents examples to illustrate the application of the back work ratio ϕ as a key parameter in the OTEC plant design and operation.

***Keywords*—** Ocean thermal energy, Ocean thermal energy conversion (OTEC), OTEC plant, plant back work ratio ϕ

Vermicomposting Amended with Microorganisms and Biochar: Phytopathogen Resistant Seedbeds for Vegetables and Heavy Metal Polluted Waste Treatment

Fuad Ameen, Ali A. Al-Homaidan

***Abstract*—**Biochar can be used in numerous biotechnological applications due to its properties to adsorb beneficial nutrients and harmful pollutants. Objectives: We aimed to treat heavy metal polluted organic wastes using vermicomposting process and produce a fertilizer that can be used in agriculture. We improved the process by adding biochar as well as microbial inoculum and biomass into household waste or sewage sludge before vermicomposting. The earthworm Eisenia fetida used in vermicomposting was included to accumulate heavy metals, biochar to adsorb heavy metals, and the microalga Navicula sp. or the mangrove fungus Acrophialophora sp. to promote plant growth in the final product used as a seedbed for Solanaceae vegetables. We carried out vermicomposting treatments to see the effect of different amendments. Final compost quality was analyzed for maturity. The earthworms were studied for their vitality, heavy metal accumulation, and metallothionein protein content to verify their role in the process. The compost was used as a seedbed for vegetables that were inoculated with a phytopathogen Pythium sp. known to cause root rot and destroy seeds. Compost as seedbed promoted plant growth and reduced disease symptoms in leaves. In the treatment where E. fetida, 6% biochar, and Navicula sp. had been added, 90% of the seeds germinated, while less than 20% germinated in the control treatment. The experimental plants had acquired resistance against Pythium sp. The metagenomic profile of microbial communities will be reported.

***Keywords*—**organic wastes, vermicomposting process, biochar, mangrove fungus.

Fuad Ameen Hasan is with the King Saud University, Saudi Arabia (e-mail: fuadameen@ksu.edu.sa).

Hydrogeochemical Assesment of Groundwater in Selected Part of Benue State Southern, Nigeria

[1,2]Moses Oghenenyoreme Eyankware (Ph.D), Christian Ogbuchi Ede
[1]Department of Geology, Ebonyi State University, Abakaliki. P.M.B. 053, Ebonyi State Nigeria
[2]Geomoses Consultancy Limited Warri, Delta State Nigeria
Email: geomoses203@gmail.com

Moses Oghenenyoreme Eyankware: ORCID.org/0000-0003-0221-1138

Abstract: Groundwater is the principal source for various use in this study area. The quality and availability of groundwater depends on rock formation within the study area. To effectively study the quality of groundwater, 24 groundwater samples were collected. The study was aimed at investigating the hydrogeochemistry of groundwater, and additionally its suitability for drinking and irrigation purposes. The following parameters were analyzed using American Public Health Association standard method: pH, turbidity, Ec, TDS, Mg^{2+}, SO_4^{2-}, NO_3^{-}, Cl^{-}, HCO_3^{-}, K^{+}, Na^{2+} and Ca^{2+}. Results obtained from Water Quality Index revealed that groundwater sample fell within good water quality that implies that groundwater is considered fit for drinking purpose. Deduced results obtained from irrigation indices revealed that Permeability Index (PI), Soluble Sodium Percentage (SSP), Sodium Percentage (Na %), Sodium Absorption Ratio (SAR), Kelly Ratio (KR), Magnesium Hazard (MH) ranges from 0.00 to 0.01, 4.04 to 412.9, 0.63 to 257.7, 0.15 to 2.34, 0.09 to 2.57 and 6.84 to 84.55 respectively. Findings from Total hardness revealed that groundwater fell within soft, moderate hard and hard category. Estimated results obtained from CSMR, RI and LSI showed that groundwater showed corrosion tendency, salinization influenced groundwater at certain sampling points and chloride and sulfate unlikely to interfere with natural formation film.

Keywords: Water, Quality, Suitability, Anthropogenic, Nigeria.

Investigating the Impact of Individual Risk Willingness and Group-Interaction Effects on Business Model Innovation Decisions. A Mixed-Method Approach

S. Müller-Sägebrecht

***Abstract*—**Today's volatile environment challenges executives to make the right strategic decisions to gain sustainable success. Entrepreneurship scholars postulate mainly positive effects of environmental changes on entrepreneurship behavior, such as developing new business opportunities, promoting ingenuity, and the satisfaction of resource voids. A strategic solution approach to overcome threatening environmental changes and catch new business opportunities is business model innovation (BMI). Although this research stream has gained further importance in the last decade, BMI research is still insufficient. Especially BMI barriers, such as inefficient strategic decision-making processes, need to be identified.

Strategic decisions strongly impact organizational future and are therefore usually made in groups. Although groups draw on a more extensive information base than single individuals, group-interaction effects can influence the decision-making process - in a favorable but also unfavorable way.

Decisions are characterized by uncertainty and risk, whereby their intensity is perceived individually differently. The individual risk willingness influences which option humans choose. The special nature of strategic decisions, such as in BMI processes, is that these decisions are not made individually but in groups due to their high organizational scope. These groups consist of different personalities whose individual risk willingness can vary considerably. It is known from group decision theory that these individuals influence each other, observable in different group-interaction effects. The following research questions arise: *i) Which impact has the individual risk-willingness on BMI decisions? And ii) how do group-interaction effects impact BMI decisions?*

After conducting 26 in-depth interviews with executives from the manufacturing industry, the mixed-method approach reveals the following results: i) Risk-averse decision-makers have an increased need to be guided by facts. The more information available to them, the lower they perceive uncertainty and the more willing they are to pursue a specific decision option. However, the results also show that social interaction does not change the individual risk willingness in the decision-making process. ii) Generally, it could be observed that during BMI decisions, group interaction is primarily beneficial to increase the group's information base for making good decisions, less than for social interaction. Further, decision-makers mainly focus on information available to all decision-makers in the team but less on personal knowledge.

This work contributes to strategic decision-making literature twofold. First, it gives insights into how group-interaction effects influence an organization's strategic BMI decision-making. Second, it enriches risk-management research by highlighting how individual risk-willingness impacts organizational strategic decision-making.

To date, it was known in BMI research that risk aversion would be an internal BMI barrier. However, with this study, it becomes clear that it is not risk aversion that inhibits BMI. Instead, the lack of information prevents risk-averse decision-makers from choosing a riskier option. Simultaneously, results show that risk-averse decision-makers are not easily carried away by the higher risk willingness of their team members. Instead, they use social interaction to gather missing information. Therefore, executives need to provide sufficient information to all decision-makers to catch promising business opportunities.

***Keywords*—**Business model innovation, group biases, group decisions, group-interaction effects, mixed-method approach, risk willingness.

S. Müller-Sägebrecht is a research associate and PhD candidate at Technische Universität Dresden, Chair of Entrepreneurship and Innovation, 01062 Dresden, Germany since 2016 (e-mail: sarah.mueller-saegebrecht@tu-dresden.de).

Berry Phase and Quantum Skyrmions: A Loop Tour in Physics

Sinuhé Perea-Puente*
(Dated: March 30, 2022)

In several physics systems the whole can be obtained as an exact copy of each of its parts, which facilitates the study of a complex system by looking carefully at its elements, separately. Reducionism offers simplified models which makes the problems easier, but "there's plenty of room...at the *mesoscopic* scale". Here we present a tour for two of its representants: Berry phase and skyrmions, studying some of its basic definitions and properties, and two cases in which both arise together, to finish constraining the scale for our mesoscopic system in the quest of quantum skyrmions, discovering which properties are conserved and which others may be destroyed.

I. INTRODUCTION

"I have here only made a nosegay of culled flowers, and have brought nothing of my own but the thread that ties them together."
Michel De Montaigne (1533-1592).

Here we try to explore some concepts from Condensed Matter and Quantum Mechanics and make a little research activity which is presented below. We will start by introducing the basic ideas from *Berry phase*, the 'harmless' quantum phase that was ignored nearly sixty years [1] and appears in the case of adiabatic limit in Quantum Mechanics. In particular, later we will study a physical case in which the next concept will appear naturally.

Next, the notion of *skyrmions*, an object in between Physics and Mathematics in the context of solitons, discovered in the context of nuclear physics [2] but later applied in many disciples will be presented with special interest in two of its properties: '*topological protection*' and '*locally energetically favourable*', concepts explained with more detail in Appendix A,B respectively, where a deeper mathematical study will be offered, from *homotopy* concept in Topology to variational minimization of the Hamiltonian.

In the third section we will present some models that merge both concepts introduced previously, Berry phase and skyrmions, in the case of an electron trespassing a smooth magnetic structure and later on studying the effects of the scale in this mesoscopic phenomena, trying to explore some properties underlying a *quantum skyrmion*.

Still nowadays an interesting open [3] debate of the necessary and sufficient ingredients for considering an object quantic; maybe is the size a critical parameter (despite macroscopic quantum phenomena), maybe its mathematical formalism (non-conmutativity observables in a Hilbert space or due to the compulsory imaginary numbers presence [4]), or perhaps is just the presence of some characteristic phenomena (e.g. presence of photons, wave-functions or entanglement).

Trying to avoid this philosophical debate, the fact is that Berry phase offers a strong candidate of the latter case and allow us to link this report with the subject: we will start by discussing the adiabatic changes, developing the 'slow' evolution of a quantum system. When this path was considered over a closed curve, we discovered that the adiabatic phase was indeed Gauge invariant and therefore has an intrinsic physic interest despite the paradigm [1] in which was initially proposed [5] and finally was re-evaluated [6] nearly sixty years later.

Although later we will see that this emergent behaviour may be exhibited in classical objects, our focus will be dynamic evolution of a quantum system that moves in the parameter space given by parameters $R(t)$, which is well-known and described by Schrödinger's equation

$$i\hbar\partial_t \left|\psi(t)\right\rangle = H(R(t)) \left|\psi(t)\right\rangle . \tag{1}$$

From here, assuming an *adiabatic* evolution, an additional phase γ arise, in the following form:

$$\left|\psi(t)\right\rangle = \sum c_n e^{-i\int_0^t E_n(t')dt'} e^{i\gamma_n(t)} \left|n; R(t)\right\rangle , \tag{2}$$

in which the sum is used to form the initial state $\left|\psi(t_0)\right\rangle = \sum c_n \left|n; R(t_0)\right\rangle$, E_n corresponds to the energy of the Hamiltonian and the 'extra' factor $\gamma_n(t)$ obeys to the *Berry phase*, which in closed loops can't be 'gauged' away and represent a truly degree of freedom of the system. The following related definitions for the *Berry connection* $\mathbf{A}^n(R)$ is:

$$\gamma_n(C) = \int_{R(0)}^{R(t)} \mathbf{A}^n(R)\cdot dR = i\int_{R(0)}^{R(t)} \left\langle n; R\right| \boldsymbol{\nabla}_R \left|n; R\right\rangle \cdot dR, \tag{3}$$

* sinuhe.perea@kcl.ac.uk

also known as *Berry vector potential*, as we will see later. From here, in the case when $R(t)$ correspond exactly to a multiple of the period, i.e., if we follow a closed path, the Berry phase is uniquely determined up to a $2k\pi$ phase due to the analytic branches of the exponential function. From *Kelvin-Stokes Theorem*, it is also possible to define a *Berry curvature* $F^n(R)$ which will correspond to the magnetic field, in the electrodynamics image, for being the rotational of the vector potential, previously introduced.

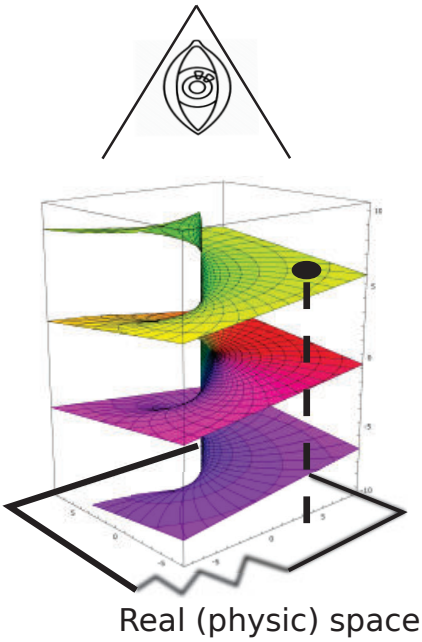


FIG. 1. In a naive mental picture, this Berry phase is represented in the branches of an analytic complex function. In the adiabatic closed case, this phase show physical value which is 'measured' (and represented as the projection into the 'physical' plane) while for faster evolution can be retrieved and stands for just random fluctuations, in similarly of some alternatives in initial *path-integral* chapter of the module.

With this basic ingredients in mind, now we introduce the main goal of this report, the *skyrmions*.

II. SKYRMIONS. ORIGIN AND FIRST DEFINITIONS

We will start by commenting the figure of Tony H. Skyrme from which the skyrmions take its name. English physicist from the second part of the last century, and member of the *Manhattan project*, he focused on nuclear physics in which he initially work on short-range forces in the nuclei, an later develop a theory of non-linear sigma model nuclear forces, lately known as the *Skyrme model* [2], where the nucleus was filled with a 'medium' caused by three pion fields, remarkably predicting the incipient *QCD* theory.

As it is well-known, classically in mechanics, particles correspond to an ever-lasting building blocks with non-zero mass, but in *QFT* approximation, in a closer look, they consist on wave-like excitations [7] of a fundamental field. The idea now is how we made this functions not to disappear, as they would naturally decay.

Here is where the concept of *soliton* will be related with the skyrmions and again by another British author, the Scottish engineering J.S Russell. In the beginnings of S.XIX, building water channels in Great Britain, he discovers that although in general dumped boats needed to be pulled because its water waves below rapidly decay, in certain narrow parts, the boat just continues autonomously its travel for really long distances (kilometres!). So something was happening with the water 'wave of translation' [8].

In a more general way, we can characterize [9] this solitons by the following properties: they are *localized* (i.e. finite energy and therefore finite size), *stable* (they have a constant shape in time as a solution of the PDE) and *algebraic* (meaning that the can interact with other solitons and recombine, collide and emerge with a phase shift or collapse).

Before we formally introduce the skyrmions, we will introduce another example of solitons, with a classical object: the pendula (which, indeed, will appear again during the next sections). Considering a single harmonic pendulum i, with given mass m, known length l and phase tilting ϕ under the force of gravity, we can describe this system using its known Lagrangian

$$L_i = \frac{1}{2}\partial_t^2\phi_i - (1 - cos\phi_i) \tag{4}$$

Here the notation ∂_u for the partial derivation with respect to u is adopted. Now consider an arbitrarily large amount (let's say n) of nearest neighbors-coupled pendula, for example joining them with elastic bands (it is also possible just to think on a *Newton's cradle*, although there the oscillations are not perpendicular to the axis of the pendula, so the image is less clear). If we now rotate 2π the first pendulum, this perturbation will be transmitted, ideally without losses, to every single neighbour till the end, in a solitary wave [10].

From mechanically describing this system, we

solve *Euler-Lagrange equations* [9] for each of the ϕ_i perturbations, so that if we consider it 'small enough', the solution for the scalar field $\phi(t, r)$ is

$$\partial^2\phi + sin\phi = 0, \tag{5}$$

in which now the ∂ notation stands for all the spatio-temporal components and receive the name of *sine-Gordon equation*, in analogy with *Klein-Gordon equation* (in the QFT notation). As usually happens in PDE, boundary conditions are extremely important. In particular, in the previous example, we started with a complete rotation $\phi_0(0) = 2\pi$ (or any integer number, negative included which would lead to clock-anticlockwise rotation) while in the final pendulum $\phi_n(0) = 0$. This toy-model can be exploited more, in higher dimensions and can also analytically solved, in what is called *kink* and *anti-kink solutions* depending on its direction, but for us is far from our objective.

The beauty idea is what comes next: *what happens if the large amount of pendula are arranged in a close circuit?* Note that the proposal seems quite similar to the previous section. In that case, we will have an infinite perturbation in the loop which travels periodically. It is the same idea as when we have a closed rope with a knot (or more) on it. We can not 'untangled' to delete the knot...it is an intrinsic *defect* (in fact it is a *topological* defect, see Appendix A for further details).

This is the concept that it is exploited in the skyrmion case, an emergent figure that appears from well-known physics in the mesoscopic regime. Forming defects that can not be easily untangled, so the system tries to evolves to integrate this defect and forming complex (exotic) structures. A well-known example is the Neel and Bloch domain walls that needs to break the spatial symmetry of a constant direction by re-absorbing smoothly the abrupt change (just exactly as the pendula above). In more dimensions, the structure has more degrees of freedom to form (restricted, however, to *Toulouse-Kleman* [11] relations) a structure that is protected (and also corresponds to a local minimum of the energy, see below). For r-dimensional materials with a single defect, the configuration that maximizes the dimension of the parameter space ($s = r - 1$) are the vortex and the skyrmions (corresponding to two and three dimensions, respectively).

The quantity that fully characterizes the skyrmions corresponds to the number of times the field configuration *wraps* the unitary sphere (in lower dimensions this is just the case of the infinite pendula, where the quantity was just the number of whole rotations given to the initial pendulum which is, in absence of losses, also conserved), and it's given by the *winding number* [12] also known as *topological charge*, *Brouwer index* or *topological quantum number* in different contexts,

$$n = \frac{1}{4\pi}\int \mathbf{F}\cdot(\partial_x\mathbf{F}\times\partial_y\mathbf{F})\,dS \tag{6}$$

where $\mathbf{F}$ stands for the smooth (continuous) field configuration and we have considered a xyz Cartesian system. The sign of it can be positive or negative, depending on the orientation of the vector fields, which leads to talk about (positive) *skyrmions*, with a positive winding number, or *antiskyrmions*, with a negative one, in analogy with the 'topological charge' image. It is invariant under smooth transformation and therefore skyrmions are *topologically protected* (see Appendix A for further details), and therefore we prove it would not decay in time.

Now two key questions may arise: first, *if the skyrmions are intrinsically stable and topological protected so they can not decay, how it's possible the reverse process, i.e, the creation of them?* And second, we have seen that an important feature was the smoothness of the functions involved. *When we decrease our scale, for example to the quantum limit, and therefore continuous functions became discrete, it is still possible to have some kind of quantum skyrmions?*

The first question relies on the fact that all the ideas presented previously are *mathematically* accurate, assuming an infinite discretization and allowing infinity values but, in physics [13], we have firm spatial limits and a infinite values can not be obtained in the lab [14] so the topological protection must be substituted with the notion of a ('big') *potential barrier*, and therefore can be created (or destroyed) applying 'enough large' perturbations. Also, *topologically stable* don't immediately means *energetically stable*, and a study of the Hamiltonian must be considered.

Some example of the energetic consideration (static limit) and time evolution (dynamics) for a magnetic skyrmion is studied in Appendix B, which would be the case study in next section, and where the second question will be fully addressed.

III. BERRY PHASE FROM SLOW ELECTRONS IN A MAGNETIC FIELD TO QUANTUM SKYRMIONS

Starting with the mathematical concept of the *parallel transport* of a vector in a closed curve and the *Gauss-Bonnet Theorem* (for example in the case

of a triangle placed in the surface of the sphere S_2) we can think about the first physical example that offers a practical example of the Berry phase, the *Foucault pendulum* (as it can be seen, oscillating phenomena usually offers interesting physics development), where the precession movement within the rotation of the Earth allow us to see a macroscopic realization of this accumulated geometric phase. From here we can move on to the next example, in a lower scale range, and where a 'lazy' electron crosses a magnetic smooth texture.

Now, we want to study the behaviour of this electron traversing a magnetic structure. In most cases, the *Fermi velocity* makes that the electron don't 'feel' the magnetization and its magnetic moment is not affected. By contrast, when we consider an *adiabatic limit*, this moment recalibrates and adjusts to the local magnetization landscape. For this, we re-consider (1) with

$$\mathbf{H} = \frac{\mathbf{p}^2}{2m} + J\frac{g\mu_B}{2}\boldsymbol{\sigma}\cdot\mathbf{m}(r,t) \tag{7}$$

and where $\mathbf{m}(r,t) = \mathbf{M}(r,t)/|\mathbf{M}|$ corresponds to the local magnetization [15], J corresponds to the exchange coupling (where its sign distinguish ferro/antiferromagnetism), the term in the fraction corresponds to the magnetic moment of the electron: the Bohr magneton μ_B, Lande factor g , and $\boldsymbol{\sigma}$ are the Pauli matrices for the spin of the fermion. From here, it is possible to analyse [16] separating the field into two different components. The first one corresponding to a uniform Zeeman magnetic field and the seconds is an additional *emergent* [16] electromagnetic fields given by the components

$$\mathbf{E_i^{eff}} = \hbar\mathbf{m}\cdot(\partial_i\mathbf{m}\times\partial_t\mathbf{m}) \tag{8}$$

$$\mathbf{B_i^{eff}} = \hbar\epsilon_{ijk}\mathbf{m}\cdot(\partial_j\mathbf{m}\times\partial_k\mathbf{m})/2 \tag{9}$$

The next step is just trying to make (7) look simpler by analysing the dot product and trying to align our reference system to just have the z-component s.t.

$$J\frac{g\mu_B}{2}\boldsymbol{\sigma}\cdot\mathbf{m}(r,t) = J^{eff}\sigma_z \tag{10}$$

with the constant $J^{eff} = Jg\mu_B\hbar m/2$. This can be achieved by considering the transformation $\psi = U(r,t)\psi^{eff}$ with $U = exp(-i\theta\boldsymbol{\sigma}\cdot\mathbf{n})$ an unitary rotation with $\theta = cos^{-1}(\mathbf{m}\cdot\mathbf{z})$ and $\mathbf{n}$ the corresponding rotation axis perpendicular to the magnetization and z. Applying the transformation to (1) we have

$$i\hbar\partial_t\left|\psi(t)^{eff}\right\rangle = H\left|\psi(t)^{eff}\right\rangle \tag{11}$$

$$\mathbf{H} = \left(\frac{(\mathbf{p}-e\mathbf{A}^{eff})^2}{2m} + e\mathbf{V}^{eff} + J\frac{g\mu_B}{2}\boldsymbol{\sigma}\cdot\mathbf{m}(r,t)\right) \tag{12}$$

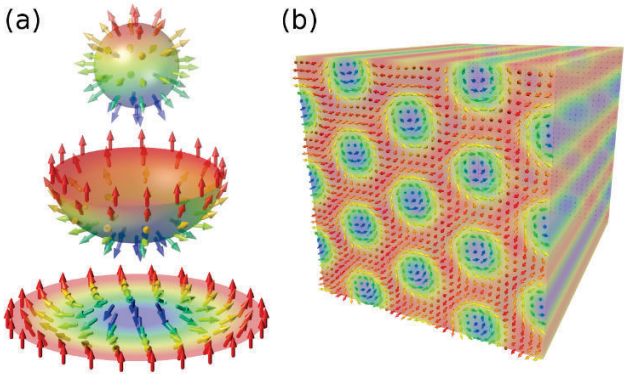


FIG. 2. (a) Hedgehog smooth unfolding in S_2 (above) to its stereographic projection into the plane (below) embedded in a ferromagnetic material. The north pole corresponds to point-to-infinity magnetization, which in this case is trivial (majority spin). The S_2 to R^2 mapping its possible thanks to Alexandroff compactification, as the difference space is numerable (just one point, the north pole itself). The arrows stand for the direction of the spin, while the colour code correspond to the in-plane component: from red (normal and positive), green for in-plane direction to blue (normal and negative). (b) Schematic bulk skyrmion lattice, requiring chiral non-centrosymmetric materials to exhibit a non-negligible DMI. Extracted from [17].

in the equation above, $(\mathbf{V},\mathbf{A})^{eff}$ appears as *effective scalar* and *vector potentials* given by [16]

$$\mathbf{V}^{eff} = -i\frac{\hbar}{e}U^\dagger\partial_\mathbf{t}U \tag{13}$$

$$\mathbf{A}^{eff} = -i\frac{\hbar}{e}U^\dagger\nabla U \tag{14}$$

respectively, in analogy with the Hamiltonian of an electron under the action of a electromagnetic field. With this, considering $c = -\hbar/e$ we can see how this operators acts on the two spin up/down states:

$$\langle\uparrow|A^{eff}|\uparrow\rangle = \left(\langle\uparrow|U^\dagger\right)ic\nabla\left(U|\uparrow\rangle\right) = i\langle\psi|c\nabla|\psi\rangle \tag{15}$$

$$\langle\uparrow|V^{eff}|\uparrow\rangle = \left(\langle\uparrow|U^\dagger\right)ic\partial_\mathbf{t}\left(U|\uparrow\rangle\right) = i\langle\psi|c\partial_\mathbf{t}|\psi\rangle\,, \tag{16}$$

and similarly with the $\downarrow$ states. Both expressions look exactly equal to the Berry potential introduced in the initial section. Also, the relation with the skyrmions appears clear with expressions (9) in which we see that exactly corresponds to the winding number when we consider as our field the magnetization, $\mathbf{F} = \mathbf{m}(r,t)$.

Here, this stable non-trivial configuration is obtained for a ferromagnetic material (let's say

wlog. with a majority of up states), it is possible to include a wrap of the magnetization in a locality without affecting the rest (see Figure 2 (a) below, where the boundary conditions are just the trivial ferromagnetic states).

Again, we see that the scale is crucial, as several components have a protagonist role that must not opaque the other components to form this exotic swirls: size of the unit cells, wavelength of the electrons, hierarchy and characteristic range of the different interactions [18]. All of above make the size of the skyrmion a critical parameter. In the adiabatic limit, this should be much bigger than the average spin-flip range of the electron to take place. Moreover, due to the additional protection of this kind of emergent fields, it is possible to perfectly control the magnetic flux on each unit cell, as long as the topological charge is quantized (as it is an topological invariant).

It is also worth mentioning that this electromagnetic fields causes a *Lorentz force* and therefore the dynamics of the skyrmion are quite interesting (see Appendix B for the theoretical derivation to the *Landau-Lifshitz-Gilbert equation* and the different ways to 'move' the skyrmion); in particular, we expect the non co-linear forces to make the trajectory non-trivial. In particular, perpendicular forces causes a topological Hall effect [19] and a *'Magnus' effect* (in analogy with hydrodynamics) that could be observed in a simulation using the software *OOMMF*. Here we develop the *ELSID* code that allow us to 'see' the dynamics of the system.

In Figure 3 we simulate a nanoscale display where a skyrmion is manually created (left) and we apply an electric current (STT) to the right to observe the trajectory of the skyrmion which deviate from horizontal direction due to the vectorial components involved. It is interesting [17] to make a classical analogy with a spinning football, in which the rotation the object causes a deviation (the original *Magnus effect*). Here, the magnetization patter acts 'as if' it rotates, as another beauty mental image, although the metaphor is not complete because here we have a monotone deviation that only makes the skyrmion collapse with the boundaries and it is not re-conduct so well as in the case of 'banana-shots'; also the spin is not conserved in this case.

From here, as now we have 'seen' the answer to the first question simulating it, we are now ready trying to solve the second question formulated in previous section.

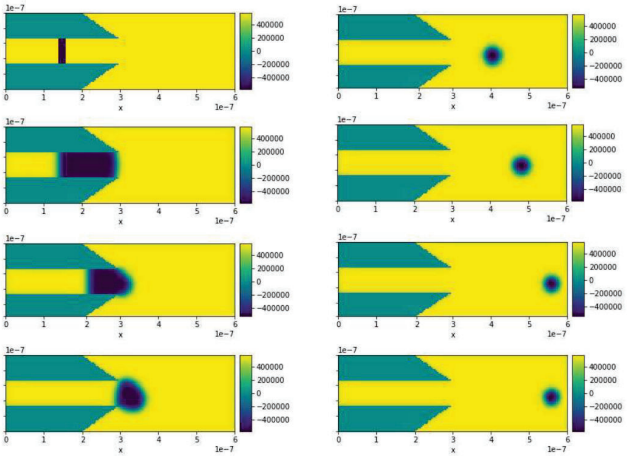


FIG. 3. Numerical simulation using OOMMF of a skyrmionic structure with the code *ELSID*. The skyrmion is created in the top left and evolves up-to-down, and left-right, and a spin-transfer-torque current is applied, causing the skyrmion to describe a trajectory, as a 'pseudo-particle'; vectorial components causes a perpendicular deviation from original applied current (left-to right) modifies the Hall signal. The size of the racetrack is 600nm long and 150nm width, with a thickness of 1nm. Many degrees of freedom can be controlled, like the channel angle transition, the intensity of the current or the initial size of the skyrmion. Effective control of the skyrmion could lead to a potential spintronic device of storage and transmission of information.

IV. THE QUANTUM LIMIT OF SKYRMIONS

Up to this point, we have seen that the scale plays a fundamental role in this kind of mesoscopic objects. The idea is now trying to see the 'boundaries' of this skyrmion theory. We could go to bigger scales, but that would abandon the scope of the modulus so we are going to put our attention in diminishing the size, to reach quantum-like scale. As in this regime the properties of the skyrmion will collide with quantum phenomena like tunnelling [20], annihilation process that destroys the topological protection in low-dimensional regimens or Heisenberg's uncertainty principle and maybe affect the winding number measurement and precision, and therefore it loses its integer-like behaviour.

But again, *is it still possible to observe, quantum-size skyrmions*?

The quantum skyrmion should be in the order of nanometres, in which some smooth properties may become more edgy, due to the discrete structure. Therefore, a quantum skyrmion should be a bound state with quantized properties that can be connected to the 'classic' definition in the con-

tinuum limit. The problem here is solving whether the minimum state corresponds to a *Landau level* available. Many additional question should be also solved [21] as the influence of quantum fluctuations, tunneling effects, or the notion of defects in this scale, to cite some examples.

Some assumptions should be initially made for the analysis of this structure. In the literature [22], many different conditions for the system are used. Here we will consider a symmetric ferromagnetic coupling for nearest-neighbour with exchange parameter J_1 (see Appendix B) and an opposite anti-ferromagnetic interaction for next-nearest neighbour with intensity proportional to J_2.

This model is applied in a hexagonal flake-like lattice with 31 sites surrounded by ferromagnetic boundaries, as is described in Figure 4. In principle, it should have been simulated using the *ELQSID* code develop, but it is currently not working. The quantum numbers of this quantum system correspond [21] to the flipped spins $N_f = S_z^{fm} - S_z$ (with $S_z^{fm} = sN = 31/2$ the trivial ferromagnetic solution) and the angular momentum due to the existence of a hexagonal rotation symmetry, and characterize by l_z.

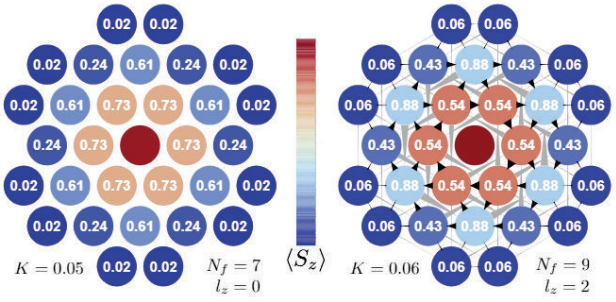


FIG. 4. Numerical simulation of quantum skyrmions in ferromagnetic structure. States are characterized by the number of flipped spins N_f and angular momentum l_z. Color map represent the z-component of the spin and numbers in the circles are related with the correlation of the spin in the plane (a) $|N_f = 7, l_z = 0\rangle$ with $K = 0.05$ represents a superposition of skyrmion and antiskyrmion and (b) $|N_f = 9, l_z = 2\rangle$ with $K = 0.06$ corresponds to a quantum skyrmion. Figure extracted from [21].

Now we need to prove that this quantum skyrmion state exists in reality and for this we evaluate whether it has a lower energy than the alternative near states, i.e

$$\langle N_f, l_z | E | N_f, l_z \rangle < \langle N_f - \lambda, l_z | E | N_f - \lambda, l_z \rangle + \lambda E_{min}$$

with $\lambda \in (1, 2, ..., N_f)$ and where E_m corresponds to the minimum energy for the change of the spin. It is also possible to simplify the RHS term by considering a *binding energy* defined as $E^B(N_f, l_z) = E(N_f, l_z) - N_f E_{min}$. A numerical simulation is showed in Figure 5, where it is seen that, under some limitations over the J_2, K parameters, negative energies are possible with (a) local minimum exactly at $N_f = 7$, which seems to be a physical quantum skyrmion realization. Nevertheless, in the case of the (b) anysotropy coupling, a large value would lead to trivial constant spin, as the line tends to decrease with N_f. Also another phenomena like spin-orbit, dipolar interactions or excitation of phonons can interfere in practical materials and change the results, although, in principle, single spin change is not allowed as we reach a local minimum.

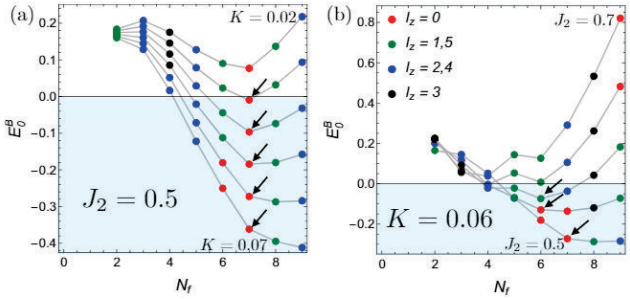


FIG. 5. Binding energy as function of the flipped spins N_f under the influence of a variation in the (a) anysotropy coefficient and (b) antiferromagnetic coupling are evaluated, where the colours represent different values of possible angular momentum. Arrows point to the lowest density (by unit) binding energy, which would cover the continuum limit. Figure extracted from [21] .

So although this unusual quantum skyrmion are robust against *quantum evaporation* for the flipping spins and therefore can exist, we have the pay-off of losing the topological protection before mentioned and also quantum PI causes the winding number of the quantum skyrmion being no longer an integer, and depends on the actual eigenvalues of N_f, l_z. To 'close the loop', Berry phase also appears in this quantum skyrmions. In particular, at least when the tunnelling of skyrmion-antiskyrmion is studied [20, 21] in the context of condensed matter physics and band structure.

V. CONCLUDING REMARKS

Within this term paper, we have introduced the skyrmion and related with the Berry phase in some electromagnetic and quantum situations.

Also, we have seen that the notion of skyrmion still can survive the quantum limit, although some crucial properties are also lost in the way. Despite the theoretical work, we have also 'seen' (simulated ourselves) the dynamic and stabilization of a skyrmion in the micrometer regime (ELSID). We wanted to also simulate the quantum-skyrmion (with ELQSID) but unfortunately we were unable to make the numerical simulations properly work, so it is currently under development.

During the report, many areas of the physics have been mentioned in the context of skyrmion and more importantly, there are few more. We initially wanted to name this term paper as *Skyrmions: A birth-cake for Physics* in reference of the whirling arrows looks like candles in a birth-cake, but that would be dispersed the aim of the term paper in the context of AQM.

Nevertheless, it is worth mentioning some of other applications of this fascinating structures: in condensed matter [23], scattering problem [24], *continuous Aharonov-Bohm effect* [25], higher dimension particle physics with *Wess-Zumino* additional term [26], superconductivity [27, 28], optics with the new optic-skyrmions as an example of *Full-Poincare-Beams* [29] useful for example of control of nanoparticles, characterizing the *Seebeck and Nernst* [30] coefficients or even explain the abnormal presence of carbon on Earth with the *Hoyle state* of Carbon-12 [12].

Mention apart, in my personal point of view, corresponds to the Micromagnetic simulations of skyrmions as potential carriers of information and storage, which is the final idea of ELSID/ELQSID codes, to develop a stable logic gate, which could easily overcome in advantages the current electronic, as another candidate of spintronics.

I hope this hybrid structure between lecture notes and research paperdon't result too chaotic and the main ideas result 'enough' clear.

[1] M. Born and V. Fock, Zeitschrift fur Physik **51**, 165.
[2] V. G. Ivancevic and T. T. Ivancevic, (2013), arXiv:1305.0613.
[3] H. Lyre, (2014), arXiv:1408.6867.
[4] M.-O. Renou, D. Trillo, M. Weilenmann, L. P. Thinh, A. Tavakoli, N. Gisin, A. Acin, and M. Navascues, (2021), arXiv:2101.10873.
[5] S. Pancharatnam, Proceedings of the Indian Academy of Sciences - Section A **44**, 247 (1956).
[6] M. Born and V. Fock, Proceedings of the Royal Society of London Physical Sciences **392**, 45 (1984).
[7] S. Bhowal, S. Satpathy, and P. Sahu, (2020), arXiv:2005.07635.
[8] L. Kelvin, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **1**, 257 (1876).
[9] C. Adam, C. Naya, K. Oles, T. Romanczukiewicz, J. Sanchez-Guillen, and A. Wereszczynski, (2020), 10.1103/PhysRevD.102.105007, arXiv:2009.05787.
[10] For a clearer visual explanation it is possible to visit https://demonstrations.wolfram.com/SystemOfPendulumsARealizationOfTheSineGordonModel/.
[11] G. Toulouse and M. Kléman, Journal de Physique Lettres **37**, 149 (1976).
[12] F. L. Roux, (2014), 10.1017/etds.2015.56, arXiv:1401.2333.
[13] J. L. Ackrill and E. Schrodinger, The Philosophical Review **64**, 317 (1955).
[14] As Kelvin said, *If you cannot measure it, you cannot improve it. When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind.*
[15] Note that here a smooth density is needed, so we are in a pre-quantic area. In the next example we will try to decrease this limit.
[16] M. I. Trukhanova, Progress of Theoretical and Experimental Physics **2020** (2020), 10.1093/ptep/ptz131.
[17] K. Everschor-Sitte and M. Sitte, Journal of Applied Physics **115**, 172602 (2014).
[18] In particular, the short range of the DMI in B20 materials allows discovering experimentally [23] the skyrmions by serendipity trying to control the helimagnetic regime in MnSi.
[19] T. Jolicoeur and B. Pandey, (2019), 10.1103/PhysRevB.100.115422, arXiv:1907.04612.
[20] S. M. Vlasov, P. F. Bessarab, I. S. Lobanov, M. N. Potkina, V. M. Uzdin, and H. Jónsson, New Journal of Physics **22**, 083013 (2020).
[21] V. Lohani, C. Hickey, J. Masell, and A. Rosch, Physical Review X **9**, 041063 (2019).
[22] O. M. Sotnikov, M. V. V., C. J., and K. M. I., (2018), arXiv:1811.10823.
[23] S. Muhlbauer, B. Binz, F. Jonietz, C. Pfleiderer, A. Rosch, A. Neubauer, R. Georgii, and P. Boni, Science **323**, 915 (2009).
[24] D. Foster and S. Krusch, Nuclear Physics B **897**, 697 (2015).
[25] V. A. Uzunova, (2019), 10.5488/CMP.22.33002, arXiv:1910.00890.
[26] I. J. R. Aitchison, (2019), arXiv:2001.09944.
[27] S. S. Pershoguba, S. Nakosai, and A. V. Balatsky, (2015), 10.1103/PhysRevB.94.064513, arXiv:1511.01842.

[28] K. Pöyhönen, A. Weststrom, S. S. Pershoguba, T. Ojanen, and A. V. Balatsky, (2016), 10.1103/PhysRevB.94.214509, arXiv:1608.08357.
[29] A. M. Beckley, T. G. Brown, and M. A. Alonso, Optics Express **18**, 10777 (2010).
[30] J. Xiao, X. Tan, X. Ye, X. Tang, and D. Ren, Organic Electronics **63**, 1 (2018).
[31] R. V. Verba, D. Navas, S. A. Bunyaev, A. Hierro-Rodriguez, K. Y. Guslienko, B. A. Ivanov, and G. N. Kakazei, Physical Review B **101**, 064429 (2020).
[32] And can also show more complex zoology of the skyrmions, with *Neel*-or *Bloch*- skyrmions, in analogy with the nomenclature for domain walls.

Appendix A: Topological aspects of skyrmions

In Section II, skyrmions were introduced, associated with the concept of *defect* and the genuine difference between a close rope with a knot and without it. Mathematically speaking, we need to establish a way of 'comparing' objects. From basic topology, the concept of *homotopy* corresponds to the solution:

Given two continuous functions $f(x), g(x)$ *defined in their respective topological spaces, both functions are homotopic iff exist a continuous function* $h(x,t)$ *s.t.* $h(x,0) = f(x)$ *and* $h(x,1) = g(x)$.

Intuitively, this corresponds to smoothly (without cut or gluing) change from one to another (it is well-known the smooth transformation of a mug into a torus- or doughnut, which allows some jokes in the popular culture). From here we can classify the infinite (in general) solutions to equations into (equivalence) classes, which are, at least, numerable.

But therefore, how can we classify these solutions? What can be a good representant for these classes? If some characteristics are *invariant* under this transformation, we can directly differentiate between solutions just by comparing these attributes. For example, the number of knots in the closed rope can be an interesting parameter, as it is impossible to transform a rope with 7 knots into one with only 3 (without cutting or gluing).

More interestingly, if we measure different index in two solutions, we can not transform one into the other. Physically this means that, in principle, we can not apply any smooth perturbation (the debate starts with what can be considered 'smooth') to convert one state into the other. So, if we measure in a certain system a topological invariant that is different of the corresponding parameter in the trivial solution, we can not transform smoothly this system into the null-case... it is *stable*! And as it has a topological origin, it is *topologically protected*.

That's exactly the case of the *winding number* from skyrmions (in reality, this is a composite of more parameters, as vorticity, helicity and polarity [31] but in principle it is not needed to split them). This new degrees of freedom causes the break of symmetries and allows to a complete zoology of these defects using the Toulouse-Kleman relation [11].

Appendix B: Energetic

Here the analysis of the statics and dynamics of skyrmions in the magnetic case, which be used in the first part of Section III is presented. We will start by considering the Hamiltonian of a magnetic skyrmion, in which its local variation minimization would lead to the different possible solutions.

$$H = -J\sum \mathbf{S}_i\mathbf{S}_j - K\sum \mathbf{S}_i^z\mathbf{S}_j^z$$
$$-B\sum \mathbf{S}_i^z - \sum_i D\cdot(\mathbf{S}_i\times\mathbf{S}_j) + H_{RKKY}$$

where the summation is extended to neighbours. $\mathbf{S}$ corresponds to the spin, and the other terms correspond to the following: The first term corresponds to *direct exchange* or short-range exchange with J representing the integral of exchange and under symmetry condition can be associated with the divergence of the magnetization. The next term is associated with the *anisotropy*, and it is related with the concept of 'easy axis' having characteristic coefficients K depending on the different symmetries in the material. In third position, we found the magnetic field or *Zeeman* energy. Later on, the *Dzaloshinki-Moriya interation*, which is a *asymmetric exchange* energy that arise due to the spin-orbit interaction in absence of spatial inversion between a ferromagnetic material and a heavy metal and favors non-uniform structures, and where D represents the coupling coefficient and its key for skyrmion stabilization. Finally, the *Ruderman-Kasuya-Kittel-Yosida* interaction or *indirect exchange* that happens between heavy materials with d, f shells with delocalized s ones, which polarized and causes a new coupling. Its mathematical expression is omitted for clarity, but essentially it is important in multi-layer structures with a material separated by thin layers of another heavier species.

The sensitive adjustment between all this coefficients in the energy minimization is a rich experimental landscape and leads to a huge variety of exotic solutions, which are a local minimum

of the Hamiltonian, with higher energy than the ground state but topologically protected. That's the case of B20 chiral structures, like MnSi, were skyrmion were first in laboratory discovered.

Going throught *Brown condition* and equations [12] it is possible to study the (*Larmor*) *precession* (again we have a oscillating pendulum for the spins), associated with the *helicity* [21], topological invariant [32] and whose analytical expression for motion is the following

$$\partial_t \mathbf{m} = \gamma_0 \cdot \mathbf{m} \times \mathbf{H}^{eff} - \alpha_G \left(\mathbf{m} \times \partial_t \mathbf{m}\right) \qquad \text{(B1)}$$

which is known as *Landau-Lifshitz-Gilbert equation*, where γ_0 is the *gyromagnetic* ratio and it is related with the frequency of the movement, while α_G is known as the *Gilbert damping* and causes that the precession slowly stops. It is possible to study this system as a equation of movement of the system (that is the reason why, sometimes, the skyrmion are consider *pseudo-particles* as they are solitons with stable, finite size and moves 'as if' it would be a particle, which is simulated in Appendix C).

And how can move an skyrmion? Basically, due to the spin-torque [17] it is possible to induce an additional term in the LLG equation related with a spin polarized current, 'pushing' the whirling structure and making it 'moving'. There are several kinds of mechanics: from in-plane *spin-transfer-torque* (STT) interaction discovered by Berger [20], to perpendicular control in multilayers avoiding exchange interaction, also called *Slonczewki STT*. This method are specially effective in B20 materials in which the threshold for unpinning the skyrmion structures is several orders of magnitude [17] easier than in other materials.

All this forces causes the complex trajectory and additional non-colinear forces affecting by Berry phase adquisition by the gyrocoupling current which will affect, for example, in Hall signal. The effective control of this interactions is a promising field as several spintronic devices, like *racetrack memories*, requires a full control on the dynamics of the skyrmions.

Finally, it is also interesting to consider thermal influence [16] to see the stability of this processes under the kinetical energy of the system and was crucial for the A-phase appearance in MnSi [23].

Appendix C: Computational simulation

The open-source software OOMMF (*Object Oriented MicroMagnetic Framework*) from ITL/NIST, currently updated to Python language in which we develop the ELSID (*E-stable Logic Skyrmion Information Device*), which code is shown below and its operated in Jupyter Notebook. Underscore symbols are ommited.

```
   import scipy.constants as cte
import numpy as math # Useful libraries
import oommfc as phys
import discretisedfield as comp
import oommf-sk-number as wind
matplotlib inline

KC = 15e-12 # exchange constant
Ms = 5.8e5 # saturation magn
AG = 0.3 # Gilbert damping
aS = AG # STT adiabatically
Px = 1.8e12 # STT factor
gg = 2.211e5 # gyromagnetic term
DD = 3.5e-3 # DMI constant
K1 = 0.8e+6 # PMA constant
uu = (0,0,1) # anistropic direction
```

Initial configuration settings and physical system parameters, extracted from [23].

```
xi = 200e-9 #initial length
xc = 200e-9 #channel length
xf = 200e-9 # final length
xt = xi+xc+xt # total length
yc = 30e-9 #channel width
yt = 150e-9 # total width
zt = 1e-9 # thickness
thdeg = 90 # (deg-rad) phase angle
thrad = math.tan(math.deg2rad(th-deg))
pi = (0,0,0) # initial point skyr
pf = (xt,yt,zt)# final point skyr
cel = (zt,zt,zt)# cell size (FEM)
mesh = phys.Mesh(p1=pi,p2=pf,cell=cel)
sys = phys.System(name="ELSID")
sys.hamiltonian = phys.Exchange(A=KC)
 +phys.DMI(D=DD, crystalclass='Cnv')
 +phys.UniaxialAnisotropy(K1=K1, u=uu)
sys.dynamics = phys.Precession(gamma=gg)
 +phys.Damping(alpha=AG) # dynamics
```

Simulation parameters and construction of the Hamiltonian and LLG equation.

```
   def Ms-none(xyz):  # Abrupt transition
 x, y, z, = xyz
 if (x<xi) and (y<yt/2-yc or y>yt/2+yc:
return 0
 else :  return Ms
def Ms-full(xyz):  # 180deg transition
 x, y, z, = xyz
 if (x<xi) and (y<yt/2-yc or y>yt/2+yc:
return 0
```

```
 if thrad>0
 if (xi<x<xi+xc) and
thrad*(x-xi)<y<(yt/2-yc)
 or (yt/2+yc)<y<(yt-thrad*(x-xi):
return 0
 else:  return Ms
 else:
 if (xi<x<xi+xc) and
(y<(yt/2-yc)+th-rad*(x-xi)
 or (yt/2+yc-thrad*(x-xi))<y<yt:
return 0
 else :  return Ms
def Ms-half(xyz):  # 090deg transition
 x, y, z, = xyz
 if (x<xi) and
(y<yt/2-yc or y>yt/2+yc:  return 0
 elif (xi<x<xi+xc) and
thrad*(x-xi)<y<(yt/2-yc)
 or (yt/2+yc<y<thrad*(x-xi)):  return 0
 else :  return Ms
def Ms-two(xyz):  # double transition
 x, y, z, = xyz
 if (x<xi) and (y<yt/2-yc or y>yt/2+yc:
return 0
 elif (xi<x<xi+xt) and thrad*(x-
xi)<y<(yt/2-yc):  return 0
 else :  return Ms
def init-all(xyz):
 x, y, z, = xyz
 if xi/2<x<xi/2+zt:
return (0.1, 0.1, -1)
 else:  return (0.1, 0.1, +1)
```

Some auxiliary function are introduced to control the shape of the transition to avoid the skyrmion to decay on the transition or to set the initial skyrmionic state, with opposite orientation with respect to the boundaries.

```
   sys.m = cal.Field(mesh,
value=initi-all, norm=Ms-full)
sys.m.z.plot plane(z=0);
minima = phys.MinDriver()
minima.drive(sys, overwrite=True)
sys.m.z.plot plane(z=0);
```

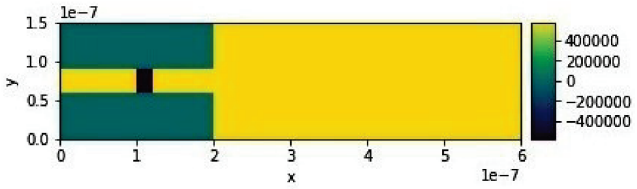


FIG. 6. b.

```
 def speedx(Px):
 return (Px*cte.hbar/2*cte.m-e)
speed = speedx(Px)
sys.dynamic += phys.STT(u=(speed,0,0),beta=aS)
tdis = phys.TimeDriver()
timer = math.ones(9)*zt/10
for ti in timer:
tdis.drive(sys, t=ti, n=10,
overwrite=True)
sys.m.z.plot plane(z=0);
oommf-data = ne.SkNumberOOMMF('ELSID.omf',
z index=0)
print('n3=',
oomf-data.compute sk number())
oommf data.plot charge density
('ELSID-skrmirmion.png'))
```

Heat Vulnerability Index (HVI) mapping in extreme heat days coupled with air pollution using principal component analysis (PCA) technique: A case study of Amiens, France

Aiman Mazhar Qureshi, Ahmed Rachid

***Abstract*—** Extreme heat events are emerging human environmental health concerns in dense urban areas due to anthropogenic activities. High spatial and temporal resolution heat maps are important for urban heat adaptation and mitigation, helping to indicate hotspots that are required for the attention of city planners. The Heat Vulnerability Index (HVI) is the important approach used by decision-makers and urban planners to identify heat-vulnerable communities and areas that require heat stress mitigation strategies. Amiens is a medium-sized French city, where the average temperature is increasing since the year 2000 by +1°C. Extreme heat events are recorded in the month of July for the last three consecutive years 2018, 2019 and 2020. Poor air quality, especially ground-level ozone, has been observed mainly during the same hot period. In this study, we evaluated the HVI in Amiens during extreme heat days recorded last three years (2018,2019,2020). The Principal Component Analysis (PCA) technique is used for fine-scale vulnerability mapping. The main data we considered for this study to develop the HVI model are (a) socio-economic and demographic data; (b) Air pollution; (c) Land use and cover; (d) Elderly heat-illness; (e) socially vulnerable; (f) Remote sensing data (Land surface temperature (LST), mean elevation, NDVI and NDWI). The output maps identified the hot zones through comprehensive GIS analysis. The resultant map shows that high HVI exists in three typical areas: (1) where the population density is quite high and the vegetation cover is small (2) the artificial surfaces (built-in areas) (3) industrial zones that release thermal energy and ground-level ozone while those with low HVI are located in natural landscapes such as rivers and grasslands. The study also illustrates the system theory with a causal diagram after data analysis where anthropogenic activities and air pollution appear in correspondence with extreme heat events in the city. Our suggested index can be a useful tool to guide urban planners and municipalities, decision-makers and public health professionals in targeting areas at high risk of extreme heat and air pollution for future interventions adaptation and mitigation measures.

***Keywords*—**Heat vulnerability index, Heat mapping, Heat health-illness, Remote sensing, Urban heat mitigation.

Aiman Mazhar Qureshi is with the University of Picardie Jules Verne, Amiens, CO 80000 France (corresponding author, phone: +33601535460; mail: aiman.mazhar.qureshi@etud.u-picardie.fr).

Ahmed Rachid is with the University of Picardie Jules Verne, Amiens, CO 80000 France. (e-mail: ahmed.rachid@u-picardie.fr).

Analysing Competitive Advantage of IoT and Data Analytics in Smart City Context

Petra Hofmann, Dana Koniel, Jussi Luukkanen, Walter Nieminen, Lea Hannola, Ilkka Donoghue

***Abstract*—**The Covid-19 pandemic forced people to isolate and become physically less connected. The pandemic has not only reshaped people's behaviours and needs but also accelerated digital transformation (DT). DT of cities has become an imperative with the outlook of converting them into smart cities in the future. Embedding digital infrastructure and smart city initiatives as part of normal design, construction, and operation of cities provides a unique opportunity to improve connection between people.

Internet of Things (IoT) is an emerging technology and one of the drivers in DT. It has disrupted many industries by introducing different services and business models, and IoT solutions are being applied in multiple fields, including smart cities. As IoT and data are fundamentally linked together, IoT solutions can only create value if the data generated by the IoT devices is analysed properly. Extracting relevant conclusions and actionable insights by using established techniques, data analytics, contributes significantly to the growth and success of IoT applications and investments. Companies must grasp DT and be prepared to redesign their offerings and business models to remain competitive in today's marketplace.

As there are many IoT solutions available today, the amount of data is tremendous. The challenge for companies is to understand what solutions to focus on and how to prioritise and which data to differentiate from competition. This paper explains how IoT and data analytics can impact competitive advantage and how companies should approach IoT and data analytics to translate them into concrete offerings and solutions in the smart city context. The study was carried out as a qualitative, literature-based research. A case study is provided to validate the preservation of company's competitive advantage through smart city solutions. The results of the research contribution provide insights into the different factors and considerations related to creating competitive advantage through IoT and data analytics deployment in the smart city context. Furthermore, this paper proposes a framework that merges the factors and considerations with examples of offerings and solutions in smart cities.

The data collected through IoT devices, and the intelligent use of it, can create competitive advantage to companies operating in smart city business. Companies should take into consideration the five forces of competition that shape industries and pay attention to the technological, organisational, and external contexts which define factors for consideration of competitive advantages in the field of IoT and data analytics. Companies that can utilise these key assets in their businesses will most likely conquer the markets and have a strong foothold in the smart city business.

***Keywords*—**Internet of Things, Data Analytics, Smart Cities, Competitive Advantage

Petra Hofmann is a Master of Science student at the LUT University in the Global Management of Innovation and Technology Programme, 53851 Lappeenranta, Finland (phone: +358 40 142 9700; e-mail: petra.melander@student.lut.fi).

Dana Koniel is a Master of Science student at the LUT University in the Global Management of Innovation and Technology Programme, 53851 Lappeenranta, Finland (e-mail: dana.koniel@student.lut.fi).

Jussi Luukkanen is a Master of Science student at the LUT University in the Global Management of Innovation and Technology Programme, 53851 Lappeenranta, Finland (e-mail: jussi.luukkanen@student.lut.fi).

Walter Nieminen is a Master of Science student at the LUT University in the Global Management of Innovation and Technology Programme, 53851 Lappeenranta, Finland (e-mail: walter.nieminen@student.lut.fi).

Lea Hannola is an Associate Professor at the LUT University in the Department of Industrial Engineering and Management, 53851 Lappeenranta, Finland (e-mail: lea.hannola@lut.fi)

Ilkka Donoghue is a post-doctoral researcher at LUT University in the Department of Industrial Engineering and Management, 53851 Lappeenranta, Finland (e-mail: ilkka.donoghue@lut.fi)

Challenges for a WPT 4 Waiting Lane Concept - Laboratory and Practical Experience

Julia Langen

***Abstract*—**This article describes the challenges of a wireless charging system for a cab waiting lane in a public space and presents a concept for solving them.

In this concept multiple cabs can be charged simultaneously and during stopping and rolling. Particular technical challenges are a coil topology that meets the EMF requirements and an intelligent control concept that allows the individual coil segments to be switched on and off.

The charging concept explained here is currently being implemented as a pilot project, so that initial results on the operation can be presented.

.

***Keywords*—**Charge lane, inductive charging solution, smart city, wireless power transfer.

J. Langen, M. Eng., is Head of Communication Development at INTIS Integrated Infrastructure Solutions GmbH, Lathen, Germany (phone: 05933-6245; fax: 05933-6220; e-mail: julia.langen@intis.de).

The Impact of an Ionic Liquid on Hydrogen Generation from a Redox Process Involving Magnesium and Acidic Oilfield Water

Mohamed A. Deyab[1,*], Ahmed E. Awadallah[1]

[1]*Egyptian Petroleum Research Institute (EPRI), Nasr City, Cairo, Egypt*

*Corresponding Author E-mail: hamadadeiab@yahoo.com

Abstract— Under various conditions, we present a promising method for producing pure hydrogen energy from the electrochemical reaction of Mg metal in waste oilfield water (WOW). Mg metal and WOW are primarily consumed in this process. The results show that the hydrogen gas output is highly dependent on temperature and solution pH. The best conditions for hydrogen production were found to be a low pH (2.5) and a high temperature (338 K). For the first time, the Allyl methylimidazolium bis-trifluoromethylsulfonyl imide) (IL) ionic liquid is used to regulate the rate of hydrogen generation. It has been confirmed that increasing the solution temperature and decreasing the solution pH accelerates Mg dissolution and produces more hydrogen per unit of time. The adsorption of IL on the active sites of the Mg surface is unrestricted by mixing physical and chemical orientation. Inspections using scanning electron microscopy (SEM), energy dispersive X-ray (EDX), and FT-IR spectroscopy were used to identify and characterise surface corrosion of Mg in WOW. This process is also completely safe and can create energy on demand.

Keywords— Hydrogen production; Mg; Waste water; ionic liquid

Hydrodynamic and Sediment Transport Analysis of Computational Fluid Dynamics Designed Flow Regulating Liner (Smart Ditch)

Saman Mostafazadeh-Fard, Zohrab Samani, Kenneth Suazo

***Abstract*—** Agricultural ditch liners are used to prevent soil erosion and reduce seepage losses. This paper introduced an approach to validate a computational fluid dynamics (CFD) platform FLOW-3D code and its use to design a flow regulating corrugated agricultural ditch liner system (Smart Ditch (SM)). Hydrodynamic and sediment transport analysis were performed on the proposed liner flow using CFD platform FLOW-3D code. The code's hydrodynamic and scour and sediment transport models were calibrated and validated using lab data with an accuracy of 94 % and 95%, respectively. The code was then used to measure hydrodynamic parameters of sublayer turbulent intensity, kinetic energy, dissipation and packed sediment mass normalized with respect to sublayer flow velocity. Sublayer turbulent intensity, kinetic energy, and dissipation in the SM flow was significantly higher than CR flow. An alternative corrugated liner was also designed and sediment transport was measured and compared to SM and CR flows. Normalized packed sediment mass with respect to average sublayer flow velocity was 27.8 % lower in alternative flow compared to SM flow. CFD platform FLOW-3D code could effectively be used to design corrugated ditch liner systems and perform hydrodynamic and sediment transport analysis under various corrugation designs.

***Keywords*—** CFD, hydrodynamic, sediment transport, ditch, liner design.

Saman Mostafazadeh Fard is with the New Mexico State University, United States (e-mail: samanmzf@nmsu.edu).

Masonry Arches Under Spreading Supports: Size Effect Law Including Self-Weight

M. Micaela, M. Pathirage, A. Gregori, G. Cusatis

***Abstract*—**This study is about the fracturing behavior of unreinforced masonry arches and vaults under spreading supports. Vaulted structures are traditionally analyzed throughlimit analysis method, that is limited in understanding the actual failure of these structures because it assumes the simultaneous formation of hinges once the thrust line reaches the edge of the masonry structure and itcompletely ignores the fracturing propagation phenomenon. Moreover, limit analysis does not capture the effect of structural size on the nominal strength due to strain-softening and damage localization. This manuscript proposes a thorough understanding of the fracturing behavior and size-effect of arches and vaults based on computational modeling and non-linear fracture mechanics concepts. The numerical data of size-effect are also analyzed using a newly developed analytical formula based on non-linear fracture mechanics theory that accounts for the self-weight. Results show a strong reduction of structural strength as the arch size increases, that is much more pronounced than the typically observed reduction due to energetical size-effect. The difference can be explained by the presence of self-weight, one of the main driving forces in the fracture and collapse of thrusting vaulted structures. The results of this study can explain the reason why in some seismic locations, small sized masonry vaulted structures remain almost undamaged under seismic excitation whereas larger ones often collapse.

***Keywords*—**Fracture mechanics, Masonry Vaulted Structures, Self-Weight, Size-effect.

M. M. Author is with University of L'Aquila, L'Aquila, AQ67100ITALY (e-mail: micaela.mercuri@graduate.univaq.it).
M. P. Author is with Northwestern University, Evanston, IL 60208-3109.
A. G. Author is with University of L'Aquila, L'Aquila, AQ 67100 ITALY.
G. C. Author is with Northwestern University, Evanston, IL 60208-3109.

Cosmic Muon Tomography At The Wylfa Reactor Site Using An Anti-Neutrino Detector

Ronald Collins, Jonathon Coleman, Joel Dasari, George Holt, Carl Metelko, Matthew Murdoch, Alexander Morgan, Yan-Jie Schnellbach, Robert Mills, Gareth Edwards, Alexander Roberts

Abstract—At the Wylfa Magnox Power Plant between 2014 – 2016 the VIDARR prototype anti-neutrino detector was deployed. It is comprised of extruded plastic scintillating bars measuring $4\,\text{cm} \times 1\,\text{cm} \times 152\,\text{cm}$ and utilised wavelength shifting fibres (WLS) and Multi-Pixel Photon Counters (MPPCs) to detect and quantify radiation. During deployment it took cosmic muon data in accidental coincidence with the anti-neutrino measurements with the power plant site buildings obscuring the muon sky. Cosmic muons have a significantly higher probability to be attenuated and/or absorbed by denser objects and so one-sided cosmic muon tomography was utilised to image the reactor site buildings. In order to achieve clear building outlines a control data set was taken at the University of Liverpool from 2016 – 2018 which had minimal occlusion of the cosmic muon flux by dense objects. By taking the ratio of these two data sets and using GEANT4 simulations it is possible to perform one-sided cosmic muon tomography analysis. This analysis can be used to discern specific buildings, building heights, and features at the Wylfa reactor site including the reactor core/reactor core shielding using ~ 3 hours worth of cosmic-ray detector live time. This result demonstrates the feasibility of using cosmic muon analysis to determine a segmented detector's location with respect to surrounding buildings, assisted by aerial photography or satellite imagery.

Keywords—Anti-neutrino, GEANT4, muon, tomography, occlusion.

I. Introduction

SINCE 2007 and the deployment of the SONGS1 prototype [1] there has been a renewed interest in the use of anti-neutrinos for reactor monitoring. A detector that meets all desirable traits for this application [2] has been developed at the University Of Liverpool, the Verification Instrument for the Direct Assay of Radiation at Range (VIDARR). A prototype of this detector was deployed at Wylfa on the northwest coast of Wales and took anti-neutrino data [3]. Whilst this detector was taking anti-neutrino data cosmic μ data was taken in accidental coincidence. This cosmic μ data should contain tomographic information of the Wylfa site and as such an investigation was launched to determine if this information could be extracted.

II. A Brief Overview of the Detector

The tonne scale anti-neutrino detector deployed at the Wylfa reactor site was $\sim 60\,\text{m}$ from the $1.5\,\text{GW}_{th}$ reactor [4]. It was a prototype of the VIDARR detector based on the ND280 electromagnetic calorimeter from T2K [3]. Using segmented extruded plastic scintillator and Hamamatsu Multi-Pixel Photon Counters (MPPCs), the detector had 49 layers of plastic scintillator which alternated being at 90° from the previous layer to form a hodoscope. The active detector region uses gadolinium sheets in-between the layers of plastic in order in order to enhance the efficiency of the inverse β decay channel by increasing the neutron capture cross-section. Muons interact with the detector as minimally ionising particles (MIPs) as determined by the Bethe-Bloch formula [5]. Muons were primarily used as a calibration tool at the beginning of the deployment [3]. This technology has also been used in cosmic μ telescopes such as the MU-RAY experiment which uses triangular prism segments [6] and the DIAPHANE experiment which uses cuboid segments [7]. During the main deployment at Wylfa cosmic muons were mostly rejected via a cosmic muon veto in order to reduce background when measuring anti-neutrinos. As a result the amount of information available is significantly less than the 20 month period of deployment at Wylfa would suggest. Instead the number of cosmic muon events is equivalent to ~ 3 hours worth of live time.

R. Collins (email R.P.Collins@Liverpool.ac.uk) is with the Department of Physics, University of Liverpool, as are J. Coleman, J. Dasari, G. Holt, C. Metelko, M. Murdoch, A Morgan and Y.Schnellbach.

R.W Mills is with the National Nuclear Lab

G. Edwards, A. Roberts are with the Wylfa Power Station

III. One-Sided Cosmic μ Tomography

Cosmic μ tomography has two distinct types: one-sided and two-sided. Two-sided tomography requires two detectors with an area much larger than the object attempting to be measured in order to reconstruct vertices from μ scattering [8]. This approach can be used to discern dense objects, such as nuclear waste [9] [10] and reactors [11] as well as ascertain damage to structures [12]. However, the size of the reactor site buildings is far greater than the VIDARR prototype and there was only one unit deployed. As a result one-sided cosmic μ tomography will be used instead, a top down example of which can be seen in figure 1 which shows how a hypothetical cuboid shaped building would look via spherical projection in ϕ to a detector. The curve shape in the occluded angles in figure 1 are due to the amount of material obscuring the cosmic μ increasing and decreasing depending on the angle. The same information can also be represented using both ϕ and θ which is shown in figure 2, the result resembles a square with a parabola on top. This is due to two main factors: a wide field of view and spherical projection onto a single point. The approach of measuring the difference in intensity or "transmission" has also been done by the MU-RAY collaboration [13] and is the method that will continue to be used in this paper. In addition, the μ radiography approach may not counter act the issues of using a segmented detector to analyse the entire cosmic

hemisphere distribution at once. Bin migration is a significant effect in the data therefore taking the ratio of expected and measured distributions may be the only way to counteract bin migration of this scale. Further credence to this is the requirement of a control data set when using one-sided cosmic μ tomography to discern reactor damage in simulation [14].

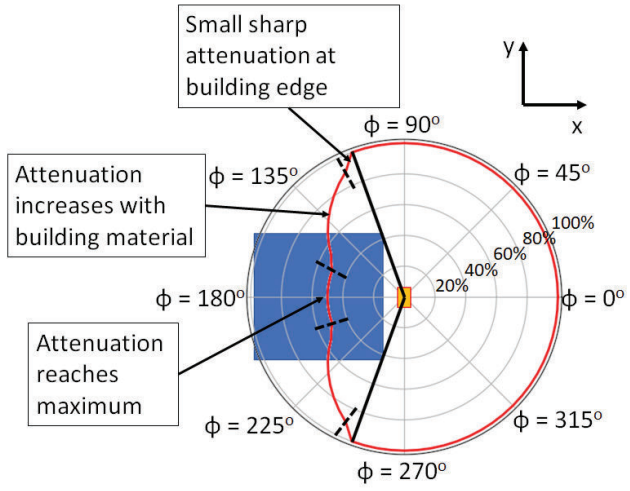


Fig. 1. A top down theoretical example of what one-sided cosmic μ tomography with the red line indicating the received μ flux in the detector with respect to ϕ. The more material that cosmic μ have to pass through the lower chance they will be detected via attenuation. Assuming a uniform density.

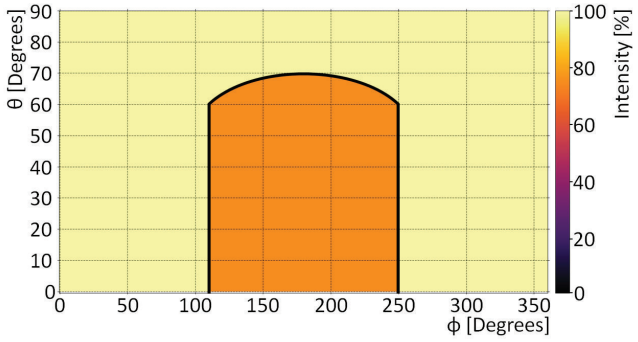


Fig. 2. A side on view of figure 1 showing how ϕ and θ project on to a detector and what outline would be expected from a cuboid of material that stops a significant amount of cosmic μ.

IV. Wylfa Reactor Site

In July 2014 the VIDARR prototype detector started taking data at the Wylfa reactor site. A cut-away of the Wylfa reactor building can be seen in figure 3, showing that the structure of the main reactor building is significantly denser in the bottom half than in the top half. The upper storage section therefore will obscure fewer cosmic μ than bottom section with the reactor. Due to the top half of the reactor buildings being mostly empty, the edges of the buildings are expected to show as indistinct, blurred drops in intensity as opposed to sharp edges to a detector analysing cosmic μ intensity. As a result, measuring the main reactor building height is more difficult than measuring the height of the closer service buildings. The position of the buildings relative to the detector, building widths, and building depths can all be located via aerial photography sourced from Google Maps (see figure 4). There are many buildings that will occlude the sky as viewed by the detector. This combined with the variable density of the site buildings means that the overlapping occlusion patterns or ”shadows” will result in a complex distribution of cosmic-ray directions recorded by the detector.

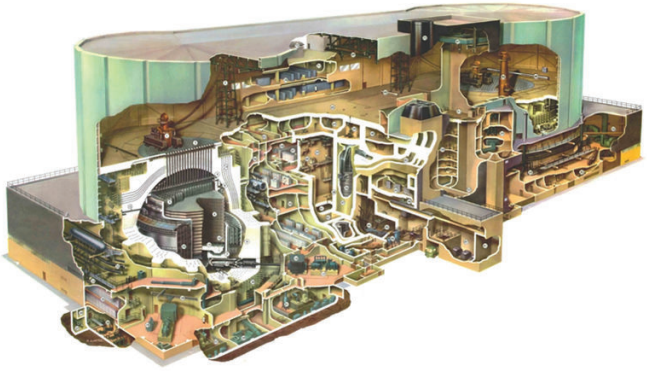


Fig. 3. A cut-away diagram from [15]. Shows the internal structure of the Wylfa from the back of the reactor buildings the bottom half of the reactor buildings have thick concrete walls whilst the upper section is mostly storage. As this is an qualitative illustration of the building for public communication, the exact details, locations of features and proportions may not match the exact physical structure.

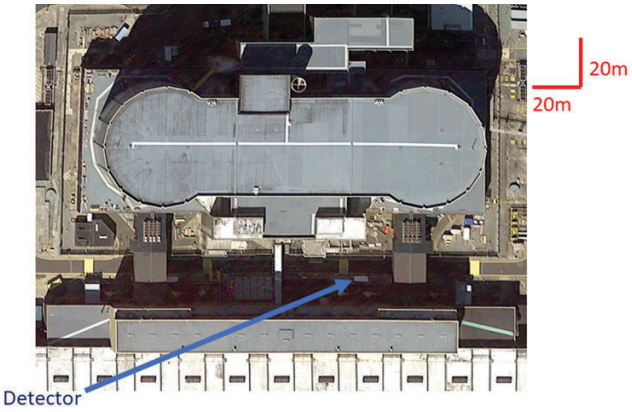


Fig. 4. Google Maps aerial photography image, displaying the power plant site as well as the approximate detector position. The detector is in the middle of many site buildings occluding the incident cosmic-rays. Key buildings shown are the reactor building, the turbine hall and the steam pipe connections between the two [16].

V. GEANT4 Geometry

In order understand the combined effect of the buildings at the reactor site GEANT4 [17] was used to simulate each

building individually. The overall setup for the buildings in GEANT4 can be seen in figure 5. Approximations of the shape of the buildings in the x-y plane in GEANT4 are estimated from the Google Maps aerial photography seen in figure 4. The building heights have to be approximated from the data measured at Wylfa, this is then compared and adjusted until the heights of the simulated shadows match the heights of the measured shadows at Wylfa. The building heights range from 25 m to 50 m and are shown in figure 5, with the steam bridges being 10 m off the ground and 5 m thick in the z axis. The reactor cores are cylinders with widths estimated from private communications to be $\sim$ 25 m (including core shielding). The height of 27.5 m is inferred from measured data.

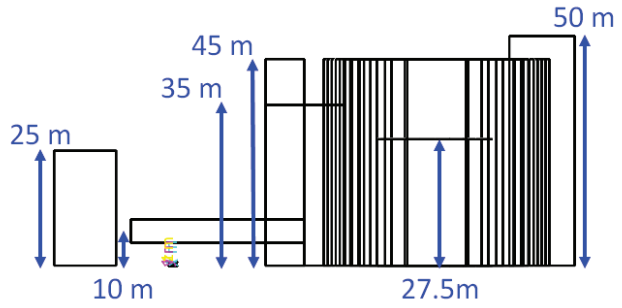


Fig. 5. A Side on X Z view of the GEANT4 geometry for simulated buildings which are 100 % concrete. The origin of the coordinate system corresponds to the detector location.

VI. Reconstruction of Cosmic μ Tracks

The cosmic μ tracks were reconstructed with a custom built tracker. The tracker has two main purposes to run as a continuous online track fitter to allow for constant calibration of the detector and to perform a full reconstruction of cosmic tracks. In this analysis, the reconstruction is optimised for the angle-of-incidence extraction as key analysis variable. Each event in the detector will contain a series of energy deposits in a single bar, each one of these deposits is referred to as a ”hit”. The tracker must therefore be able to determine if a cosmic track is present and if it is which hits are part of the cosmic μ track. The reconstruction steps for fitting each individual track is as follows:

1) **Remove Small Events:** If $<$ 8 are hit above a threshold of 0.693 MeV discard that event (efficiency $\sim$ 99 % for simulation, efficiency $\sim$ 99 % for measured)
2) **Energy Threshold:** Exclude all hits $<$ 0.345 MeV
3) **Track Width:** Exclude hits that are 4 bars away from any other hit
4) **Hit Threshold:** Exclude events that have $<$ 4 bars per side that are above a 0.69 MeV threshold (efficiency $\sim$ 98 % for simulation, efficiency $\sim$ 95 % for measured)
5) **Basic Fit:** Estimate track gradient and intercept using top and bottom hit of the event
6) **First Fit:** First fit of the track (using the GNU simplex algorithm [18])
7) **Track Width 2:** Exclude hits that are 4.5 bars away from the track
8) **Second Fit:** second fit of the track (again using the GNU simplex algorithm [18])
9) **Track Width 3:** Exclude any hits that are 0.5 bars away from the track
10) **Third Fit:** Third and final fit of the track (again using the GNU simplex algorithm [18])
11) **Exclude Bad Events:** Events that have ¡ 50 % of the total energy as signal energy are removed (efficiency $\sim$ 95 % for simulation, efficiency $\sim$ 65 % for measured)
12) **Exclude Empty Events:** Any Events that have no signal energy are removed (used if other cuts are disabled)

VII. Ratios of Measured and Simulated Data Sets

The ratio of the Wylfa data set and the background measurement is taken to determine the buildings shadows on a bin-by-bin basis. Figure 6 shows the distortion inherent to analysing the cosmic μ hemisphere with a cuboid segmented detector. This distortion has 3 main components: Firstly, the cosmic μ distribution is a hemi-spherical distribution which is being analysed by a cuboid shaped detector. Secondly, the VIDARR prototype segments were 4 cm $\times$ 1 cm $\times$ 152 cm long, resulting in a limited angular resolution causing subsequent bin migration. This bin migration is most clearly seen in bins $\phi = 0°$, $\phi = 90°$, $\phi = 180°$, $\phi = 270°$ in figure 6. Thirdly, there is significant scattering as the cosmic μ interact with the atmosphere and buildings which causes blurred edges. Therefore, the deficit in cosmic-ray flux is less pronounced than naively expected for such large and dense buildings, as shown in figure 6. In figure 6 the shadows in the Wylfa data are faint even when comparing side by side with the free sky data set taken at Liverpool.

According to the CRY library [19] 162000 μ^- and 174000 μ^+ are produced in 2820 s for 1 m^2. VIDARR's top surface area is 1.52 m $\times$ 1.52 m = 2.31 m^2 resulting in $\sim$ 275 muons s^{-1} the Wylfa data contains $\sim$ 3 $\times$ 10^6 events and so represents only $\sim$ 3 hours worth of live time. At the Wylfa reactor site there are several buildings and features that are expected to produce shadows: the Main reactor building, the two service buildings at the front, the two service towers, the stream bridges connecting the reactor to the turbine hall, the turbine hall itself (behind the detector), and the reactor cores. By simulating each building/feature in GEANT4 it is possible to obtain an outline of each of the buildings as it would be visible to the detector. Then the combined effect of all relevant buildings at the site can be simulated taking the ratio both with and without the buildings highlighting the outline in ϕ and θ as it appears to the detector. This can then be compared to the ratio of the Wylfa and Liverpool data set is (whilst normalising the Liverpool distribution accordingly). Both of these 360° panoramics are produced in figure 7, the expected simulated distribution matches the measured distribution very well. This suggests that the measured shadows shown in figure 7 are likely an accurate representation of the Wylfa reactor site. Data below 17.5° in θ in figure 7 is excluded as shadows in the Liverpool data become prominent and the statistics fall sharply. The simulated data below 17.5° in θ in figure 7 is also removed so the comparison is 1:1.

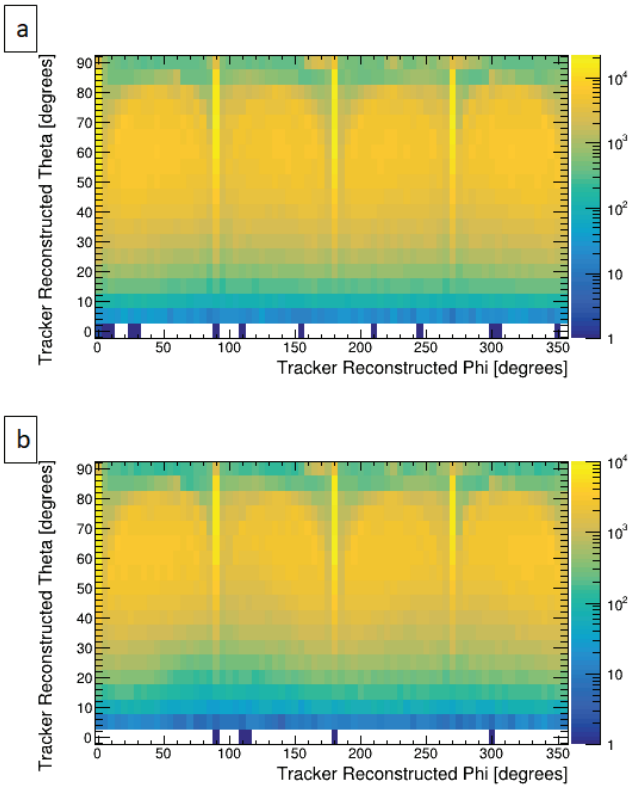


Fig. 6. (a) The azimuthal angle (ϕ) and the polar angle (θ) for the Liverpool cosmic muon distribution. (b) ϕ and θ for the Wylfa cosmic muon distribution. Both distributions are warped considerably due to the shape and segmentation of the detector. The Liverpool data have shadows below 20° θ. The Wylfa data have shadows that extend from 0° θ – 65° θ.

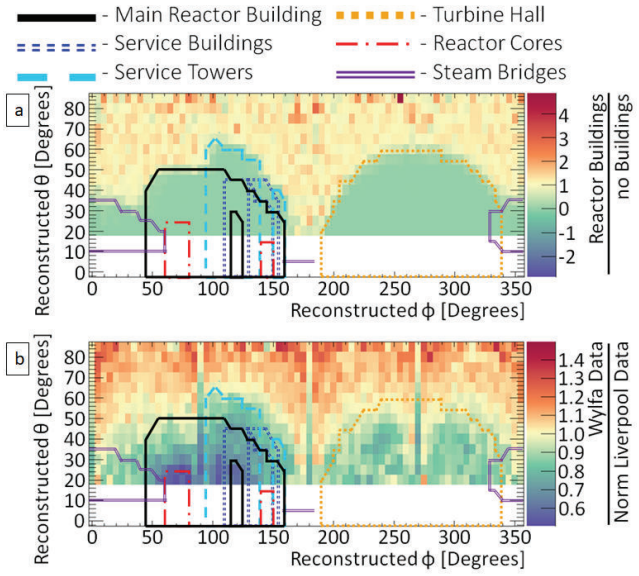


Fig. 7. (a) Ratio of simulated GEANT4 data with and without the shadow. The shadows are shown in green (ratio $<$ 1). In order to determine which components of the shadow are caused by which building each building was simulated on its own and outlined. (b) Ratio of the Wylfa and Liverpool data shadows are seen with a values $\sim$ $<=$ 1 coloured yellow and green. The components of the buildings and features match well as shown by the outlines.

VIII. Results

Whilst the composite shadow shape from simulation matches the measured results well the feature of most interest is the near reactor core / reactor core shielding which is shown by the red rectangle in figure 7. The reactor core / reactor core shielding shadow can be seen around the core at lower angles of θ = 20° and θ = 25° highlighting the core's position effectively. With the other buildings accounted for in the shadow it is reasonable to assume that this feature is the reactor core/reactor core shielding. These features have been mapped out as a top down projection with the main reactor building, reactor core / reactor core shielding, and turbine hall being shown in figure 8 with the near side of each building being clearer than the far side, which is due to scattering of cosmic μ increasing with distance.

IX. Application to Reactor Monitoring

Anti-neutrinos cannot be obscured by any material. Due to their small cross section of the order 10^{-42} cm^2 [20]. As a result one possible approach to manipulate the measured anti-neutrino flux is to physically move the detector. For example moving the detector twice as close would correlate to an anti-neutrino rate four times higher than if the power generation were to remain constant due to the inverse square

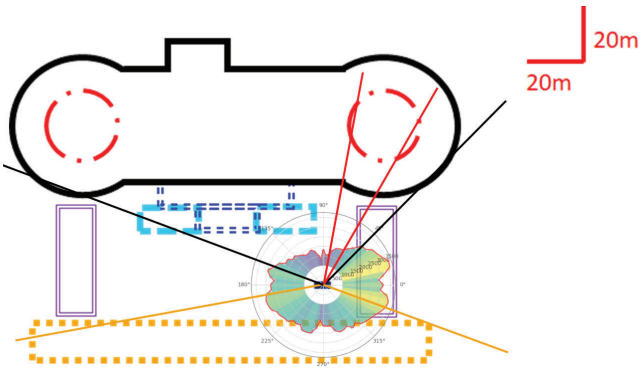


Fig. 8. Trace of the Wylfa reactor site (figure 4) showing the key buildings that cast muon shadows onto the detector with an overlay of measured data from the Wylfa reactor site highlighting the main reactor building turbine hall and reactor core area. The key is the same as figure 7.

law. However, it would be possible for an unscrupulous reactor site to reduce the observable reactor power in the anti-neutrino flux by moving the detector away from the reactor. However, reactors do not reach peak power output instantly. It took $\sim$ 2 weeks for the Wylfa reactor to reach peak power output [3]. To reduce power output, move the detector accordingly, extract illicit material, and return the power to peak power output in ¡ 1 week should be infeasible for most reactor sites. Therefore, running the detector in cosmic mode for 1 hour a week should prove unobtrusive to neutrino measurement of the core whilst allowing imaging of the surrounding site to perform an in-detector position survey of the detector relative to its surroundings. Thus enhancing the capability of the detector as an instrument aiding in non-proliferation.

X. Conclusion

The VIDARR prototype has observed a deficit in cosmic-ray flux or "shadows" at the Wylfa reactor site that correspond to building positions extracted from aerial photography. Then the building positions were validated using simulations and the height as well as sub-features of the buildings, including the shadow of the near reactor core were determined. Including the shadow of the near reactor core/reactor core shielding seen in figure 7. In addition this is possible with only ~ 3 hours worth of data at the Wylfa reactor site. Moving forward the VIDARR detector will be able to switch into cosmic mode for 1 hour a week when deployed which should be sufficient to determine the position of the detector whilst not being detrimental to the reactor monitoring process. Higher cumulative live time will also allow more detailed imaging of the surroundings, allowing for higher position accuracy as well as imaging features of interest. In future more detailed positional data and on-site building measurements will also be utilised to constrain tomographic measurements more effectively.

Acknowledgements

With thanks to the Science and Technology Facilities Council (STFC) for their support
And to the Royal Society for their support
And to Magnox for letting the collaboration on-site
And to NNL for helping us understand the reactor site

References

[1] N. Bowden, A. Bernstein, M. Allen, J. Brennan, M. Cunningham, J. Estrada, C. Greaves, C. Hagmann, J. Lund, W. Mengesha, T. Weinbeck, and C. Winant, "Experimental results from an antineutrino detector for cooperative monitoring of nuclear reactors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 572, no. 2, pp. 985 – 998, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900206024326

[2] I. Headquarters, "Final report: Focused workshop on antineutrino detection for safeguards applications," October 2008. [Online]. Available: www.lefigaro.fr/assets/pdf/AIEA-neutrino.pdf

[3] J. Carrol *et al.*, "Monitoring reactor anti-neutrinos using a plastic scintillator detector in a mobile laboratory," *Pre-print*, Nov 2018. [Online]. Available: https://arxiv.org/abs/1811.01006

[4] S. Jensen and E. Nonbøl, "Description of the magnox type of gas cooled reactor (magnox)," 1999.

[5] P. Zyla *et al.*, "Review of Particle Physics," *PTEP*, vol. 2020, no. 8, p. 083C01, 2020.

[6] A. Anastasio *et al.*, "The mu-ray detector for muon radiography of volcanoes," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 732, pp. 423–426, 2013, vienna Conference on Instrumentation 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900213007705

[7] J. Marteau, J. d. B. d' Ars, D. Gibert, K. Jourde, J.-C. Ianigro, and B. Carlus, "Diaphane: muon tomography applied to volcanoes, civil engineering, archaelogy," *Journal of Instrumentation*, vol. 12, no. 02, p. C02008–C02008, Feb 2017. [Online]. Available: http://dx.doi.org/10.1088/1748-0221/12/02/C02008

[8] L. J. Schultz, G. S. Blanpied, K. N. Borozdin, A. M. Fraser, N. W. Hengartner, A. V. Klimenko, C. L. Morris, C. Orum, and M. J. Sossong, "Statistical reconstruction for cosmic ray muon tomography," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 1985–1993, 2007.

[9] G. Jonkmans, V. Anghel, C. Jewett, and M. Thompson, "Nuclear waste imaging and spent fuel verification by muon tomography," *Annals of Nuclear Energy*, vol. 53, p. 267–273, Mar 2013. [Online]. Available: http://dx.doi.org/10.1016/j.anucene.2012.09.011

[10] Y. Guangliang, A. Clarkson, D. Hamilton, M. Hoek, D. Ireland, R. Johnstone, R. Kaiser, T. Keri, S. Lumsden, D. Mahon *et al.*, "Application of muon tomography to encapsulated nuclear waste," in *2015 12th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, vol. 3. IEEE, 2015, pp. 1359–1362.

[11] H. Miyadera, K. N. Borozdin, S. J. Greene, Z. Lukić, K. Masuda, E. C. Milner, C. L. Morris, and J. O. Perry, "Imaging fukushima daiichi reactors with muons," *Aip Advances*, vol. 3, no. 5, p. 052133, 2013.

[12] E. Niederleithinger, S. Gardner, T. Kind, R. Kaiser, M. Grunwald, G. Yang, B. Redmer, A. Waske, F. Mielentz, U. Effner *et al.*, "Muon tomography of the interior of a reinforced concrete block: First experimental proof of concept," *Journal of Nondestructive Evaluation*, vol. 40, no. 3, pp. 1–14, 2021.

[13] F. Ambrosino *et al.*, "The MU-RAY project: detector technology and first data from mt. vesuvius," *Journal of Instrumentation*, vol. 9, no. 02, pp. C02 029–C02 029, feb 2014. [Online]. Available: https://doi.org/10.1088/1748-0221/9/02/c02029

[14] A. Erlandson, K. Boniface *et al.*, "One-sided muon tomography—a portable method for imaging critical infrastructure with a single muon detector," *CNL Nuclear Review*, vol. 7, no. 1, pp. 1–9, 2018. [Online]. Available: https://doi.org/10.12943/CNR.2016.00014

[15] N. E. I. Magazine, "Wylfa magnox, wylfa, anglesey, uk. wall chart insert," *Nuclear Engineering International Magazine*, vol. 39, 1965.

[16] "Google Maps wylfa aerial image," https://www.google.com/maps/place/Wylfa+Power+Station+Reactor+No.1/@53.4169844,-4.4836539,396m/data=!3m1!1e3!4m5!3m4!1s0x48644f5f73d2cfeb:0x3c76316391d4330!8m2!3d53.4174802!4d-4.482983, longitude = 53.4169844, latitude = -4.4836539, Google given access: 2018 / 06 / 28, accessed: 2021 / 08 / 02.

[17] S. Agostinelli *et al.*, "GEANT4–a simulation toolkit," *Nucl. Instrum. Meth. A*, vol. 506, pp. 250–303, 2003.

[18] M. Galassi, J. Davies *et al.*, *GNU scientific library*. Network Theory Limited, 2002.

[19] C. Hagmann, D. Lange, and D. Wright, "Cosmic-ray shower generator (cry) for monte carlo transport codes," in *2007 IEEE Nuclear Science Symposium Conference Record*, vol. 2, 2007, pp. 1143–1146.

[20] P. Vogel and J. F. Beacom, "Angular distribution of neutron inverse beta decay, $\bar{\nu_e} + p \rightarrow e^+ + n$," *Physical Review D*, vol. 60, no. 5, Jul 1999. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.60.053003

Health Communication: Solution or Confusion in Determining Health Information During Pandemic in 2019-2022

K. Y.S. Putri, kinkinsubarsa@unj.ac.id
Neneng Siti Silfi Ambarwati nenengsitisilfi@unj.ac.id
Ifan Iskandar, ifaniskandar@unj.ac.id
S Bekti Istiyanto, bekti.istiyanto@unsoed.ac.id
Yasir Riady, yasir@ecampus.ut.ac.id
Khalda Noer Maidah, KhaldaNoerMaidah_1410619021@mhs.unj.ac.id
Harisa Ninda Siti Ramadhany, HarisaNindaSitiRamadhany_1410619043@mhs.unj.ac.id

Abstract— Society needs health information during this two-year pandemic. However, people get hoaxes and misleading information in the news and mainstream media. The research question is whether health communication is a solution or confusion in determining health information in the 2019-2022 pandemic. This research aims to find health communication as a solution or confusion in determining health information in the 2019-2022 pandemic. The literature review used in this research is the concepts of health communication, hoax information, and consumer attitudes. Several previous studies expressed support for this research. Quantitative research method by distributing questionnaires to respondents in several areas in Indonesia, such as Ambon, Makassar, Bangka, Bandung, and Bekasi. The results of the validity and reliability of the research variables are valid and reliable. The descriptive test of the two variables is positive. Health communication is positive in all its dimensions. And the attitude is positive in all dimensions in it. So it can be continued to the next analysis. The results of linear regression of research variables influence these two variables. And health information helps provide information to society through social media is a solution to helping their health. At the beginning of the pandemic, people did not know how to identify hoaxes or facts. Over time, through research assistance, society can identify hoaxes or facts. Suggestions from research results, society must be accompanied continuously by the latest conditions in developments around them.

Keywords— Hoax, Health Information, Health Communication, Significant Word.

I. INTRODUCTION

Health Communication is a way used to distribute information, and influence and motivate individuals and even institutions in the field of health. Health Communication refers to the interaction between health and individual behavior. When used correctly, health communication can affect people's attitudes. The main objective of Health communication is to achieve health improvement in practice and activity status.[1]

Since the Covid-19 pandemic, it has had a bad impact on all aspects of activities in Indonesia. Statistical data were collected from the official website of Covid-19 Indonesia, from March 15, 2020, to May 15, 2022, the total number of Covid-19 cases in Indonesia was 6,050,776 with 5,889,534 people recovering and 156,458 people dying. [2]

The scale of the crisis and the government's response has been offset by the huge flow of information regarding Covid-19 in news coverage, television press conferences, and others. People need health information during this two-year pandemic. But those in the new media and mainstream media people get information that is a hoax and misleading. The hoax and misleading news related to Covid-19 cause excessive fear in the community. On the other hand, vigilance is needed both personally and in groups in preventing the transmission of the virus. [3]

Not a few of this news is used by irresponsible people to benefit. The news even led to legal proceedings resulting from this misleading hoax news. Data information was taken from the Directorate General of Informatics Applications of the Ministry of Communication and Informatics Indonesia,

[1] Paramasari and Nugroho, "Health Communication Strategy in An Effort to Build Public Participation during the Covid-19 Pandemic," *Journal of Communication Pearl Lenses* 5, no. 1 (2021): 123–132.

[2] Unit Assigment Handling, "COVID-19 Handling Task Force. The Covid-19 Situation in Indonesia," last modified 2022, https://covid19.go.id/artikel/2022/05/15/situasi-covid-19-di-indonesia-update-15-mei-2022.

[3] Finset, Bosworth, and Butow, "Effective Health Communication – a Key Factor in Fighting the COVID-19 Pandemic.," *Patient Education and Counseling* 103, no. 5 (2020): 873–876.

from May 6, 2019, to April 4, 2020, there were 2,161 issues of Covid-19 hoaxes in Indonesia. [4][5]

Based on the We Are Social report, the number of active social media users in Indonesia in 2019 was 150 million and until 2022 there were 191 million social media users in Indonesia. That number has increased by 12.35% compared to 2021 which was 170 million people. Seeing the trend, the number of social media users in Indonesia continues to increase every year. However, its growth has fluctuated from 2014-to 2022. [6]

Several previous studies have expressed support for this study. So the formulation of the problem is how much influence health communication has on health information during the 2019-2022 pandemic on people's attitudes toward receiving health information during the pandemic? And the purpose of the research is to find out the magnitude of the influence of Health communication in health information during the 2019-2022 pandemic on people's attitudes toward receiving health information during the pandemic.

II. LITERATURE

Health communication, the study of health-related communication (Ratzan, Payne, & Bishop, 1996), is a relatively new subfield but one of the fastest-growing areas of communication. The formal definition of health communication put forward by the U.S. Department of Health and Human Services (2000): Health communication is the art and technique for informing, influencing, and motivating individual audiences, institutions, and the public about important health issues. The scope of health communication includes disease prevention, health promotion, health care policy, and health care business as well as improving the quality of life and health of individuals in society.

In general, hoaxes are interpreted as fake news. According to Mc Dougall in stating that hoaxes are untruths that are deliberately created to cover up the truth. The phenomenon of hoaxes in the world of mass media is not new. A hoax is considered a hoax that is used so that someone wants to believe something wrong. Hoaxes aim to create, shape, and shepherd public opinion. Others are regarded as jokes or just a fad. [7]

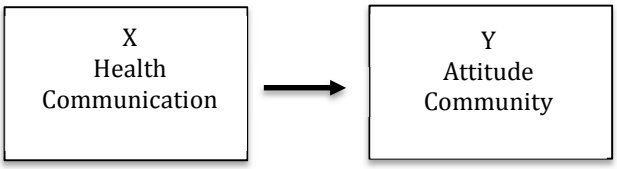
Gradually hoaxes become dangerous information and have a negative impact. In today's modern era, the dissemination of information is done quickly and when a piece of information does not pay attention to the ethics of the news, the public will be confused to distinguish which information is true and what misinformation is wrong. The hoax information in the health sector circulated causing panic and doubt in the community, such as information related to the pandemic. [8]

Secord & Backman define attitude as a certain regularity in terms of feelings (affection), thinking (cognition), and the predisposition of a person's actions (Konami) towards an aspect of the surrounding environment. People are more likely to believe hoaxes if the information is by their opinions, beliefs, or attitudes. The information circulating will naturally cause positive feelings in the reader, when his beliefs or opinions get affirmations. This causes people to believe something and not hesitate to spread it back regardless of whether the information received is right or wrong.[9]

According to Healthy People 2010, Health communication is an art and technique used to inform, influence, and motivate individuals, institutions, and the public about important issues in the field of health in improving the quality of health. This explains that health communication is increasingly recognized as a necessary element in efforts to improve personal and public health.[10]

The research theory hypothesis is that there is an influence of health communication in health information during the 2019-2022 pandemic on people's attitudes toward receiving health information during the pandemic.

Research model :



Ambon, Makassar, Bandung, and Bekasi

III. METHOD

[4] Benski, Goto, and Reich, *Developing Health Communication Materials During a Pandemic*, *Frontiers in Communication*, 2020.

[5] Kominfo Indonesia, *Covid-19 Hoax Issue.*, 2022.

[6] S. Kemp, "DIGITAL 2022: INDONESIA.," *Https://Datareportal.Com/Reports/Digital-2022-Indonesia*.

[7] E. J. Thaib, "Hoax In Social Media and Its Threats To Islamic Moderation In Indonesia," *Proceedings of International Conference on Da'wa and Communication* 2, no. 1 (2020).

[8] K. Y.S. Putri et al., "Transformation of Health Communication Literacy in the Pandemic Era," *Informasi* 51, no. 1 (2021): 93–110.

[9] D. R. Rahadi, "User Behavior And Hoax Information On Social Media," *Journal of Management and Entrepreneurship* 5, no. 1 (2017): 58–70, https://doi.org/10.26905/jmdk.v5i1.1342.

[10] Endah Endrawati, "Penerapan Komunikasi Kesehatan Untuk Pencegahan Penyakit Leptospirosis Pada Masyarakat Desa Sumberagung, Kecamatan Moyudan, Sleman, Yogyakarta ," *Jurnal Komunikasi* 7, no. 1 (2015): 1–25.

Respondents to the study numbered 287, quantitative methods. Research areas in Indonesia are Ambon, Makassar, Bangka, Bandung and Bekasi. The results of the validity of health communication variables are the value KMO = .894 Sign = .000 and the attitude variable KMO = .831 Sign = .000. alpha Cronbach health communication variable reliability values .862 and attitude .854.

Table 1

Validity and Reliability

Variable	Validity	Reliability
Health Communication	KMO = .894 Sign = .000	.862
Community Attitude	KMO = .831 Sign = .000	.854

Because both variables have met the research requirements, the next statistical analysis is continued.

IV. RESULT

Descriptive research results of health communication research all dimensions of health communication and public attitudes are positive value, according to the research figure (Table 2).

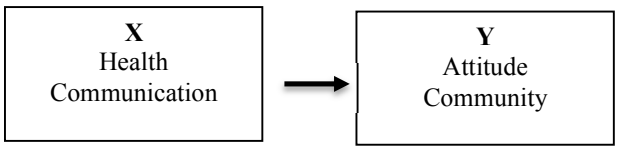
Table 2

Descriptive Research Results

Dimensions of Health Communication	Mean	Attitude Dimensions	Mean
Disease prevention	.879	Cognitive	.837
Health Promotion	.782	Affective	.794
Health Maintenance Policy	.865	Conative	.836
Business Regulation in the Field of Health	.845		
Renewing the quality of individuals in a community of society by considering aspects of science	.891		
Ethics	.866		

Based on the results of descriptive research on the dimensions of health communication, namely disease prevention has a mean value of 0.879, health promotion has a mean value of 0.782, health maintenance policy has a mean value of 0.865, business regulation in the field of health has a mean value of 0.845, updating the quality of individuals in a community of society by considering aspects of science has a mean value of 0.891, and ethics has a mean value of 0.866. Health communication is a technique to inform and influence the decisions of individuals and groups to improve health. In this study, techniques to inform and influence individual decisions use dimensions of health communication such as disease prevention, health promotion, and health maintenance, to update the quality of individuals taking into account aspects of science. Attitude is a certain regularity in terms of feelings (affection), thinking (cognition), and the predisposition of a person's actions (Konami) towards an aspect of the surrounding environment. Based on the results of descriptive analysis research, the attitude dimension consists of cognitive which has a mean value of 0.837, affective which has a mean value of 0.794, and conative value which has a mean value of 0.836.

The results of the study with the next statistical analysis are by the alternative hypothesis of the great influence of health communication in health information during the 2019-2022 pandemic on people's attitudes toward receiving health information during the pandemic. The following model images of the results of the study are as follows:



r = .876

Sign = .000

Fig. 1 Research Results Model

Based on the results of the study showed that the value of r = 0.876 and the sign value = 0.000 so it can be known that variable X, namely health communication influences variable Y, namely people's attitudes. Health communication aims to inform and influence individual and group decisions to improve health. An attitude is the behavioral predisposition of the cognitive, affective, and conative components of an object of attitude. The Cognitive Component consists of all aspects of cognition that a person has about a particular object. While the component of cognition that a person has towards an object, the affective component consists of all aspects of one's feelings and feelings of emotion towards an object. While the conative component consists of a person's readiness to react or a certain tendency to behave towards an object. Health communication that provides information can influence people's attitudes toward health cognition, the feeling, and readiness of the community in health to increase. (Princess et al., 2022)

V.CONCLUSION

Descriptive research results of research of all dimensions are of positive value. So that the results of the next statistical analysis are a big influence, meaning that it is the influence of large health communication on public attitudes toward health information during the 2019-2022 pandemic.

VI. REFERENCE

[1.] Benski, Goto, & Reich. (2020). Developing Health Communication Materials During a Pandemic. *Frontiers in Communication*.

[2.] David, E. R. (Eribka), Sondakh, M. (Mariam), & Harilama, S. (Stefi). (2017). The Influence of Vlog Content in Youtube on the Formation of Communication Science Students ' Attitudes, Faculty of Social and Political Sciences, Sam Ratulangi University. *Acta Diurna*, *6*(1), 93363.

[3.] Endrawati, E. (2015). Application of Health Communication for the Prevention of Leptospirosis Disease in the Community of Sumberagung Village, Moyudan District, Sleman, Yogyakarta. *Journal of Communication*, *7*(1), 1–25.

[4.] Finset, Bosworth, & Butow. (2020). Effective health communication – a key factor in fighting the COVID-19 pandemic. Patient *Education and Counseling*, *103*(5), 873–876.

[5.] Kemp, S. (2022, December 15). *DIGITAL 2022: INDONESIA*. https://datareportal.com/reports/digital-2022-indonesia

[6.] Kominfo Indonesia. (2022). *Covid-19 Hoax Issue*.

[7.] Paramasari, & Nugroho. (2021). Health Communication Strategy in An Effort to Build Public Participation during the Covid-19 Pandemic. *Journal of Communication Pearl Lenses*, *5*(1), 123–132.

[8.] Putri, K. Y. S., Mulyana, D., Istiyanto, S.B., Zakiah, L., & Hafni, R. D. (2021). Transformation of health communication literacy in the pandemic era. *Information*, *51*(1), 93–110. https://doi.org/10.21831/informasi.v51i1.38290

[9.] Putri, K. Y. S., Sutjipto, V. W., Anindhita, W., & Romli, N. A. (2022). *Digital Literacy Hoax Information in Indonesian Tourism Area*. *2*(1), 1–11.

[10.] Rahadi, D. R. (2017). User Behavior And Hoax Information On Social Media. *Journal of Management and Entrepreneurship*, *5*(1), 58–70. https://doi.org/10.26905/jmdk.v5i1.1342

[11.] COVID-19 Handling Task Force. (2022, May 15). *The Covid-19 Situation in Indonesia (Update as of May 15, 2022)*. https://covid19.go.id/artikel/2022/05/15/situasi-covid-19-di-indonesia-update-15-mei-2022

[12.] Thaib, E. J. (2020). HOAX IN SOCIAL MEDIA AND ITS THREATS TO ISLAMIC MODERATION IN INDONESIA. *Proceedings of International Conference*

on Da'wa and Communication, *2*(1).Benski, Goto, and Reich. *Developing Health Communication Materials During a Pandemic*. *Frontiers in Communication*, 2020.

[13.] Endrawati, Endah. "Penerapan Komunikasi Kesehatan Untuk Pencegahan Penyakit Leptospirosis Pada Masyarakat Desa Sumberagung, Kecamatan Moyudan, Sleman, Yogyakarta ." *Jurnal Komunikasi* 7, no. 1 (2015): 1–25.

[14.] Finset, Bosworth, and Butow. "Effective Health Communication – a Key Factor in Fighting the COVID-19 Pandemic." *Patient Education and Counseling* 103, no. 5 (2020): 873–876.

[15.] Kemp, S. "DIGITAL 2022: INDONESIA." *Https://Datareportal.Com/Reports/Digital-2022-Indonesia*.

[16.] Kominfo Indonesia. *Covid-19 Hoax Issue.*, 2022.

[17.] Paramasari, and Nugroho. "Health Communication Strategy in An Effort to Build Public Participation during the Covid-19 Pandemic." *Journal of Communication Pearl Lenses* 5, no. 1 (2021): 123–132.

[18.] Putri, K. Y.S., Deddy Mulyana, S Bekti Istiyanto, Linda Zakiah, and Rayni Delya Hafni. "Transformation of Health Communication Literacy in the Pandemic Era." *Informasi* 51, no. 1 (2021): 93–110.

[19.] Rahadi, D. R. "User Behavior And Hoax Information On Social Media." *Journal of Management and Entrepreneurship* 5, no. 1 (2017): 58–70. https://doi.org/10.26905/jmdk.v5i1.1342.

[20.] Thaib, E. J. "Hoax In Social Media and Its Threats To Islamic Moderation In Indonesia." *Proceedings of International Conference on Da'wa and Communication* 2, no. 1 (2020).

[21.] Unit Assigment Handling. "COVID-19 Handling Task Force. The Covid-19 Situation in Indonesia." Last modified 2022. https://covid19.go.id/artikel/2022/05/15/situasi-covid-19-di-indonesia-update-15-mei-2022.

Component Interface Formalization in Robotic Systems

1st Anton Hristozov
Polytechnic Institute
Purdue University
West Lafayette, Indiana, USA
ahristoz@purdue.edu

2nd Dr. Eric Matson
Polytechnic Institute
Purdue University
West Lafayette, Indiana, USA
ematson@purdue.edu

3rd Dr. Eric Dietz
Polytechnic Institute
Purdue University
West Lafayette, Indiana, USA
jedietz@purdue.edu

4th Dr. Marcus Rogers
Polytechnic Institute
Purdue University
West Lafayette, Indiana, USA
rogersmk@purdue.edu

***Abstract*—Components are heavily used in many software systems including robotics systems. The growth of sophistication and diversity of new capabilities for robotic systems presents new challenges to their architectures. Their complexity is growing exponentially with the advent of AI, smart sensors and the complex tasks they have to accomplish. Such complexity requires a more rigorous approach for the creation, use and interoperability of software components. The issue is exacerbated because robotic systems are becoming more and more reliant on third party components for certain functions. In order to achieve this kind of interoperability, including dynamic component replacement, we need a way to standardize their interfaces. A formal approach is desperately needed for specifying what an interface of a robotic software component should contain. This study performs and analysis of the issue and presents a universal and generic approach to standardizing component interfaces for robotic systems. Our approach is inspired by well established robotic architectures such as ROS, PX4 and Ardupilot. The study is also applicable to other software systems that share similar characteristics with robotic systems. We consider the use of JSON or Domain Specific Languages (DSL) development with tools such as Antlr and automatic code and configuration files generation for frameworks such as ROS and PX4. A case study with ROS2 is presented as a proof of concept for the proposed methodology.**

***Index Terms*—CPS, robots, software architecture, interface, ROS, autopilot**

Comparing UV-based and O_3-based Advanced Oxidation for Removal of Emerging Contaminants from Food Processing Digestate Sludge

N. Moradi*, C.M. Lopez-Vazquez **, H. Garcia Hernandez**, F. Rubio Rincon**, D. Brdanovic**, Mark van Loosdrecht***

***Abstract*—**Advanced oxidation processes have been widely used for disinfection, removal of residual organic material, and for the removal of emerging contaminants from drinking water and wastewater. Yet, the application of these technologies on sludge treatment processes has not gain enough attention; mostly, considering the complexity of the sludge matrix. In this research, ozone and UV/H_2O_2 treatment were applied for the removal of emerging contaminates from a digestate supernatant. The removal of the following compounds was assessed:(i) salicylic acid (SA) (a surrogate of non-stradiol anti-inflammatory drugs (NSAIDs)), and (ii) sulfamethoxazole (SMX), sulfamethazine (SMN), and tetracycline (TCN) (the most frequent human and animal antibiotics). The ozone treatment was carried out in a plexiglass bubble column reactor with a capacity of 2.7 L; the system was equipped with a stirrer and a gas diffuser. The UV and UV/H_2O_2 treatments were done using on a LED set-up (PearlLab beam device) dosing H□O□. In the ozone treatment evaluations, 95 % of the three antibiotics were removed during first 20 min of exposure time, while an SA removal of 91 % occurred after 8 hours of exposure time. In the UV treatment evaluations, when adding the optimum dose of hydrogen peroxide (H_2O_2:COD molar ratio of 0.634), 36% of SA, 82% of TCN, and more than 90 % of both SMX and SMN were removed after 8 hours of exposure time. This study concluded that O_3 was more effective than UV/H_2O_2 removing emerging contaminates from the digestate supernatant.

***Keywords*—**Digestate sludge, Emerging contaminants, Ozone, UV-AOP.

*Department of Water Supply, Sanitation and Environmental Engineering. IHE Delft Institute for Water Education, 2611 AX Delft, The Netherlands. Department of Biotechnology, Delft University of Technology. n.moradi@un-ihe.org** Department of Water Supply, Sanitation and Environmental Engineering. IHE Delft Institute for Water Education, 2611 AX Delft, The Netherlands. c.lopezvazquez@un-ihe.org

***Department of Biotechnology, Delft University of Technology, the Netherlands.

A Review Paper on Data Security in Precision Agriculture using Internet of Things

Tonderai Muchenje
Department of Information Technology Tshwane University of Technology Pretoria, South Africa
Muchenjet@tut.ac.za

Xolani Mkhwanazi
Department of Information Technology Tshwane University of Technology Pretoria, South Africa
xolanifanasibili@gmail.com

Abstract

Precision agriculture uses a number of technologies, devices, protocols and computing paradigms to optimize agricultural processes. Big data, artificial intelligence, cloud computing and edge computing are all used to handle the huge amounts of data generated by precision agriculture. However, precision agriculture is still emerging and has a low level of security features. Furthermore, future solutions will demand data availability and accuracy as key points to help farmers and security is important to building robust and efficient systems. Since precision agriculture comprises a wide variety and quantity of resources, security addresses issues such as compatibility, constrained resources, and massive data. Moreover, conventional protection schemes used in the traditional internet may not be useful for agricultural systems, creating extra demands and opportunities. Therefore, this paper aims at reviewing the state of the art of precision agriculture security, particularly in open field agriculture, discussing its architecture, describing security issues, presenting the major challenges and future directions.

Keywords

Precision agriculture, Security, IoT

I. INTRODUCTION

Smart farming is defined as a set of tools and methods that, based on IT technologies development, results in significantly different production technology solutions. Furthermore, they can be applied to make farming more efficient and effective while taking environmental and sustainability into account. The agricultural sector plays an important role throughout the globe. The rapid growth of information and communication technologies is strongly influencing the structure and procedures of contemporary agriculture. Today the utilization of data tools has become commonplace, without their application, we cannot imagine our world. Agriculture is the most vital provider of food and provides significant role in economic growth. According to (Sharma et al.,2021) it is expected that the global demand for food must increase to 70 percent by 2050 to meet demand. However, the current production suffices to feed the entire globe population. Moreover, researchers discovered that 500 million people still

suffer from malnutrition, and about 821 million go to bed without food. According to (Sharma et al.,2021) the world population is expected to increase by over 2 billion people mostly in the African countries. Furthermore, this increase in population represents a challenge to reach the goal of ensuring that no one is in hunger. So, the expectations influence the global demand for food. It might be difficult to meet 40% of water demands by 2030 due to irregularity of weather and the degradation of about 20% of arable land is going to reduce food supply. Moreover, food growth or production requires lot of resources than currently available and more accurate and sustainable systems to increase cultivation rates and decrease the use of natural resources. Furthermore, systems that can be used to provide accurate and sustainable data is artificial intelligence, cloud, and edge computing, can help to keep, store, and analyze big data that is generated.
However, it is important that producers, processors, and farmers use high-technological tools or systems to optimize production conditions use resources efficiently and reduce waste. The use of IoT farming technologies and precision technologies can be a good solution. Smart farming technologies include data collection and processing to improve crop fields and food quality. For instant a moisture sensor placed in the ground can be used to reduce the application of excess water. Moreover, such technologies make it possible to create the right environment conditions for plants in greenhouses or even animals. Therefore, information technology assets are threatened by serious threats, by significant cybersecurity risks. Furthermore, my research aims to improve security risk in smart farming using IoT.

In this paper section A. gives an overview of data privacy and data confidentiality . B. assuring information integrity. C. sensors and actuators connects to the network and gather data.

II. SYSTEM ARCHITECTURE

Since 1900, agricultural mechanization(between 1900 and 1930) has introduced machines and implements to computerize work, increasing laborers productivity. Precision agriculture has undergone several revolutions, which have improved its efficiency and profitability. According the Green Revolution (about the 1960s) enabled agriculturalists to use new crop variations and agrochemicals.

Precision agriculture systems use sensors to collect climate or environmental data. Sensors connect to a constrained border device, called a gateway, which provides limited computational capabilities to a local computer through a network connection, frequently wireless. Moreover, the local computer receives data from the gateway, stores it in a database, and shows processed information on a web page. Furthermore, Local systems did not integrate with external systems or the Internet.

However general purpose systems, just implement IoT technologies and resources or design web-services , alert services , traceability resources and control on the cloud . Even though there are several solutions in Agriculture, they are still immature and provide a low level of intelligence. Additionally, many of these proposals rely on sensors and actuators to send data to the gateway, which limits automation. Consequently, local systems rarely integrate with the internet , although they do sometimes store data in the cloud.

Moreover, the perception layer consists of sensors, GPS, tags RFID, cameras, actuators, and any other devices that collect data from the farm environment and acting to modify them. Furthermore, these devices lack the computational capacity to process or store data and perform at the edge or the cloud. Therefore, the network layer communicates with edge resources through wireless sensor networks (WSN).

a). data security strategy involves technical safeguards that ensure data confidentiality, integrity, and availability, while a data privacy strategy involves the use of personal information and managed.

b). assuring information integrity, security resources, accuracy, and validity involves preventing unauthorized changes or deletions.

c). sensors in a wireless sensor and actuator network(WSAN) gather information about their environment. In an autonomous or human-controlled environment, all elements communicate wirelessly.

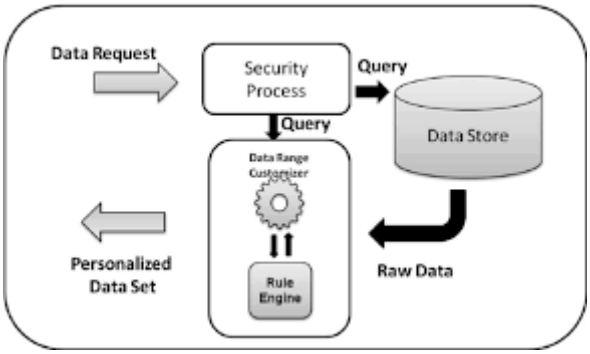


Figure 1:

The diagram above is summarized to show the network architecture on how security detection will be applied to secure data.

The network edge layer may comprise a variation of resources like security measures, data filters, decision-making capability, diversified processing, in-out interface, and thus the gateway. Including one or more resources at the sting depends on the features of the appliance. Some appliances support only the retransmission of knowledge, while others have the computational capability to perform more tasks. However, robust gateways can process data, make decisions, and send commands to actuators and data to the cloud. Moreover, the network Service Provider (ISP) connects the gateway to the cloud. Furthermore, the cloud processes and stores data to offer end-users with information and services. Hence, processing is also a challenge, considering the big mass of data produced by the perception devices reaching the massive Data world, and also the financial cost of processing within the cloud.

Moreover, processing everything within the cloud, as proposed by many solutions, implies enormous bandwidth requirements and high financing costs. it should be advantageous to use a sturdy gateway and perform a component of the processing at the sting. Moreover, moving an element of the subsystems to the sting may reduce the financial costs of smart farming. Furthermore data consumed or pre-processed at the sting saves bandwidth and should reduce the computing resources required from the cloud, protects privacy, and preserves the battery lifetime of some devices. Hence,

the cloud could store and process massive data, make decisions, and interact with the user.

However, processing big data to make decisions at the sting may require tools like technology. Moreover, the improvement of devices and communication technologies will make it possible for more computational resources to be integrated into systems. Furthermore, such integration aims at meeting different demands of agricultural automation, farm management, and precision farming. Solutions must evolve to management systems rather than just monitoring, which could end in new challenges and possibilities. Hence, another issue is that the safety of information. From the detection up to the storage and decision-making within the cloud, it's mandatory to produce data privacy, reliability, and accuracy. Therefore, Security issues in precision agriculture are a major challenge and may be detailed

III. RELATED WORK

This part of the section explains some of the related work on data security using Internet of Things and detecting possible attacks on the network. Furthermore, many authors have contributed on this area of precision agriculture.

Reviews shows that attacks on hardware: unprotected vulnerabilities of IoT and cyber-physical devices may be exploited by professional attackers using specialized tools. For instance, according (Beno's et al.,2021) state that they can use side channel which can violate privacy, confidentiality and authenticity when attack poorly designed IoT and cyber-physical systems. Therefore, a gap has been recognized for a need of side channel attacks: that aims at gathering unauthorized information regarding the implementation detail of a system and monitoring physical parameters of data, such electrical current or bandwidth because attacks of this type violate confidentiality of the system.
Some researchers conducted studies on data integrity which guarantees information not to be changed during storage or transmission.

The reviews shows that integrity of precision agriculture has been a topic in most research projects including confidentiality which protects data against unauthorized access says (Zhang et al., 2020). Thus, this interest showed a gap and importance of security and data privacy.

A study conducted by Barreto et al (2018). Takes an empirical approach towards the identification or pointing of cyber-security challenges. However, it is pointed that the most important work of the scope of improving security is those focusing on threats. Among these reviews we can mention one the study of Demestichas et al (2018), which fails to develop a taxonomy. Therefore, privacy is required to keep unauthorized permission to other users' information from being accessed. Consequently, physical attacks can compromise data and lead to the violation of privacy.

Study reveals that trust makes it impossible for users to spoof another identity says (Trendov et al.,2019). However, researchers are still conducting a study on authenticity in precision agriculture, because in the absence of proper provisions for security, can be exposed to variety of attacks that may exploit the environment of precision agriculture, and related smart information systems, cause harm, unauthorized change, and destruction to the data.

Reviews on the other hand on security features and other related technologies have been of interest to most researchers such as (Sanchez et al.,2019) in recent years. Especially the roles of internet of things (IoT) and artificial intelligence (AI) have received great attention. Moreover, challenges related to the security and privacy have been studied without presenting taxonomy of threats.

Researchers have conducted reviews on different structures of precision agriculture and some of these surveys are based on AI techniques, the design of sensors and sensor networks to support smart farming. The application of hardware usage such as raspberry Pi and the technical aspects such as imaging techniques and routing protocols of the systems. hence it was discovered that few researchers have reviewed literature on threats to precision farming. For instance, Boghossian et al. have presented a study on vulnerabilities of precision agriculture as well as the related threats with limited threat mitigation strategies for data collection and privacy.

Research of (Tahaseen and Moparthi, 2021) shows that the explosive growth of IoT systems and large data further fuel the transformation of traditional agriculture into digital, knowledge-based agriculture, where data collection, communication and data-based decision-making are taking place at a rapid speed. Moreover, IoT devices assist with data collection sensors connected to tractors, trucks, fields, soil and plants to collect real time data that analysts have instant access to. For instant data collected by intelligent climate monitoring systems can be used to map weather conditions. Sensors can monitor temperature, soil moisture humidity and other variables and even adjust environmental variables to enhance optimal plant growth

The literature review by (Nigam et al.,2019) identifies a problem related to precision agriculture that they must take cybersecurity risks seriously, because it affects the mechanical and digital aspects the same. It is identified that most vulnerabilities are similar to those found in industrial applications of technology such as use of data security, sensors, drones, and controls. Moreover, agriculture businesses are vulnerable to cyberattacks of malware, spearfishing, technology corruption and data breaches for economic manipulation. However, farmers do not see their use of technology as being a danger and susceptible to attacks, they are focused on the potential benefits

The study reveals that systems connected to the internet can easily fall victim to cybercriminals, who on the other hand can recover the data they have acquired and on the other hand paralyze the entire automated systems and cause any damage in this regard says (Sahoo et al.,2018). Furthermore, potential agricultural attacks can cause an unsafe and product-less farming environment. Attackers in closed system such greenhouse, livestock farm, and environmental conditions can be modified to destroy the plants and animals

study shows a gap that more robust gateways can process data, send commands to actuators and data to the cloud and make decision. The cloud computing processes and stores data to provide end-users with information and services. However, data processing is a challenge, considering a large mass of data produced by the perception devices to reach the big data and cost of processing in the cloud says (Kemeny et al,.2017). Therefore, processing data in the cloud as discovered by many solutions implies big bandwidth requirements and enormous financial costs.

IV. RESEARCH METHODS USED

This review is based on the quantitative methodology. The designs used to procure the results is mainly focused on data collection in precision agriculture using the internet of things. The literature review played an important role in incorporating different algorithms that can be used to detect data accuracy and data integrity. Additionally, this was obtained through the use of journals, articles and literature reviews, where we get to find what other authors have found through their experiments. Furthermore, the use of sensors, actuators and drones has made it easier for data capturing and monitoring data storage. Therefore, my aim is to optimize the current systems designed for collecting data by incorporating internet of things and adding data privacy.

V. THE PROPOSED SOLUTION

The enhanced data integrity encryption (EDIE) algorithm has been proposed for enhancing data security using IoT in precision agriculture. The proposed system is designed to incorporate detecting protocol kw2000 for ensuring security privacy and integrity. The goal of this section is to describe a method for balancing data access, as required by the GGIAR Open Access and Data management (OADM) policy (CGIAR 2014), with data privacy goals, as described in the CGIAR Responsible data Guidelines (CGIAR 2019). Therefore, proposed solution will reduce cyberattacks by implementing NMAP, SolarWinds security event manager and Netsparker which will make it difficult for attackers to invade the confidentiality data.

VI. CONCLUSION AND FUTUREWORK

The agricultural techniques modernization is important to increase production rates and preserve natural resources. Precision agriculture can enhance farming tasks by providing efficient control of actuators, optimizing utility and resource use, managing production, maximizing profit, and minimizing costs. However, to accomplish this goal, smart systems must include more computational capabilities, such as edge computing, handling massive data, artificial intelligence resources, and security features. Moreover, security requires special attention as constrained devices generate a large volume of data and forward them to the gateway or the cloud. Therefore, the farming system must protect the data from the detection through to decision-making and storage.

Even though many security threats can affect agricultural systems, they still incorporate a few security resources. This is because these solutions are still in their early stages of development. Furthermore, there are only automation resources implemented, and these have few computational resources. Hence, security features are not yet on the list of system requirements. However, reaching an additional level of precision agriculture demands solutions with security mechanisms that give them enough reliability and accuracy to implement these systems on a large scale. As precision agriculture creates an extra set of challenges, it also presents fresh research opportunities both in security and in other areas of computer science.

ACKNOWLEDGMENT

I would just like to extend my gratitude to all participants in this research for allowing me to take up their time and ask sometimes seemingly pointless questions, and especially Dr. T Muchenje for guiding my research. Furthermore, I would like to thank some

of my classmates who have been valuable in sharing materials and explaining some complicated parts of the research.

REFERENCES

[1] Aldossary, S., and Allen, W.: 'Data security, privacy, availability and integrity in cloud computing: issues and current solutions', International Journal of Advanced Computer Science and Applications, 2016, 7, (4), pp. 485-498 I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[2] Apolinário, F., Pardal, M., and Correia, M.: 'S-Audit: Efficient Data Integrity Verification for Cloud Storage', in Editor (Ed.)^(Eds.): 'Book S-Audit: Efficient Data Integrity Verification for Cloud Storage' (IEEE, 2018, edn.), pp. 465-474.

[3] Boissard, P., Martin, V. & Moisan, S. (2016) A cognitive vision approach to early pest detection in greenhouse crops. Journal of Computers and Electronics in Agriculture. 2(3). pp. 81-93

[4] Karimi, Y. Prasher, O. Patel, M. & Kim, H. (2014) Application of support vector machine technology for weed and nitrogen stress detection in Precision Agriculture. Journal of Computers and Electronics in Agriculture. 51 (1–2). Pp.99–109M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[5] Yang, C., Chen, X., and Xiang, Y.: 'Blockchain-based publicly verifiable data deletion scheme for cloud storage', Journal of Network and Computer Applications, 2018, 103, pp. 185-193

[6] Hunt, T., Zhu, Z., Xu, Y., Peter, S., and Witchel, E.: 'Ryoan: A distributed sandbox for untrusted computation on secret data', ACM Transactions on Computer Systems (TOCS), 2018, 35, (4), pp. 1-32

[7] Srbinovska, M. Gavrovski, C. Dimcev, V. Krkoleva, A. & Borozan, V. (2015) Environmental parameters monitoring in precision agriculture using wireless sensor Esposito, C.: 'Interoperable, dynamic and privacy-preserving access control for cloud data storage when integrating heterogeneous organizations', Journal of Network and Computer Applications, 2018, 108, pp. 124-136

[8] Chadwick, D.W., Fan, W., Costantino, G., De Lemos, R., Di Cerbo, F., Herwono, I., Manea, M., Mori, P., Sajjad, A., and Wang, X.-S.: 'A cloud-edge based data security architecture for sharing and analysing cyber threat information', Future Generation Computer Systems, 2020, 102, pp. 710-722

Studies of Ectoparasite Bionomics Among Household Animals in Ebonyi State

**Uhuo, Cosmas Augustus, Ude Ifeanyi Augustine, Nwokporo, Nkiruka, Obealor Ogochukwu
Department of Applied Biology, Ebonyi State University Abakaliki.

***Abstract*—** The study of ectoparasites bionomics among household animals in Ebonyi State was conducted in Ebonyi central district in 2019 using standard entomological techniques. A total of 300 animals was examined at different predilection sites and hygiene for ectoparasite bionomics. The result revealed that of the 539(179.7%) ectoparasite recovered, sheep had the highest prevalence 215(39.80%) followed by goats 191(35.40 %) and lastly dog 133(24.66 %). There was a mixed infestation among the animals examined of which sheep was infested with 55(47.83 %) ticks, goat 50(43.48 %) and dog 40(34.78 %). These ectoparasites had a high infestation rate in the animals assessed with very poor hygiene and nutritional standard, therefore control and treatment of these infested animals should be taken seriously because of the damage caused on these animals. The study also revealed that most ectoparasites were choice specific in infestation. It is therefore recommended that further research on the occurrence and spread of ectoparasitic diseases should be carried out to cover other areas of Ebonyi state.

***Keywords*—** Ectoparasites, Tick, Bionomics, sheep, goat, dog, Ebonyi State.

Mentha Piperita Formulations in Natural Deep Eutectic Solvents: Phenolic Profile and Biological Activity

Tatjana Jurić[1,*], Bojana Blagojević[1], Denis Uka[1], Ružica Ždero Pavlović[1], Boris M. Popović[1]

***Abstract*—** Natural deep eutectic solvents (NADES) represent a class of modern systems that have been developed as a green alternative to toxic organic solvents, which are commonly used as extraction media. It has been considered that hydrogen bonding is the main interaction leading to the formation of NADES. The aim of this study was phytochemical characterization and determination of the antioxidant and antibacterial activity of *Mentha piperita* leaf extracts obtained by six choline chloride-based NADES. NADES were prepared by mixing choline chloride with different hydrogen bond donors in 1:1 molar ratio following the addition of 30% (w/w) water. The mixtures were then heated (60 °C) and stirred (650 rpm) until the clear homogenous liquids were obtained. The *Mentha piperita* extracts were prepared by mixing 75 mg of peppermint leaves with 1 mL of NADES following by the heating and stirring (60 °C, 650 rpm) within 30 min. The content of six phenolics in extracts was determined using HPLC-PDA. The dominant compounds presented in peppermint leaves - rosmarinic acid and luteolin 7-O-glucoside, were extracted by NADES at a similar level as 70% ethanol. The microdilution method was applied to test the antibacterial activity of extracts. Compared with 70% ethanol, all NADES systems showed higher antibacterial activity towards *Pseudomonas aeruginosa* (Gram -), *Staphylococcus aureus* (Gram +), *Escherichia coli* (Gram -), and *Salmonella enterica* (Gram -), especially NADES containing organic acids. The majority of NADES extracts showed a better ability to neutralize DPPH radical than conventional solvent and similar ability to reduce Fe^{3+} to Fe^{2+} ions in FRAP assay. The obtained results introduce NADES systems as the novel, sustainable, and low-cost solvents with a variety of applications.

***Keywords*—**Antibacterial activity, antioxidant activity, green extraction, natural deep eutectic solvents, polyphenols.

[1] Chemistry&Biochemistry Laboratory, Department of Field and Vegetable Crops, Faculty of Agriculture, University of Novi Sad, Trg Dositeja Obradovića 8, 21000 Novi Sad, Serbia

*Corresponding author: Tatjana Jurić (tatjana.juric@polj.uns.ac.rs)

Novel Delivery Systems for Fruit Polyphenols

Boris M. Popović, Tatjana Jurić, Bojana Blagojević, Denis Uka, Ružica Ždero Pavlović

***Abstract*—** Green solvents are environmentally friendly and greatly improve the sustainability of chemical processes. There is a growing interest in the green extraction of polyphenols from fruits. In this study, we consider three Natural Deep Eutectic Solvents (NADES) systems based on choline chloride as a hydrogen bond acceptor and malic acid, urea, and fructose as hydrogen bond donors. NADES systems were prepared by heating and stirring, ultrasound, and microwave (MW) methods. Sour cherry pomace was used as a natural source of polyphenols. Polyphenol extraction from cherry pomace was performed by ultrasound-assisted extraction, microwave-assisted extraction and compared with conventional heat and stirring method extraction. It was found that MW-assisted preparation of NADES was the fastest, requiring less than 30 s. Also, MW extraction of polyphenols was the most rapid, with less than 5 min necessary for the extract preparation. All three NADES systems were highly efficient for anthocyanin extraction, but the most efficient was the system with malic acid as hydrogen bond donor (yield of anthocyanin content was enhanced by 62.33% after MW extraction with NADES compared with the conventional solvent).

***Keywords*—**Anthocyanins, green extraction, NADES, polyphenols

B. M. P. Author is with the Faculty of Agriculture, University of Novi Sad, Novi Sad, 21000 Serbia (phone:+381214853424; e-mail: boris.popovic@polj.uns.ac.rs).

T. J. Author is with the Faculty of Agriculture, University of Novi Sad, Novi Sad, 21000 Serbia. (e-mail: tatjana.juric@polj.uns.ac.rs).

B. B. Author is with the Faculty of Agriculture, University of Novi Sad, Novi Sad, 21000 Serbia. (e-mail: bojana.blagojevic@polj.uns.ac.rs).

D. U. Author is with the Faculty of Agriculture, University of Novi Sad, Novi Sad, 21000 Serbia. (e-mail: ukadenis5@gmail.com).

R. Ž. P. Author is with the Faculty of Agriculture, University of Novi Sad, Novi Sad, 21000 Serbia. (e-mail: ruzica.zdero@polj.uns.ac.rs).

Genomic Sequence Representation Learning: An Analysis of *K*-Mer Vector Embedding Dimensionality

James Jr.Mashiyane, Risuna Nkolele, Stephanie J. Müller, Gciniwe S. Dlamini, Rebone L. Meraba, Darlington S. Mapiye

Abstract—When performing language tasks in natural language processing (NLP), the dimensionality of word embeddings is chosen either ad-hoc or is calculated by optimizing the Pairwise Inner Product (PIP) loss. The PIP loss is a metric which measures the dissimilarity between word embeddings and it is obtained through matrix perturbation theory by utilizing the unitary invariance of word embeddings. Unlike in natural language, in genomics, especially In genome sequence processing, unlike in natural language processing, there is no notion of a "word", but rather, there are sequence substrings of length k called *k*-mers. *K*-mers sizes matter and they vary depending on the goal of task at hand. The dimensionality of word embeddings in NLP has been studied using the matrix perturbation theory and the PIP loss. In this paper, the sufficiency and reliability of applying word-embedding algorithms to various genomic sequence datasets is investigated to understand the relationship between the *k*-mer size and their embedding dimension. This is completed by studying the scaling capability of three embedding algorithms, namely Latent Semantic analysis (LSA), Word2Vec, and Global Vectors (GloVe) with respect to the *k*-mer size. Utilising the PIP loss as a metric to train embeddings on different datasets, we also show that Word2Vec outperforms LSA and GloVe in computing accurate embeddings as both the *k*-mer size and vocabulary increase. Finally, the shortcomings of natural language processing embedding algorithms in performing genomic tasks are discussed.

Keywords—word embeddings, *k*-mer embedding, dimensionality reduction

I. Introduction

IN computational genomics, the use of *k*-mer representation has been widely explored in analyzing genomic sequence data [1]–[6]. *K*-mer representations are obtained by taking a genomic sequence of length L and splitting it into smaller substrings, called *k*-mers, using a sliding window of s (default value of $s = 1$). These *k*-mers are treated as basic tokens of analyses in a similar way that words are treated as tokens in many traditional natural language processing (NLP) applications. There are several ways of numerically encoding *k*-mers for use in machine learning analysis. These include methods such as term frequency-inverse document frequency (TF-iDf) vector representation, one-hot encoding vector and continuous distributed vector [1], [3], [5], [7], [8], [9]. In Machine Learning, *k*-mers are regarded as "features" of importance, thus can be represented in categorical, single numerical representation or in a vector representation.

One-hot encodings have several disadvantages. First, the distance between the word vectors is equal, which is not a true representation of the relationship between several word vectors. Another disadvantage of using one-hot encodings is that as the *k*-mer size increases, the dimensions of the *k*-mer vectors also increases (i.e $d = 4^k$), leading to data sparsity and computational complexity for downstream machine learning (ML) models, since most ML models are either linear or quadratic functions of the embedding vector [1], [10], [11]. Continuous distribution vector representation of *k*-mers has also been widely used for ML tasks [1], [3], [5]. In these studies, however, the dimensions of the embedding are chosen *ad hoc* and commonly derived using the popular embedding algorithm Word2Vec [12]. Using Word2Vec, embedding dimensions between 100 and 300 are often chosen due to the large influence of NLP research [12].

In NLP research, word embeddings aid in the investigation of a wide range of problems such as language modelling tasks which include text classification, text summarization, and machine translation [13]–[20]. In all of these applications, word embeddings serve as fundamental inputs to downstream ML models. Choosing the optimum dimensionality of embeddings can have computational consequences and may lead to sub-optimal results when using ML models built on embeddings. For example, if a small dimensionality of embeddings is chosen, the vector tends to not be representative of the data, while choosing a large dimensionality may lead to over-fitting, model complexity and slow training time [10], [21].

Conventionally, the dimensions of word embeddings are chosen *ad hoc* between the standard chosen values of 100 and 300 [12]. This was until recent work introduced the Pairwise Inner Product (PIP) loss, a novel metric to calculate dissimilarity between word embeddings [10], [21]. Using the PIP loss and matrix factorization methods, the optimum dimensionality of embeddings may be obtained, thereby reducing noise by trading off spectral signal [10], [21]. In this paper, this investigation of embedding dimensionality is extended to genomic sequence data. Unlike in NLP where there is a well-defined token and the main task is to determine the embedding dimensionality, analysis involving genomic sequence data require that the tokens either be chosen *ad hoc* or determined using resampling methods [9].

To the best of our knowledge, no studies to date have investigated the relationship between the *k*-mer size and its corresponding optimum embedding dimensionality in genomic sequences. Thus, there is a need to study these methods that don't only optimize the *k*-mer size, but also find the optimum dimensionality of the *k*-mer embeddings so that these may be utilized with the latest ML algorithms

James Jr. Mashiyane is with the Accelarated Discovery(IAD) team at IBM Research Africa(South Africa), University of thw Witwatersrand, Braamfontein, Johannesburg, 2001 South Africa e-mail: james.mashiyane1@ibm.com.

J. Doe and J. Doe are with Anonymous University.

and allow for improved performance. We limit our study to equal-length *k*-mer sizes, studies of dimensionality of variable *k*-mer sizes is left for future research.

The contributions of this study are as follows:

- Using a *k*-mer set of *k* = (3, 4, 5, 6 and 7), experiments using matrix factorization methods are conducted and the PIP loss is optimized to obtain the optimum dimensions for each *k*-mer size.
- It is shown that as the value of the *k*-mer increases, the embedding dimensions and the PIP loss also increase, thus showing that for large values of *k*, the PIP loss is not sufficient to calculate the dimensions of the *k*-mers.
- It is also shown that as the value of α increases, the dimensions of *k*-mers are not affected. Thus, for genomic sequence data, the Latent Semantic Analysis (LSA) algorithm with $\alpha = 0$ performs better compared to Word2Vec and GloVe.
- Finally, an approximation of dimensionality with respect to the *k*-mer size is provided.

For the rest of the paper, the dimensionality of *k*-mer embeddings is calculated for a vocabulary size of n. As in [10], [21], the assumption that unitary *k*-mer embeddings are unitary invariant (two word embeddings are similar if one is a unitary transformation i.e a rotation of the other) are maintained, thereby limiting this study to word embedding algorithms that can be either explicitly or implicitly formulated as a low rank matrix approximation.

A. General Computational workflow

Understanding the multi-dimensional relationship between sub-sequences in sequence data (i.e. *k*-mers) may promote the understanding of features driving genome classification and clustering tasks in genomic sequence analysis. One way to understand such multi-dimensional relationships between the sub-sequences, is to represent them as continuous vector distributions. Choosing the dimensionality of such vector distributions is crucial not only in capturing as much inter-relationship between the sub-sequences, but also minimizing computational complexity.

B. Explicit and Implicit Matrix factorization for document and word embeddings

Here, a review of how both explicit and implicit matrix factorization methods have been used to build word embeddings for various NLP tasks is provided. For explicit matrix factorization, embeddings are defined as,

$$E = f_{\alpha,k}(M) = U_{.,1:k} D^{\alpha}_{1:k,1:k}, \tag{1}$$

where M is computed from the data and is different for the different embedding algorithms and different NLP tasks. For tasks such as document classification [22] or information retrieval [23], one embedding algoprithm that is widely used is the Latent Semantic Analysis/Indexing (LSA/I) which uses bag of words(BoW) model to create a term document frequency matrix M, matrix [24]. In the term-document frequency matrix, the words are represented by rows, and the columns are represented by documents. since the LSA algorithm uses an occuenrence matrix M, in order to learn latent topic relationships, the model uses Singular Value Decomposition(SVD).

Another way to compute M is the word co-occurrence matrix M which is computed using the truncated Single Value Decomposition (SVD), and widely used approximations of M are the Pointwise Mutual Information (PMI) matrix [25] and its variants such as the positive PMI (PPMI) matrix or the shifted positive PMI (SPPMI) matrix [26]. Embeddings can be obtained by using matrix factorization by changing the value of α ([26]–[29]. In addition, different values of α have demonstrated improved performance for different natural language tasks [26], [29]. In the present study, the value of α is not pre-specified and the embeddings are investigated by varying α and choosing dimensionality corresponding to the best value of α.

Other embedding algorithms such as Word2Vec [12] and GloVe [30] have been widely used in various NLP tasks. The Word2Vec algorithm's objective is to maximize the co-occurrence (words-context) likelihood which is defined as,

$$l = \sum_{i=0}^{n} \sum_{j=i-w, j\neq i}^{i+w} log\left(\sigma(v_i v_j^T)\right), \tag{2}$$

where,

$$\sigma(x) = \frac{e^x}{1+e^x}, \tag{3}$$

and v_i are the word vectors and w is the context window. Word2Vec has also been shown to increase performance by adding other techniques such as negative sampling, where the word-context pairs that exist on the data are considered as positive examples [10], [12], [26]. However, for every single positive example found on the data, a number of negative examples are also generated by randomly sampling "fake" word-context pars. Thus, the objective is to both maximize the likelihood of positive word-context co-occurrence and also minimize the likelihood of negative examples [26], [29]. The Word2Vec algorithm with negative sampling (SGNS) has the objective:

$$l = \sum_{w\in V_w} \sum_{c\in V_c} (w,c)\left(log(\vec{w}\cdot\vec{c}) + k\cdot E_{c_N\sim P_D} log\sigma(-\vec{w}\cdot\vec{c})\right), \tag{4}$$

where w is the word vectors and c is the context vectors [26]. SGNS is implicitly factorizing the co-occurrence matrix

$$W\cdot C^T = PMI(w,c) - logk. \tag{5}$$

Thus, the SGNS is factorizing the PMI matrix shifted with a constant. In contrast, the objective of the GloVe algorithm is to maximize the log count matrix raised to an exponential parameter $\gamma \in [0,1]$.

C. The Pairwise Inner Product (PIP) loss

Let E be a $m \times d$ embedding matrix, and let be a $n \times k$ embedding matrix, where $k \leq d$. To measure dissimilarity between the two word embeddings, using the PIP loss.

the PIP loss is defined as,

$$||PIP(E) - PIP(\tilde{E})|| = ||EE^T - \tilde{E}\tilde{E}^T|| \quad (6)$$

$$= \sqrt{\sum_{i,j}(\langle v_i, v_j\rangle - \langle \tilde{v_i}, \tilde{v_j}\rangle)^2}, \quad (7)$$

where the PIP matrix ($PIP(E)$) is defined as,

$$PIP(E) = EE^T, \quad (8)$$

$E \in \mathbb{R}^{n\times d}$ are $\tilde{E} \in \mathbb{R}^{n\times k}$ two word embeddings [10], [21]. The PIP loss is invariant under unitary transformations, which means that for any two embedding matrices E_1 and E_2, such that, $E_1 = UE_2$, where U is a unitary matrix, then,

$$PIP(E_1) = PIP(E_2). \quad (9)$$

The PIP loss measures the vector relative position shifts between E_1 and E_2, and thus is coordinate independent [10], [21]. Furthermore, the PIP loss is a metric of functionality similarity. A small PIP loss between E_1 and E_2's similarity and compositionality is small, See Appendix A, [10], [21]. In the following sections, the use of matrix perturbation is described for the analysis of the PIP loss between the ground truth embedding and the trained embedding for each k-mer size and each dataset. In order to be able to compare the two embeddings, we need to know the ground truth embedding, as shown in [10], [21], the ground truth embedding can be computed by taking advantage of unitary invariance. Since we know that embeddings are unitary invariant, one can generate unitary invariant matrices as estimating of the ground truth embedding.

D. Matrix perturbation Theory

Here, it is described how matrix perturbation theory may be used to analyse the PIP loss and dimensionality of embeddings. Connecting the PIP loss and matrix factorization of vector embeddings using matrix perturbation theory reveals a bias-variance trade-off in dimensionality selection of word embeddings [10], [21]. The bias-variance trade-off can be used to explain the hypothesis of optimum dimensionality in vector word embeddings. Considering the following embedding procedure:

$$f_{\alpha,k}(M) = U_{\cdot,1:k} D^{\alpha}_{1:k,1:k}, \quad (10)$$

M is the input matrix, and $f_{\alpha,k}$ is the embedding matrix using SVD. In practice, M is estimated from the data. Thus, instead of having an accurate M, the estimation of M results in a noisy $\tilde{M}$. If M is observed from an infinitely large corpus, then $E = f_{\alpha,d}(M)$ is the ground truth embedding. If M is noisy($\tilde{M} = M+Z$), when M is estimated from a finite dataset, especially observed in training, the trained embeddings are given by $\tilde{E} = f_{\alpha,k}(\tilde{M})$. Notice that M has a rank d, and $\tilde{M}$ has rank k. The dimensionality of word embeddings has been thoroughly investigated for NLP tasks, where the PIP loss between the ground truth and the trained embeddings is minimized [10], [21]. Similar assumptions on the functionality of *k*-mer embeddings are used in this study. For example, if the PIP loss is close to 0, then the trained embeddings are similar to the ground truth embeddings. Both parameters α, and k are studied for each *k*-mer size ranging from 3 to 7.

In order for the the PIP loss to be minimized, the ground truth embedding needs to be known [21]. Word embeddings are unitary invariant, and leveraging the knowledge that the PIP loss measures the relative distance between two embeddings, it is possible to generate matrices with unitary norms to approximate the ground truth embedding [10]. How matrix perturbation theory is used to understand dimensionality through a bias-variance trade-off will now be described.

Definition: Principal Angles

Suppose there exists orthogonal matrices $X, Y \in \mathbb{R}^{n\times k}$ with $k \leq n$. Let $X^TY = UDV^TT$ be the SVD of X^TY. Therefore,

$$D = \cos(\Theta) = diag(\cos(\theta_1), \ldots, \cos(\theta n)), \quad (11)$$

where, $\Theta = (\theta_1, \theta_2, \ldots, \theta_n)$ are the principal angles between subspaces $(\mathcal{X}, Y)$ which are spanned by columns of X and Y.

Lemma 1: Let $X_0 \in \mathbb{R}^{n\times k}$ and $Y_1 \in \mathbb{R}^{n\times(n-k)}$ be the orthogonal matrices. The SVD of the inner product between X_0 and Y_1 is given by,

$$\text{SVD}(X_0^TY_1) = U_0 \sin(\Theta)\hat{V_1}^T, \quad (12)$$

where Θ are the principal angles between X_0 and the orthogonal complement of Y_1, which is denoted by Y_0.

Lemma 2: Let $X, Y \in \mathbb{R}^{n\times n}$ be two orthogonal matrices, and suppose $X = [X_0, X_1]$ and $Y = [Y_0, Y_1]$, where $X_0, Y_0 \in \mathbb{R}^{n\times k}$ and $X_1, Y_1 \in \mathbb{R}^{n\times(n-k)}$. Therefore,

$$||X_0X_0^T - Y_0Y_0^T|| = c||X_0^TY_1||, \quad (13)$$

where $c = 1$ for 2-norm and $\sqrt{2}$ for Frobenius norms. The proofs of these Lemmas may be found in work by [10], [21], [31].

Theorem 1: Bias-variance trade-off($\alpha = 0$)
Suppose $E \in \mathbb{R}^{n\times d}$ and $\hat{E} \in \mathbb{R}^{n\times k}$ have orthonormal columns, then the PIP loss between E and $\hat{E}$ is given by,

$$||PIP(E) - PIP(\hat{E})||^2 = d - k + 2||\hat{E}^TE^{\perp}||^2. \quad (14)$$

Theorem 2: Bias-variance trade-off ($0 < \alpha \leq 1$)
Let $E = U_{\cdot,1:d} D^{\alpha}_{1:d,1:d} \in \mathbb{R}^{n\times d}$, $\hat{E} = \tilde{U}_{\cdot 1:d}\tilde{D}^{\alpha}_{1:k,1:k} \in \mathbb{R}^{n\times k}$ be the ground truth and trained embeddings, respectively, assuming that E and $\hat{E}$ no longer have orthonormal columns, with $k \leq d$. Suppose that $D = diag(\lambda_i)$ and $\tilde{D} = diag(\tilde{\lambda_i})$, then

$$||PIP(E) - PIP(\hat{E})|| \leq \sqrt{\sum_{i=k+1}^{d} \lambda_i^{4\alpha}} + \sqrt{\sum_{i=1}^{k} (\lambda_i^{2\alpha} - \tilde{\lambda}_i^{2\alpha})^2} \tag{15}$$
$$+ \sqrt{2} \sum_{i=1}^{k} (\lambda_i^{2\alpha} - \tilde{\lambda}_i^{2\alpha}) ||\tilde{U}_{\cdot,1:i} U_{\cdot,i:n}||.$$

Theorems 1 and 2 show the existence of the bias-variance trade-off between the signal lost and the noise perturbing the subspaces [10], [21]. The terms $(d-k)$ and $\sqrt{\sum_{i=k+1}^{d} \lambda_i^{4\alpha}}$ are the bias terms. The second term of 15 is the variance. As k increases, $\hat{E}^T E^{\perp} \rightarrow f(k)$, where f which is directly proportional to k, also increases. The second term of 15 is the variance in the signal magnitudes, while the third term is the variance in the signal directions. Practically, it is difficult to estimate signal directions, and thus, Theorem 3 is also needed.

Theorem 3: Bias-variance trade-off as signal spectrum to noise ratio
Let M be a symmetric signal matrix with spectrum $\{\lambda_i\}_{i=1}^{d}$. Let Z be a symmetric random noise matrix with iid, zero mean, and variance σ^2. Suppose $\tilde{M} = M + Z$ for any $0 \leq \alpha \leq 1$ and $k \leq d$ and the ground truth and estimated embeddings are given by,

$$E = U_{\cdot,1:d} D^{\alpha}_{1:d,1:d}, \hat{E} = \tilde{U}_{\cdot,1:d} \tilde{D}^{\alpha}_{1:d,1:d}, \tag{16}$$

where the SVDs are given by $M = UDV^T$ and $\tilde{M} = \tilde{U}\tilde{D}\tilde{V}^T$. Therefore, when $\alpha = 0$,

$$\mathbb{E}\,||EE^T - \hat{E}\hat{E}^T|| \leq \sqrt{d - k + 2^2 \sum_{r \leq k, s > d} (\lambda_r - \lambda_s)^{-2}}, \tag{17}$$

and when $0 < \alpha \leq 1$,

$$\mathbb{E}\,||EE^T - \hat{E}\hat{E}^T|| \leq \sqrt{\sum_{i=k+1}^{d} \lambda_i^{4\alpha}} + 2\sqrt{2n}\alpha \sqrt{\sum_{i=1}^{k} \lambda_i^{4\alpha-2}} \tag{18}$$
$$+ \sqrt{2} \sum_{i=1}^{k} (\lambda_i^{2\alpha} - \lambda_{i+1}^{2\alpha}) \sigma \sqrt{\sum_{r \leq i < s} (\lambda_r - \lambda_s)^{-2}}. \tag{19}$$

Theorem 3 shows that the bias-variance trade-off is essentially a trade-off between signal spectrum and noise power. The theorem shows the expectation of the PIP loss with no signal directions: only signal magnitudes and noise [10], [21]. If dimensionality is small, a substantial amount of signal power is lost, increasing the magnitude of the first term with a high bias. If the dimensionality is large, the second and third terms become larger as the noise increases, thus increasing the variance.

II. Materials and Methods

A. Data Collection

Three genomic datasets were evaluated in this study, namely a dataset of 16S ribosomal ribonucleic acid (rRNA) gene sequences, the beta subunit of the RNA polymerase (rpoB) gene of the bacterium *Mycobacterium tuberculosis* (*M. tb*), and whole genome assemblies of the Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) responsible for the current COVID19 pandemic. For this study, ethical approval was not required as the datasets used were sourced from publicly accessible databases and contain no personally identifying information.

The 16S rRNA gene sequence dataset was sourced from the National Institute of Health Human Microbiome Project (Available at http://hmpdacc.org/HM16STR/) (Huttenhower et al., 2012; Jane et al., 2009). Participants provided biological samples from four body sites (oral, airways, gut, and vagina) for up to three time points. These body sites were selected in the initial study to provide varying levels of spatial and biological proximity to each other, as well as to their importance in several health conditions potentially driven by the composition of the human microbiome (Huttenhower et al., 2012; Jane et al., 2009). The 16S rRNA gene sequences in FASTA format were used in this study and are referred to as the "16S rRNA" dataset.

Paired-end sequencing reads for the *M. tb* genomes were downloaded from the European Nucleotide Archive (https://www.ebi.ac.uk/ena) and aligned to the H37Rv reference genome for *M. tb* (https://www.ncbi.nlm.nih.gov/) using the Burrows-Wheeler Aligner version 0.7 [32]. Following sequence alignment, duplicate reads were marked and indexed using Picard [33]. Thereafter, reads spanning the rpoB gene were extracted using the positional coordinates [start:759387; end:763362], sourced from Mycobrowser [19] with the addition of a flank on either side of the gene. These reads spanning the rpoB gene were then assembled into contigs using the SPAdes genome assembler v3.13.0 with default parameters. These contigs were used as input data for this study and are referred to as the "TB" dataset. Finally, whole genome assemblies of the SARS-CoV-2 were downloaded from the GISAID database (https://www.gisaid.org/) and belong to the "COVID19" dataset.

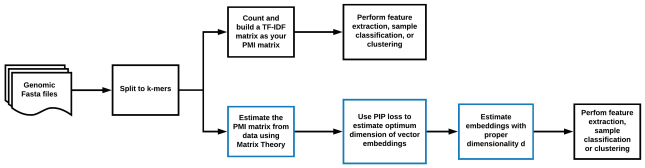
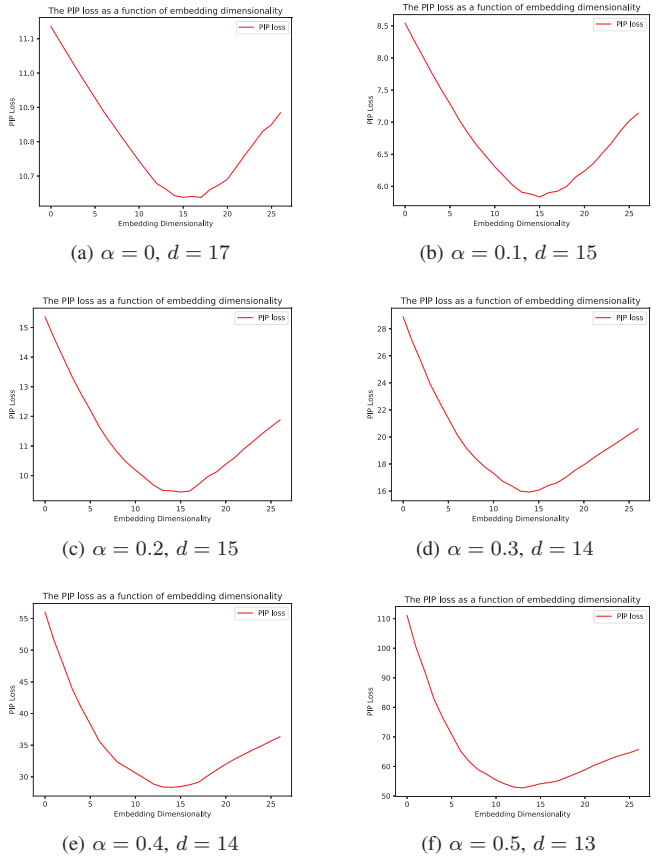


Fig. 1: Optimum dimensionality selection computational work-flow

Figure 1 show the computational work-flow from a fasta file to obtain optimum dimensionality.

III. k-MER SIZE AND DIMENSIONALITY OPTIMIZATION

Theorem 3 shows how the PIP loss between the trained embeddings and ground truth embeddings may be calculated. Estimation is required for the dependent variables, spectrum $D = \{\lambda\}_1^d$ and noise σ. Dimensionality selection may be described as finding the optimal k^* value when the left hand side is a minimum. The work of [10], [21] is followed for estimating the spectrum $D = \{\lambda\}_1^d$ and noise σ. Another approach is to simulate the spectrum and noise using Monte-Carlo methods [10], [21]. It was found that Monte-Carlo simulations are better at estimating these matrices since their distribution is usually around the mean. In this work, Monte-Carlo simulations are also used to estimate the signal and noise for the genomic datasets.



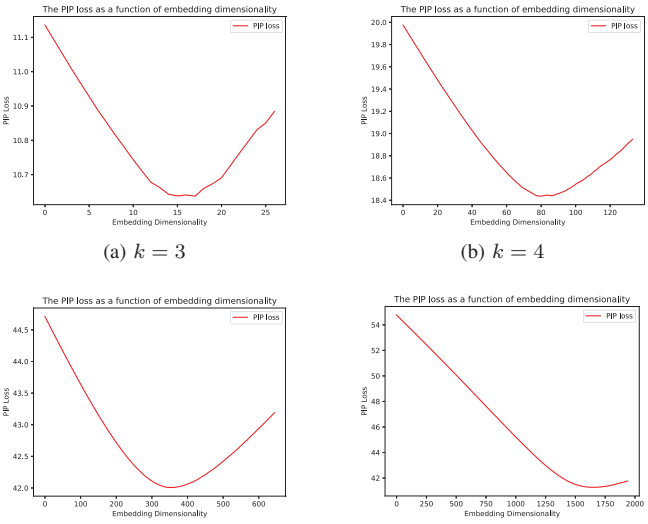
(a) $\alpha = 0$, $d = 17$ (b) $\alpha = 0.1$, $d = 15$

(c) $\alpha = 0.2$, $d = 15$ (d) $\alpha = 0.3$, $d = 14$

(e) $\alpha = 0.4$, $d = 14$ (f) $\alpha = 0.5$, $d = 13$

Fig. 2: PIP error plotted against dimensionality d for the 16S rRNA dataset using the LSA algorithm $k = 3$, $n = 64$ for values of α ranging from 0 to 0.5.

In order to perform calculation of the PIP loss, sequences from the FASTA files of each dataset were split into k-mers of sizes (3, 4, 5, 6 and 7) with no filtering and encoded into random distributed vector representations. These word vectors are inputs to an embedding algorithm (LSA, word2vec, GloVe) to estimate a PMI signal matrix and noise [10], [21], followed by calculation of the PIP loss using Monte-Carlo simulations. Figure 2 and 3 shows the PIP loss against dimensionality for each k-mer size. It was found for all datasets that using



(a) $k = 3$ (b) $k = 4$

(c) $k = 5$ (d) $k = 6$

Fig. 3: 16S rRNA dataset: $\alpha = 0$

the LSA embedding algorithm for the different k-mer sizes, dimensionality was unaffected by varying the value of α.

IV. EXPERIMENTS

The experiments performed in this study of computing optimum dimensionality in k-mer embeddings for genomic sequences are described here. Three different embedding algorithms, namely LSA, word2vec, and GloVe were explored on three genomic datasets, namely a 16S rRNA dataset, TB dataset, and a COVID-19 dataset.

A. Latent Semantic Analysis

For the LSA algorithm, the signal matrix is estimated from the data using the TFi-DF matrix.

TABLE I: LSA: 16S rRNA dataset

k	3	4	5	6	7
n	64	467	1023	4756	16430
rank	22	138	997	2224	5678
$d(\alpha = 0)$	14	80	993	1526	3546
$d(\alpha = 0.1)$	16	78	992	1437	3301
$d(\alpha = 0.2)$	15	73	994	1373	3141
$d(\alpha = 0.3)$	16	72	992	1326	3190
$d(\alpha = 0.4)$	14	69	993	1223	3205
$d(\alpha = 0.5)$	14	69	992	1223	3207

Tables I, II, and III show the results using the LSA algorithm on the 16S rRNA, TB, and COVID-19 datasets, respectively. As observed, changing the value of α at each k-mer value, does not affect the dimensionality of embeddings, unlike in NLP different values of α would work for different languages

TABLE II: LSA: TB dataset

k	3	4	5	6	7	8
n	64	256	954	2727	5250	7037
rank	54	203	823	2385	4907	6914
$d(\alpha = 0)$	49	191	795	2332	4850	6897
$d(\alpha = 0.1)$	49	191	798	2343	4861	6896
$d(\alpha = 0.2)$	50	190	792	2380	4906	6899
$d(\alpha = 0.3)$	49	188	792	2384	4906	6913
$d(\alpha = 0.4)$	51	190	821	2384	4906	6913
$d(\alpha = 0.5)$	47	188	822	2384	4906	6913

TABLE III: LSA: COVID19 dataset

k	3	4	5	6	7
n	64	256	1023	3756	10696
rank	63	244	997	3689	10499
$d(\alpha = 0)$	62	244	997	3684	10478
$d(\alpha = 0.1)$	62	242	992	3683	10478
$d(\alpha = 0.2)$	62	241	994	3682	10476
$d(\alpha = 0.3)$	62	241	992	3682	10478
$d(\alpha = 0.4)$	62	241	993	3681	10476
$d(\alpha = 0.5)$	62	241	992	3682	10478

and the value of α does affect the stability of embedding dimensionality [21], [26]. In addition, as the value of the *k*-mer size increases, the vector dimensionality of these *k*-mer embeddings also increase. In statistics and NLP, the LSA algorithm is known to show state-of-the-art performance in small n, small p(small number of rows, and small number of columns) settings. In this study, it was observed that the LSA algorithm does indeed give lower dimensionality in smaller values of the *k*-mer sizes but only in the 16S rRNA dataset.

We notice that the LSA shows good performance for small k-mer values in the 16S rRNA dataset, on the TB dataset, the LSA algorithm is not able to capture enough relationships between the k-mers. For the TB dataset, unlike the LSA algorithm, Word2Vec shows good performance with dimensionality value of 14. Since it was shown from [10], [21] that word2vec is robust to overfitting, it means for the TB dataset, if we increase dimensionality by a small amount, to try and capture higher order relationships between the k-mers, word2vec would still show optimum performance. gloVe show similar performance to LSA with large dimesnioanlity values.

B. Word2Vec with Skip-Gram

For the word2vec algorithm, the signal matrix is estimated from the data using the PMI matrix and a context window of five.

TABLE IV: Word2Vec: 16S rRNA dataset

k	3	4	5	6	7
n	64	467	1538	4756	16430
rank	43	204	550	1629	5678
$d(\alpha = 0.5)$	23	82	208	693	1050

TABLE V: Word2Vec: TB dataset

k	3	4	5	6	7
n	64	256	954	2727	5250
rank	28	115	517	2264	4907
$d(\alpha = 0.5)$	14	48	324	2263	3356

TABLE VI: Word2Vec: COVID19 dataset

k	3	4	5	6	7
n	64	256	1023	3756	10000
rank	62	212	938	3671	9819
$d(\alpha = 0.5)$	61	206	928	3661	9804

C. GloVe

For the GloVe algorithm, the signal matrix is estimated from the data using the log-count matrix [10].

TABLE VII: GloVe: 16S rRNA dataset

k	3	4	5	6	7
n	109	467	1538	4756	16430
rank	59	269	744	2367	5678
$d(\alpha = 0.5)$	41	101	493	1502	9743

TABLE VIII: GloVe: TB dataset

k	3	4	5	6	7
n	64	256	954	2727	5250
rank	35	210	753	2281	4858
$d(\alpha = 0.5)$	24	197	690	2280	4857

TABLE IX: Dimensionality results using the GloVe algorithm on the COVID19 dataset

k	3	4	5	6	7
n	64	256	1023	3756	10000
rank	52	245	994	3668	9839
$d(\alpha = 0.5)$	47	241	989	3659	9824

V. Conclusion

VI. Discussion and Future work

In this study, it was demonstrated how to obtain the optimum dimensionality using the LSA, word2vec, and GloVe embedding algorithms for each *k*-mer size using three different datasets. Different *k*-mer sizes from 3 to 7 are shown as well as the PIP loss against dimensionality with different values of alpha. For genomic data using the LSA algorithm, overparameterization is observed in smaller values of α unlike in NLP tasks where larger values of α I-II-III. The LSA algorithm showed better performance over word2vec and GloVe in smaller *k*-mer sizes, as the *k*-mer size increases, word2vec shows better performance and stability in computing optimum dimensionality using the PIP loss. Tables IV-V-VI shows results for word2vec. For word2vec with skip-gram, the rapid change in dimensionality from *k*-mer size 4 to 7 is not observed compared to LSA and GloVe as shown in Tables I-II-III and VII-VIII-IX.

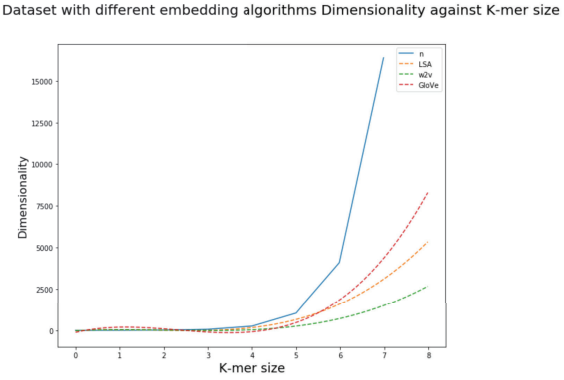


Fig. 4: 16S rRNA with different embedding algorithms, LSA, word2vec, and GloVe

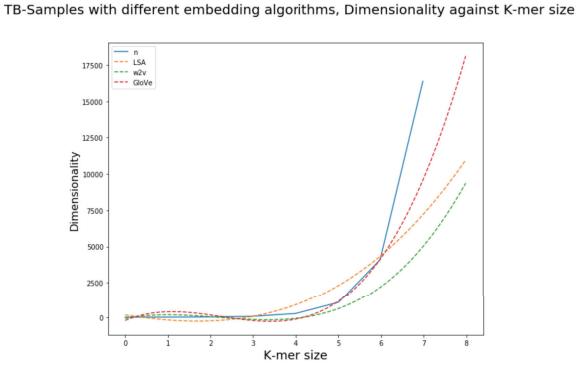


Fig. 5: TB Samples with different embedding algorithms, LSA, word2vec, and GloVe

In order to also understand the relationship between the dimensionality of *k*-mer embeddings and the *k*-mer size, the optimum dimensionality against the *k*-mer size for the three datasets using the LSA algorithm was investigated as shown in Figure 4. The blue line is the graph of the one hot encoding

representation $n = 4^k$. The data points are represented as dots. It was observed that for smaller *k*-mer values i.e $k \leq 4$, dimensionality values are of the same order of magnitude, regardless of the dataset or embedding algorithm. This was also observed in literature, for very small values of the *k*-mer size, the datasets usually end up with a lot of redundancy and thus leading to very low accuracies in computational genomics tasks such as species classification or clustering [7], [9]. In Figure 4, cubic polynomials on the dimensionality and *k*-mer values are shown. Using dimensionality selection,it is possible to reduce the dimensionality size from around $\approx 4^k$, to $d^* = \beta k^3$, where β is a scaling parameter, that scales different for different datasets. To better understand how it changes depending on different genomic datasets and different embedding algorithms, different values of β were not tested as it was beyond the scope of this work. The optimum dimensionality against *k*-mer size for each dataset is also shown in 4-5 using the three embedding algorithms. Figure 4 and 5 shows these results for the 16S rRNA and TB datasets. Similarly smaller values of the *k*-mer size shows not much difference between the embedding algorithms, all the dimensionality values are in the same order of magnitude. It was also observed that when the *k*-mer size increases beyond 5, for different embedding algorithms, dimensionality of *k*-mer embeddings starts to diverge. Lastly, as the *k*-mer size increases, the word2vec algorithm showed better performance in yielding lower dimensionality values compared to LSA and GloVe.

In this work, analyses centred around *k*-mer vector embedding dimensionality with different values of k-mer $\in [3, 4, 5, 6, 7]$ from three genomic sequence datasets. The optimum dimensionality of k-mers using three different algorithms, namely the LSA, Word2Vec and GloVe algorithms were computed. It was observed that as the value of α was varied, the optimum dimensionality value remained stable across all values of the k-mer. It was also found that as the value of the k-mer increased, so the meaning of the vocabulary increased, and the optimum dimensionality also increased. LSA shows optimal performance when *k*-mer is small, as *k*-mer increases, Word2Vec shows optimal performance in dimensionality. This study shows promissing research opportunities in building genomic tailored embedding algorithms by extending ML embedding algorithms such as word2vec and SGNS. [12], the study also presents opportunities in studying the PMI matrices estimated from genomic data using implicit matrix factorization methods. This is left for future research.

REFERENCES

[1] P. Ng, "dna2vec: Consistent vector representations of variable-length k-mers," *arXiv preprint arXiv:1701.06279*, 2017.

[2] E. A. Rødland, "Compact representation of k-mer de bruijn graphs for genome read assembly," *BMC bioinformatics*, vol. 14, no. 1, p. 313, 2013.

[3] A. Fiannaca, M. La Rosa, R. Rizzo, and A. Urso, "A k-mer-based barcode dna classification methodology based on spectral representation and a neural gas network," *Artificial intelligence in medicine*, vol. 64, no. 3, pp. 173–184, 2015.

[4] Y. Guo, K. Tian, H. Zeng, X. Guo, and D. K. Gifford, "A novel k-mer set memory (ksm) motif representation improves regulatory variant prediction," *Genome research*, vol. 28, no. 6, pp. 891–900, 2018.

[5] S. Woloszynek, Z. Zhao, J. Chen, and G. L. Rosen, "16s rrna sequence embeddings: Meaningful numeric feature representations of nucleotide sequences that are convenient for downstream analyses," *PLoS computational biology*, vol. 15, no. 2, p. e1006721, 2019.

[6] D. Kimothi, P. Biyani, J. M. Hogan, A. Soni, and W. Kelly, "Learning supervised embeddings for large scale sequence comparisons," *PloS one*, vol. 15, no. 3, p. e0216636, 2020.

[7] E. Asgari and M. R. Mofrad, "Continuous distributed representation of biological sequences for deep proteomics and genomics," *PloS one*, vol. 10, no. 11, 2015.

[8] M.-f. Zhu, J. Dong, and D.-S. Cao, "rdnase: R package for generating various numerical representation schemes of dna sequences," 2016.

[9] E. Asgari, K. Garakani, A. C. McHardy, and M. R. Mofrad, "Microphenо: predicting environments and host phenotypes from 16s rrna gene sequencing using a k-mer based representation of shallow subsamples," *Bioinformatics*, vol. 34, no. 13, pp. i32–i42, 2018.

[10] Z. Yin and Y. Shen, "On the dimensionality of word embedding," in *Advances in Neural Information Processing Systems*, 2018, pp. 887–898.

[11] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey *et al.*, "Google's neural machine translation system: Bridging the gap between human and machine translation," *arXiv preprint arXiv:1609.08144*, 2016.

[12] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.

[13] A. Jacovi, O. S. Shalom, and Y. Goldberg, "Understanding convolutional neural networks for text classification," *arXiv preprint arXiv:1809.08037*, 2018.

[14] X. Chen and C. Cardie, "Multinomial adversarial networks for multi-domain text classification," *arXiv preprint arXiv:1802.05694*, 2018.

[15] M. Yousefi-Azar and L. Hamey, "Text summarization using unsupervised deep learning," *Expert Systems with Applications*, vol. 68, pp. 93–105, 2017.

[16] S. Song, H. Huang, and T. Ruan, "Abstractive text summarization using lstm-cnn based deep learning," *Multimedia Tools and Applications*, vol. 78, no. 1, pp. 857–875, 2019.

[17] R. Nallapati, B. Zhou, C. Gulcehre, B. Xiang *et al.*, "Abstractive text summarization using sequence-to-sequence rnns and beyond," *arXiv preprint arXiv:1602.06023*, 2016.

[18] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.

[19] A. Kapopoulou, J. M. Lew, and S. T. Cole, "The mycobrowser portal: a comprehensive and manually annotated resource for mycobacterial genomes," *Tuberculosis*, vol. 91, no. 1, pp. 8–13, 2011.

[20] A. Vaswani, S. Bengio, E. Brevdo, F. Chollet, A. N. Gomez, S. Gouws, L. Jones, Ł. Kaiser, N. Kalchbrenner, N. Parmar *et al.*, "Tensor2tensor for neural machine translation," *arXiv preprint arXiv:1803.07416*, 2018.

[21] Z. Yin, "Understand functionality and dimensionality of vector embeddings: the distributional hypothesis, the pairwise inner product loss and its bias-variance trade-off," *arXiv preprint arXiv:1803.00502*, 2018.

[22] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep learning based text classification: A comprehensive review," *arXiv preprint arXiv:2004.03705*, 2020.

[23] K. Krippendorff, *Content analysis: An introduction to its methodology*. Sage publications, 2018.

[24] C. Quan and F. Ren, "Unsupervised product feature extraction for feature-oriented opinion determination," *Information Sciences*, vol. 272, pp. 16–28, 2014.

[25] A. Sorokina, A. Karipbayeva, and Z. Assylbekov, "Low-rank approximation of matrices for pmi-based word embeddings," *arXiv preprint arXiv:1909.09855*, 2019.

[26] O. Levy and Y. Goldberg, "Neural word embedding as implicit matrix factorization," in *Advances in neural information processing systems*, 2014, pp. 2177–2185.

[27] J. Manotumruksa, C. Macdonald, and I. Ounis, "Matrix factorisation with word embeddings for rating prediction on location-based social networks," in *European Conference on Information Retrieval*. Springer, 2017, pp. 647–654.

[28] Y. Li, L. Xu, F. Tian, L. Jiang, X. Zhong, and E. Chen, "Word embedding revisited: A new representation learning and explicit matrix factorization perspective," in *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[29] O. Levy, Y. Goldberg, and I. Dagan, "Improving distributional similarity with lessons learned from word embeddings," *Transactions of the Association for Computational Linguistics*, vol. 3, pp. 211–225, 2015.

[30] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543.
[31] G. W. Stewart, "Matrix perturbation theory," 1990.
[32] H. Li, "Aligning sequence reads, clone sequences and assembly contigs with bwa-mem," *arXiv preprint arXiv:1303.3997*, 2013.
[33] P. L. De Jager, Y. Ma, C. McCabe, J. Xu, B. N. Vardarajan, D. Felsky, H.-U. Klein, C. C. White, M. A. Peters, B. Lodgson *et al.*, "A multi-omic atlas of the human frontal cortex for aging and alzheimer's disease research," *Scientific data*, vol. 5, p. 180142, 2018.

Language in Court: Ideology, Power and Cognition

Mehdi Damaliamiri

***Abstract*—** Undoubtedly, the power of language is hardly a new topic; indeed, the persuasive power of language accompanied by ideology has long been recognized in different aspects of life. The two and a half thousand-year-old Bisitun inscriptions in Iran, proclaiming the victories of the Persian King, Darius, are considered by some historians to have been an early example of the use of propaganda. Added to this, the modern age is the true cradle of fully-fledged ideologies and the ongoing process of centrifugal ideologization. The most visible work on ideology today within the field of linguistics is "Critical Discourse Analysis" (CDA). The focus of CDA is on "uncovering injustice, inequality, taking sides with the powerless and suppressed" and making "mechanisms of manipulation, discrimination, demagogy, and propaganda explicit and transparent." possible way of relating language to ideology is to propose that ideology and language are inextricably intertwined. From this perspective, language is always ideological, and ideology depends on the language. All language use involves ideology, and so ideology is ubiquitous – in our everyday encounters, as much as in the business of the struggle for power within and between the nation-states and social statuses. At the same time, ideology requires language. Its key characteristics – its power and pervasiveness, its mechanisms for continuity and for change – all come out of the inner organization of language. The two phenomena are homologous: they share the same evolutionary trajectory. To get a more robust portrait of the power and ideology, we need to examine its potential place in the structure, and consider how such structures pattern in terms of the functional elements which organize meanings in the clause. This is based on the belief that all grammatical, including syntactic, knowledge is stored mentally as constructions have become immensely popular. When the structure of the clause is taken into account, the power and ideology have a preference for Complement over Subject and Adjunct. The subject is a central interpersonal element in discourse: it is one of two elements that form the central interactive nub of a proposition. Conceptually, there are countless ways of construing a given event and linguistically, a variety of grammatical devices that are usually available as alternate means of coding a given conception, such as political crime and corruption. In the theory of construal, then, which, like transitivity in Halliday, makes options available, Cognitive Linguistics can offer a cognitive account of ideology in language, where ideology is made possible by the choices a language allows for representing the same material situation in different ways. The possibility of promoting alternative construals of the same reality means that any particular choice in representation is always ideologically constrained or motivated and indicates the perspective and interests of the text-producer.

***Keywords*—** power, ideology, court, discourse.

Mehdi Damaliamiri Damaliamiri is with the Independent Researcher, Iran (e-mail: damaliamiri@yahoo.com).

Sarcasm Detection in Speech Using Transfer Learning

X. Gao, S. Nayak

***Abstract*—**Sarcasm is one of the most commonly used figurative languages in daily conversations. We can perceive and convey it through audible cues, e.g., pitch, intensity, speed of speaking, etc. Related cues have been explored extensively in the literature, however, these cues have not been investigated completely in automatic sarcasm detection in speech. In this paper, a methodology of implementing pre-trained Convolutional Neural Networks (CNNs) is used to detect sarcasm in speech. The multimodal dataset MUStARD is used as a target dataset in this study. Precisely, two selected CNNs pre-trained models are Xception and VGGish. They are respectively trained on images and sounds datasets. The results of this study show that Xception which has its convolutional layers and pooling layers retrained performs better than VGGish which is applied as a feature extractor in the experiment. Both of the models achieve higher F-score compared to the baseline Support Vector Machines (SVM) model by 20% and 11% in unimodal sarcasm detection in speech.

***Keywords*—**sarcasm detection, CNNs, transfer learning, speech recognition, human-computer interaction

X. Gao is with the University of Groningen – Campus Fryslan, Wirdumerdijk 34, 8911 CE Leeuwarden (phone: +31 685216556; e-mail: xiyuan.gao@rug.nl).

S. Nayak is with the University of Groningen – Campus Fryslan, Wirdumerdijk 34, 8911 CE Leeuwarden (e-mail: s.nayak@rug.nl).

Language, Identity and Culture: A Study of Language Maintenance and Shift Among Amabhaca Raised in Umzimkhulu, Kwazulu Natal

Yanga Lusanda Praiseworth Majola

***Abstract*—** Umzimkhulu is a small town in KwaZulu-Natal (KZN), previously Bantustan of Transkei. The people in this town speak isiBhaca, a language which is closely related to three Nguni languages (viz. IsiXhosa, IsiZulu and siSwati), but has no official status in South Africa. IsiZulu and isiXhosa are used for official purposes such as education, health, religious and government departments even though the people of Umzimkhulu regard themselves as speakers of isiBhaca and this affects isiBhaca speakers at many levels, the first one being the confusion of having two languages used in different sectors and secondly, having two powerful languages which seem to overpower isiBhaca and this affects the identity of isiBhaca speakers. The primary objectives of the study were to gain an understanding of the importance of language to the cultural identity of amaBhaca and to also discover if isiBhaca is or can be maintained of there is a possible shift towards isiZulu or isiXhosa and to establish the attitudes held by isiBhaca First Language (L1) speaking citizens of Umzimkhulu towards isiBhaca. The study employed a mixed-methods approach. It is useful to consider the full range of responsibilities for data gathering in any study and to organise these methods by their degree of predetermined nature, their use of closed-ended questioning and their focus for numeric versus non-numeric data analysis. Data were gathered from key informants identified as speakers of isiBhaca who were born and bred in Umzimkhulu. The study found that much as isiBhaca is regarded as a dialect of isiXhosa it is linguistically closer to isiZulu than it is to isiXhosa, even though most respondents felt that isiBhaca is a language of its own and the continuous use and empowerment of isiZulu and isiXhosa in Umzimkhulu particularly in the professional settings is detrimental to isiBhaca and this subsequently has the potential of endangering isiBhaca existence and might lead to its attrition. The study further revealed that isiBhaca was mostly used for traditional or informal situations in Umzimkhulu and its speakers mostly had to accommodate speakers of isiZulu and isiXhosa in most cases as these are the standard languages around it, although they were fearful that their cultural identity was in trouble.

***Keywords*—** language, identity and culture, language maintenance and shift, language and dialects.

Yanga L. P. Majola is with the Tshwane University of Technology, South Africa (e-mail: MajolaYLP@tut.ac.za).

Investigating The Onomastic Principles of Naming an Official Language: The Case of the Sepedi and Northern Sotho Language Names

Tebogo Jacob Rakgogo

Tshwane University of Technology

Soshanguve South Campus

Faculty of Humanities

Department of Applied Languages

Pretoria, South Africa, 0182

E-mail address: Rakgogotj@tut.ac.za

**Corresponding author email: rakgogotj@tut.ac.za*

Abstract

The main focus of the study embraced the investigation the onomastic principles and also the socio-cultural factors to be taken into consideration when naming an official language. The fundamental objective of the study was to consider the two names *Sepedi and Sesotho sa Leboa*, to decide which one should be used and to determine which name is supported by the investigated onomastic principles of naming a language (official). A mixed-methods research approach, including both quantitative and qualitative methodologies was employed in the study. Survey questionnaires, interviews, observations and text analysis were mainly considered as methods of data collection. Research was conducted involving a group of Higher Education students and lecturers; employees at the Sesotho sa Leboa National Lexicography Unit and at the Sesotho sa Leboa National Language Body; employees at the Pan South African Language Board; employees at the National Department of Arts and Culture and its sub-departments in Limpopo and Gauteng Provinces. The findings of the study revealed that neither the language names Sepedi nor Sesotho sa Leboa is supported by the onomastic principles of naming a language. An overwhelming majority of the participants in this study opined that the language under onomastic scrutiny should be reconsidered and replaced with a neutral and inclusive name that cannot in any way be associated with the influence of power and politics as well as with ethnic and dialectal status of a particular speech community. In summary, it was found in this study that the language under scrutiny was not properly named in the first place, since both Sepedi and Sesotho sa Leboa were found to have strong ties with colonialism as concomitant part of the Apartheid regime.

Keywords: Sociolinguistics, onomastics, dialect(s), standard language(s), Sepedi, Northern Sotho, Sesotho sa Leboa, ethnic pride, colonialism, political power and influence.

Hiding the Word to Learn its Meaning Distribution

Eduardo A. M. Freitas

***Abstract*—** In Natural Language Processing, it is common to represent a word as a fixed vector or a contextualized vector in a sentence. But a single multidimensional vector can be insufficient to explain the complexity of word relations. It is easy to find a list of word meanings in a dictionary. But semantic cognition is about perception, memory, and meaning relations. The proposal is to create a meanings distribution hiding, fully or partially, the word under study. Before encoding the words, the approach starts replacing all instances of the understudy word with a special and unique token. Instead of many meanings of a word, the training uses many unique tokens of a meaning. Thus, the learning process becomes heavily dependent on neighbor words relationships. Inflections of the word under study and subwords can enrich the training in general. The algorithm looks for the best number of unique tokens clusters corresponding to the number of meanings. Gaussian models and fuzzy clustering can predict mixtures or cluster intersections, respectively. At the end of the learning process, each token has its location and part of the line context stored, making it feasible to validate the results. Then, the results confirm the effectiveness of meaning distribution to improve semantic cognition study.

***Keywords*—** Pattern recognition, Word representation, Polysemy, Clustering analysis, Semantic cognition, Language complexity.

Eduardo Antonio Mello Freitas is with the ISCTE-University Institute of Lisbon, Portugal (e-mail: eamfs@iscte-iul.pt).

The Spatiality of Gender Equality and Women's Leadership Positions in New Uzbekistan

Juraev Zuhriddin* • Young-Jin Ahn**

* Ph.D. student, Department of Geography, Chonnam National University
** Professor, Department of Geography, Chonnam National University

Contents

Abstract

This study is to contribute to the number of scientific works in social geography. Under the term "New Uzbekistan", we will examine the impact of reforms on gender equality based on spatial analysis. Since this is the first time our study has been conducted, so, we first covered the theoretical background in more detail. We have made the field of feminism and gender studies scientifically linked to our research. We analyzed the impact of Gender equality and feministic approaches on the issues of women's leadership positions. The purpose of this paper is to analyze the work that needs to be done in the context of gender studies and to identify the research that needs to be done in the future. In addition, there are limited sources on feminist geography and gender study works in Central Asia, particularly in Uzbekistan. We then examined the proportion of women in the workplace in the new Uzbekistan and the spatiality of leadership positions for women. In the concluding section, we have provided our comments and suggestions.

Keywords: gender studies, women leadership positions, spatiality, feminist geography, gender equality

1. Introduction

There are many criteria in the world that reflect the prestige and cultural level of the state and society. But there is a criterion that clearly reflects the spiritual maturity of any nation and people, which is determined by the attitude of society to women. In the "New Uzbekistan"[1], the implementation of essential reforms, the protection of human rights and interests in the process of building a strong civil society, and the rule of law; ensuring equal rights and opportunities for men

[1] Note: The term "New Uzbekistan" was introduced by the current President. The reason for this is the policy of reviving the work of scientists-scholars from Uzbekistan, who founded the First and Second Renaissance periods in Uzbek national history.

and women have become relevant issues. However, despite the progress made in recent years, there are still old stereotypes about women. It still "contributes" to the ceiling of women's rights in society. Fair gender interpretation is one of the professional and ethical values such as clarity, impartiality, and fairness. In fact, respect for women is an ancient value of the Uzbek nation. In the distant and recent past, some famous women introduced the voice of the Uzbek Nation to the World. Although the issue of women in Uzbekistan has risen to the level of public policy, various social or cultural stereotypes and/or barriers still hinder women from achieving equality in the workplace and society, especially in leadership positions. Because the equal participation of men and women in the system of governance is one of the bright signs of the modernization of society. As an independent subject of international relations, Uzbekistan is still practically covering in fulfilling its obligations under major international human rights treaties. You can see it in our analysis. In this article, we examine the spatial dimensions of gender relations and gender equality in Uzbekistan. Under the term "New Uzbekistan", we will examine the impact of reforms on gender equality based on spatial analysis. Our purpose is to analyze the work that needs to be done in the context of gender studies through this research, to identify the research that needs to be done in the future and the research that will be effective. In addition, there are lacking of sources on feminist geography and gender studies in Central Asia, particularly in Uzbekistan.

In this study, the spatiality of women's status in workplaces, especially their advancement in leadership positions will be examined. Statistical data analysis is provided. In particular, the extent to which gender equality was ensured in the sectors is examined. The data were taken from the website of the National Statistics Committee of Uzbekistan, the information texts of the Ministry of Labor and Employment, and the press service of regional governments. Data analysis showed that the position of women in the workplace (sectoral) is still lower in quantity than that of men. For the data, in some statistics, although the numbers are given clearly, which areas are not clearly stated. In some cases, the field (sector) is clearly stated, and the percentages are given in general terms instead of clear numbers. Such shortcomings have made this study difficult to accurately examine the spatiality of women's occupations.

This study is to contribute to the number of scientific works in this field. Since this is the first time our study has been conducted, so, we first covered the theoretical background in more detail. We have made the field of feminism and gender studies scientifically linked to our research. We analyzed the impact of Gender equality and feministic approaches on the issues of women's leadership positions. We also mentioned the historical process to get to the topic. We then examined the proportion of women in the workplace in New Uzbekistan and the spatiality of leadership positions for women. In the concluding section, we have provided our comments and suggestions.

2. Theoretical framework

2.1. Feminism and Gender studies

For much of the last century, the liberal, materialist, and radical approaches inherent in feminism have been almost completely replaced by a new cross-cutting approach. Many axes of marginalized identity exist under the noses of former feminists, all new to read through a lens that shows potential oppression, bigotry, injustice, and grievances - a partnership in a system of power and privilege became a theory of need. It does this by calling on different oppressed tribes to support each other: first under the banner of "allyship" and then "solidarity" - both of which are perceived as problematic theories in "centralizing" the needs of privileged "allies" at the expense of oppressed minority groups of ever-increasing specificity (Stevi Jackson 2001)[2]. These changes redirected many feminist scholarships and activism, as surprisingly, activists adopted a new, "increasingly complex" approach called "intersection," which combined many forms of identity theory[3]. This change was so rapid and profound that by the early 2000s, many scientific articles had emerged emphasizing the need for materialist and radical approaches to feminist theory (Helen Pluckrose et al., 2020). Among the main texts of feminism are Simone de Beauvoir's book "The Second Sex" (1949)[4], an 'innovative book' that emphasized that women were created based on cultural notions of inferiority to men, and Betty Frieda's "The Feminine Mystique" (1963)[5] stated that women were done with housework and motherhood. Kate Millet's "Sexual Politics" (1978) cited familiar issues. All these texts correspond to radical feminism, where femininity is culturally constructed and loaded by men (in top-down power dynamics) and advocates the revolutionary overthrow of patriarchy (James Lindsay et al., 2020). Besides that, Judith Lorber's essay "Shifting Paradigms and Challenging Categories" (2006) portrays how Marxist feminism saw women as a class[6]. Looking at this issue, in the early 1970s and 1980s, against the backdrop of eliminating workplace inequality, Marxist feminists "worked hard" to show that the exploitation of housewives is an integral part of the capitalist economy. The 'reform' of the Soviet Union in Central Asia, especially Uzbekistan, can prove this (Marxism-Leninism)[7]. According to Lorber's essay, this feminist perspective presents 'metanarratives' about men, women, and society based on a simple oppressive male/female duality (Gallas, 2010)[8]. Perhaps, from a postmodern point of view, "these changes are important" because the individual positions and "women's" positions that connect them differ greatly in time and space. Therefore, in the early 2000s, the predominant view of feminism was that talking about "women" and "men" was quite inconsistent (Helen Pluckrose et al., 2020). Overall, you have several perspectives that emphasize the traces of a conceptual shift from feminism to gender studies. Throughout the period of postmodernism applied, the unification

[2] Stevi Jackson, 2001, "Why a Materialist Feminism Is (Still) Possible—and Necessary," Women's Studies International Forum 24, no.3–4.
[3] Helen Pluckrose et al., 2020, Cynical Theories., Book, pp. 139-147., USA, https://lccn.loc.gov/2019054122
[4] The Second Sex - Wikipedia
[5] The Feminine Mystique | Betty Friedan, Kirsten Fermaglich, Lisa Fine | W. W. Norton & Company (wwnorton.com)
[6] Judith Lorber, 2006, 'Shifting Paradigms and Challenging Categories' Shifting Paradigms and Challenging Categories on JSTOR volume 53, No.4, p.462.
[7] Marxism–Leninism - Wikipedia
[8] Alexander Gallas, 2010, Dichotomy, Dualism, Duality: An investigation into Marxist conceptualizations of structure and agency., Academia.edu p.86.

of various minority status groups under the single flag of oppression was regarded as the only "right" way to practice feminism. Feminism, on the other hand, has been superseded by gender studies under queer theory and has embraced cross-reactivity as a kind of grand unified theory of social power and social injustice (Cynical Theories, 2020). When it comes to Central Asia, after the post-colonial, democracies of the region's countries, the notion of civil society, along with human rights rhetoric, is much more ambiguous. That is, feminist approaches in the region have done little work. However, gender equality was preserved in the constitutions with the independence of the Central Asian Republics. But they accepted the consent of the UN Convention on the Elimination of All Forms of Discrimination against Women, the Beijing Declaration, and the Platform for Action (1990-2000). Thus, feminism is not openly prevalent in the region. Gender studies, on the other hand, have become more visible in recent years (UNDP, UN Women, USAID, Human rights watch, etc.).

2.2. Feminist theories and Gender equality

Feminism raised terms called "men and masculinity" while fighting to eliminate oppression. Almost all feminist theorists highlighted this term to examine patriarchy matters in society. While analyzing the role of women in leadership positions and their spatiality in this article, we believe that some scientific support is needed for views on the term "men and masculinity". Feminists and gender study theorists and researchers of this field claim that there are "need-to-solve" problems with this terminology when women are unable to take leadership positions. So, the study of men and masculinity often relies comprehensively on the concept of hegemonic masculinity[9]. Because hegemonic masculinity is aggressive and competitive masculinity that maintains a man's dominance over women and is different from the hegemonic (dominant and powerful) discourse about what it means to be "genuine" (Cynical Theories, 2020). "Men and masculinity scholars" are primarily male, but they study masculinity within the framework of feminists (James Lindsay, 2020). The subject of women's leadership positions in Central Asia is no different from the challenges the world community faced. We had female leaders in history, and we have them in our time, but the scale of gender matters has become more apparent – prompting speculation about the need to study and find solutions to the problem. It was at a time when feminists denounced patriarchy that the issue also showed its grandeur. In one study of women in Central Asia, *Aculai (et al., 2009)* stated the high proportion of women owning property in the CIS (including Uzbekistan)[10]. *Mujahed and Atan (2017)* discussed the issue of how women face restrictions in getting administrative positions. Cultural attitudes, gender stereotypes, work, family conflicts, and national traditions formed by Islamic teachings are highlighted as barriers, however, we have a disagreement with this last part. Also, in Central Asia, specifically in Uzbekistan, those mentioned barriers sound familiar as some of the limitations to women's advancement as a leader. But in

[9] Note: It has been developed by Australian gender theorist Raywin Connell. See at http://www.raewynconnell.net/p/masculinities_20.html

[10] Aculai, E., A. Bulgac, V. Gryga, O. Krasovksa, and O. Linchevskaya (2009) "Researching the Gender Aspect in Business Development (the Case of Moldova, Ukraine, and Belarus)." Lex ET Scientia International Journal 16(1).

Uzbekistan, social stereotypes can stand out greater than cultural and/or religious prejudice. A related study belongs to *Alice H.Eagly, et al., (2003),* who emphasized that "while women have some advantages in a typical leadership style, they suffer some shortcomings from a prejudiced assessment of leadership abilities, especially in the context of male organizations". Nevertheless, more and more women are being promoted to leadership positions at all levels, including elite leadership positions. So, we can see that women are more likely than men to demonstrate leadership in an effective style in today's situation[11]. So, categories will be needed in the spatial study of the topic of women's leadership. These include the categories of family, work, society, and regional tradition, as well as religious views. However, the fact that women also have child-raising responsibilities (Guarcello et al. 2005)[12] may not allow them to fully function as leaders. Although some feminists have argued that fathers and mothers are equally responsible for the upbringing of children (McCarthy, 2006)[13], in terms of national customs and values, the work of raising children is more important and higher position than the leadership for women. However, this high status has been considered as an honor (Hamza Yusuf et al., 2003)[14].

2.3. Theoretical background

Although Uzbekistan, as a young state, has boarded on extensive reforms to build a free civil society (after 1992), significant results in gender equality have been great during the "New" Uzbekistan (since 2017). The number of women leaders is growing in politics, education, economics, and business fields. However, access to credit is one of the factors that may hinder the establishment and growth of women`s businesses (Welter and Smallbone 2010)[15], which is a problem in Uzbekistan too. This is because women are more prone to risk than men and may contribute to gender disparities in business-related borrowing (Sattar 2012)[16]. A spatial analysis of "the prevalence of gender norms in education and economic growth" in the new Uzbekistan "at the national level" (Rani and Bonu, 2009)[17] is included in this topic. This is because gender studies and geography research in an integrated way are not enough (or do not exist). The spatiality of gender relations in the Central Asian region was explored by *Stefan Shutte (2014)* under the theme

[11] Alice H.Eagly, and Linda L.Carli, 2003, "The female leadership advantage: An evaluation of the evidence" Elsevier, Leadership Quarterly, volume 14, issue 6, pp., 807-834., https://doi.org/10.1016/j.leaqua.2003.09.004
[12] Lorenzo Guarcello, Fabrizia Mealli, and Furio Camillo Rosati, (2010), "Household Vulnerability and Child Labor: The Effect of Shocks, Credit Rationing and Insurance" January edition, Journal of Population Economics 23(1): pp. 169-198 DOI:10.2139/ssrn.453900 Source-Re-Pec.
[13] McCarthy, Cormac (2006), The Road, New York: Vintage. [Google Scholar]
[14] Hamza Yusuf Hanson, John Taylor Gatto, Nabila Hanson and Dorothy Sayers (2003) Educating Your Child in Modern Times: How to Raise an Intelligent, Sovereign & Ethical Human Being Paperback.
[15] Friederike Welter, David Smallbone (2010) Institutional Perspectives on Entrepreneurial Behavior in Challenging Environments., https://doi.org/10.1111/j.1540-627X.2010.00317
[16] Sattar, Sarosh (2012) Addressing the Gender Gap in Europe and Central Asia. Europe and Central Asia knowledge brief; issue no. 54. World Bank, Washington, DC. © World Bank., https://openknowledge.worldbank.org/handle/10986/17102 License: CC BY 3.0 IGO."
[17] Manju Rani, Sekhar Bonu (2008) Attitudes Toward Wife Beating: A Cross-Country Study in Asia., https://doi.org/10.1177/0886260508322182

of how the women's agencies contribute to household life (women in Afghan urban areas)[18]. But it also did not mention the position of women in leadership. *Jiekuan Zhang (et al., 2020)* examined the relationship between tourism and gender equality in CIS countries (including Uzbekistan)[19]. In our study on the spatiality of women's leadership, certainly, we rely on this source in terms of the importance of these categories. This is because gender equality in the economy, education, and employment sectors can vary significantly. The examination of spatiality also includes medicine. On this topic, *Kotsenas Amy (2017)* studied gender inequality in leadership positions in this field of medicine by comparing the number of women entering the medical field with the number of women working in leading roles there[20]. However, in the case of New Uzbekistan, there is no analysis regarding the spatiality of women's leadership positions in this sector. *Derks, Ellemers, et al., (2012)* focused on research on women in leadership positions. According to them, as a group, women may adopt different leadership styles than men. Women are also more likely to develop skills such as competitiveness and toughness[21]. In addition, (according to the authors) there are issues in this regard that can affect the career opportunities of women as individuals and/or groups. *Buvinić et al. (2013)* focused on the fact that women face more restrictions than those who move to leadership positions. Therefore, models and scientific development are needed to provide a clear starting point for developing strategies to enhance the effectiveness of women in leadership positions. We believe that the spatiality of the position of women in the leadership position of our topic will also help resolve this problem.

3. New Uzbekistan and Gender Equality

3.1. Historical approach – Female leadership is in our bloodline

Respect for women has long been valued in Uzbekistan. Even female leaders have stood out in history. One of them is a brave woman – the army general named Tamaris. She later ruled the country after her husband's death (Herodotus, The Histories)[22]. Timur and the Timurid princes ruled Central Asia, Northern Afghanistan, partly Iran, and Iraq from the second half of the 14th century to the first quarter of the 16th century. Like princes, princesses of the Timurid dynasty also produced talented politicians, motivational speakers, mature scholars, and sensitive poets[23]. For instance, the wife of the great emperor Amir Temur (1336-1405) – Sarai Mulk (1341-1408), King Hussain Baykaro's (1438-1506) wife – Khadija (1451-1511), wife of Shah Jahan (1592— 1666) -

[18] Stefan Schutte (2014) "Living with patriarchy and poverty: women's agency and the spatialities of gender relations in Afghanistan" October 2014 Gender Place and Culture A Journal of Feminist Geography 21(9) DOI:10.1080/0966369X.2013.832661

[19] Jiekuan Zhang et al., (2020) Tourism and gender equality: An Asian perspective https://doi.org/10.1016/j.annals.2020.103067 Annals of Tourism Research Volume 85, November 2020, 103067

[20] Amy L. Kotsenas 2017, "Women in Leadership: Why So Few and What to Do About It" Journal of the American College of Radiology, Volume 14, Issue 4, April 2017, Pages 450.

[21] Naomi Ellemers, Floor Rink, Belle Derks, Michelle K. Ryan (2012) Women in high places: When and why promoting women into top positions can harm them individually or as a group (and how to prevent this) Journal name: Research in Organizational Behavior., Elsevier., Volume 32, Pages 163-187. https://doi.org/10.1016/j.riob.2012.10.003

[22] https://en.wikipedia.org/wiki/Histories_(Herodotus)

[23] https://en.wikipedia.org/wiki/Timurid_Empire

Mumtaz Mahal (1594— 1631), daughter of Aurangzeb Alamgir (1618 1707) - Zebunnisa (1639-1702), etc.[24] In Soviet Uzbekistan (1980-1990) there were women leaders. For example, women have been appointed district governors in Andijan and Samarkand provinces. But one woman has been appointed as the province's mayor (current Khorezm region). She was the only woman in the USSR to rule the province at that time (1988)[25]. After the disintegration of the Soviet Union in 1991, it immediately came into conflict with the elements of Soviet heritage that ensured women's equality. The post-independence government of Uzbekistan under the first President continued to promote women's equality. Uzbekistan has enshrined gender equality in its constitution and other laws and has taken administrative measures to enhance the status of women. According to the decree, the chairwoman of the Republican Women's Committee is the Deputy Prime Minister[26], and the representatives of the committee at the regional, district and city levels serve as deputy of governors. In turn, the activities of women leaders in the religious domain are officially legalized in Uzbekistan. Such women are called "otin" because this position has existed since Islam came to the region[27]. In the selection of such women, attention is drawn to the fact that they have a reputation among their community and are advanced in both secular and religious knowledge.

3.2. Politics and reforms on gender equality

Until 2017, gender equality was one of the "prohibited topics" in Uzbekistan (Telegraph web press), the subject has been blacklisted for many years. Recently it has been released. On the eve of International Women's Day (March 8), a special commission on gender equality was established (2019). Two important laws have entered into force, the Guarantee of Equal Rights and Opportunities for Women and Men and the Protection of Women from Oppression and Violence (Lex-uz). Structural and institutional changes are underway, and legislation is improving ("Yuksalish" Nationwide action Press service, 2021). Besides, the State Award for Gender Equality was announced in the fields (2021). Several organizations have already won the first award (Media Press releases). For the first time in the history of Uzbekistan, a woman leader has been appointed Chairman of the Senate "Oliy Majlis"of the Republic. This means she is the second most senior leader in the country. To increase women's representation in parliament, a 30 percent quota has been set for female candidates for the Legislative Chamber. However, in Uzbekistan, quotas for women were legalized at 30% of the total number of candidates nominated by political parties (1995). Also, for the first time in the history of Uzbekistan, a woman was able to put her name as a candidate-elect for the presidency of the Republic of Uzbekistan (2021). The "Women's register book" system was introduced (2020). The Women's Register book system lists 900,000 women in need of socio-economic, medical, legal, and psychological assistance. There are currently more than 630,000 recorded women in the register book, 200,000 of whom are

[24] Bartold, Vasily (2020). Turks. 12 Lectures on the History of the Turks of Central Asia. Moscow: Yurayt Publishing house. p. 181.
[25] https://en.wikipedia.org/wiki/Uzbek_Soviet_Socialist_Republic
[26] https://lex.uz/ru/docs/-5884143
[27] http://old.muslim.uz/index.php/rukn/maktub

unemployed. Despite these measures, there are still cases of crime among women or domestic violence and harassment that women face.

3.3. The spatiality of gender (equality)

The population of Uzbekistan is over 35 million people. According to the State Statistics Committee, 17,5 million of them are men and 17,4 million are women[28]. The territory of Uzbekistan includes 12 regions, the Republic of Karakalpakstan and the city of Tashkent (figure 2), and they are divided into districts and cities subordinate to the region (figure 1).

Figure 1. Administrative division: Population density according to the administrative structure (2022)

Province▲	ISO 3166-2	Capital	Area (km²)	Population	Density (pers/km²)
Andijon	UZ-AN	Andijon	4 200	3 127 683	744,7
Buxoro	UZ-BU	Buxoro	39 400	1 923 934	48,8
Farg´ona	UZ-FA	Farg´ona	6 800	3 752 034	551,8
Jizzakh	UZ-JI	Jizzax	20 500	1 382 060	67,4
Namangan	UZ-NG	Namangan	7 900	2 810 843	355,8
Navoiy	UZ-NW	Navoiy	110 800	997 100	9,0
Qaraqalpaqstan Respublikasi	UZ-QR	Nukus	160 000	1 898 351	11,9
Qashqadaryo	UZ-QA	Qarshi	28 400	3 280 418	115,5
Samarqand	UZ-SA	Samarqand	16 400	3 877 355	236,4
Sirdaryo	UZ-SI	Guliston	5 100	846 260	165,9
Surxondaryo	UZ-SU	Termiz	20 800	2 629 135	126,4
Toshkent	UZ-TO	Toshkent	15 300	2 941 908	192,3
Toshkent City	-	-	327	2 571 668	7 864,4
Xorazm	UZ-XO	Urganch	6 300	1 866 493	296,3
Total			**442 227**	**33 905 242**	**76,7**

Source: Geo-ref-net, 2022 (photo credit)

Uzbekistan is one of the world's "youngest" countries in terms of aging. In Uzbekistan, the proportion of men and women of basic marriage age is almost equal. This equality continues in the reproductive age groups of women.

Figure 2. Map of population density by region - 2020

[28] https://stat.uz/uz/59-foydali-ma-lumotlar/5859-o-zbekiston-aholisi-3

Source: Geo-ref-net, 2020 (photo credit)

Thus, Uzbekistan has been pursuing a policy of gender equality since its independence. However, until 2017, this policy did not show some practical high results, but it shows significant changes in the period of "New" Uzbekistan. Although feminist views are not allowed into the region, there are reforms around women's rights. As for the direct promotion of women to leadership positions, the positivity is high during this period. According to the Women's Power Index 2020, women are currently running governments in 21 of the 193 UN Member Countries. The rating also reflects the level of women's ministerial positions in Uzbekistan. In 2020, the figure was 8 percent. (158th place). The level of women's representation in the Uzbek parliament is 29% (61st place). The United Nations estimates that the "required mass" of women in decision-making should be 30%[29]. The representation of women at this level was also established by the 1995 Beijing Action Platform (UNDP). The level of women's involvement in the country's social and political life can also be assessed by the number of registered female candidates in the elections to the lower and upper chambers of parliament. In Uzbekistan, the figure is 41% (14th rank)[30].

Table 1. Fields where a woman is a leader (2022)

Fields where woman is leader (first person)	
Sector name	**Exactly name of the field**
Political-governmental	Chairwoman of the Board of the "Oliy Majlis" – parliament of Uzbekistan
	Counsellor to the President of the Republic
	Chairwoman of the Women's Committee of Uzbekistan

[29] https://www.un.org/en/ga/about/background.shtml
[30] https://www.cfr.org/article/womens-power-index#chapter-title-0-2

	Minister of Preschool Education
Social and sports sector	Director of "Family Research" Center
	Head of Board of Directors of the "Healthy Generation" Foundation
Finance and business	Chief Executive Officer of "Aloka" Bank
	Director of the Tashkent International Arbitration Center
	Chairwoman of the Association of Entrepreneur Women of Uzbekistan
Art and culture	Director of the State Museum of Art of Uzbekistan
	Director of the State Museum of History of Uzbekistan
Education and science	Director of the Institute of Polymer Chemistry and Physics of the Academy of Sciences of Uzbekistan.
	Chairwoman of "Pharmaceutics" JSC
	Rector of the Tashkent State Institute of Oriental Studies

Source: "Yuksalish" Nationwide Center – 2022

According to official sources, from 2014 to 2018, the share of women in leadership positions in Uzbekistan was less than 1%. After 2020, the share of women in leadership positions will increase to 7% (Single portal of interactive services, 2021)[31]**.** However, there is no clear data in which fields this 7% exact figure is given (Table 1). We consider the distribution of working women in the sectors below. However, the lack of statistical analysis of the data is one of the points that need to be addressed within the topic.

Under current law (Lex-uz online, 2022)[32], women and men are treated equally in Uzbekistan. Currently, there is one woman deputy prime minister, one minister, and three deputy ministers in Uzbekistan (Media analytic report news-2021). But in practice, women are more likely to hold leadership positions in the social sphere. Today, there are 1,405 women leaders in the country (Governmental analytical reports, 2021). 9,309 neighborhood chairpersons and women deputy chairpersons have been promoted to "local women's activists" (local governments 'press service, 2021). There are also women leaders in the business and entrepreneurship sectors who are taking the initiative (Open data portal, 2021)[33]. For instance, "there are 36,295 small businesses and micro-firms under the leadership of women" (Davr web press)[34]. In the political sphere, in 2020-2022, out of 170 district governors and 25 mayors in Uzbekistan, only 6 women are holding a leadership position (Official report, media release)[35]. During the current policy (since 2017), female district governors appeared. The first female governor was appointed in Boyovut district of Syrdarya province (figure 3, red mark). Later, a woman was appointed governor of Alat district

[31] http://old.my.gov.uz/uz/authority/143 Note: Former chairwoman reported to the media.
[32] https://lex.uz/uz/docs/-5884084
[33] https://data.egov.uz/eng
[34] https://davr24.uz/archives/7926
[35] https://kun.uz/uz/news/2022

in Bukhara province, Bostan district in Andijan province, Karshi district in Kashkadarya province, Turakurgan district in Namangan province, and Kattakurgan district in the Samarkand province (figure 3). However, in the history of independent Uzbekistan, no women governor, ambassador, or head of law enforcement has been appointed. The number of women leaders in NGOs is high, but overall, they are low. (Family and Women's Committee, Official Written Information, 2022)[36].

3.3.1. The spatiality of women's positions at work

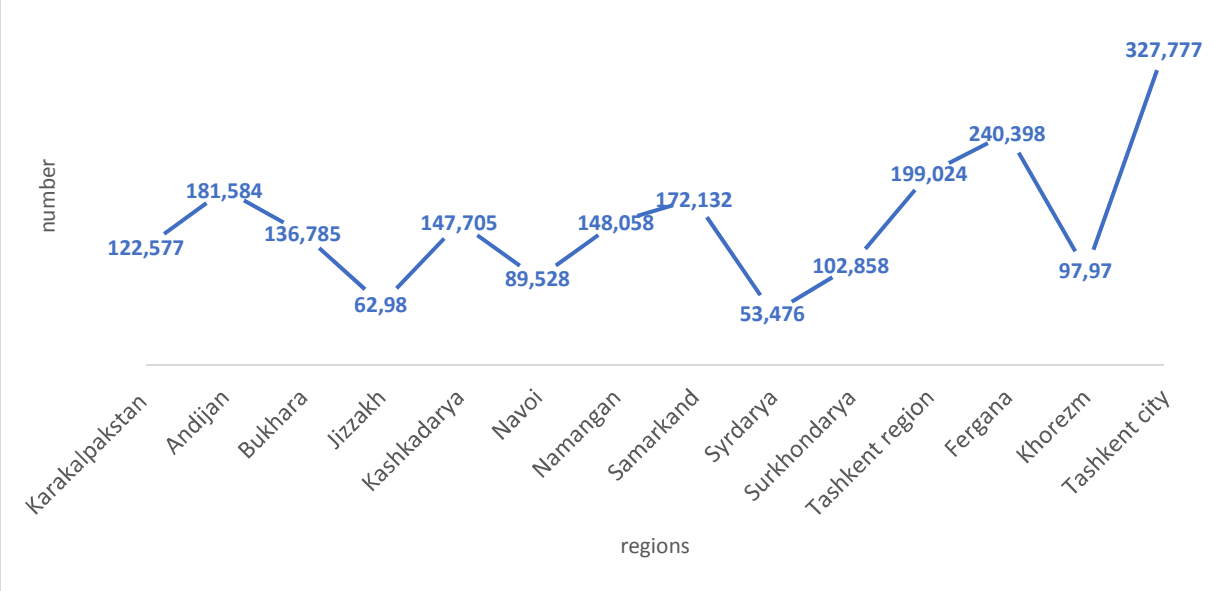
Figure 3. Female governors appointed locations

Made by the author of the article, through Google map

In general, according to official sources and officials, the number of women leaders in Uzbekistan is 26.6%. The highest percentage of women leaders is in Karakalpakstan (33.8%) and the lowest in Jizzakh (17%). Female members of parliament: deputies and senators exist at 29% (the norm is at least 30%). In other words, 48 women in the Legislative Chamber - 33%; In the Senate, 23 women are working 23.7% (Family and Women's Committee, Official Written Information, 2022)[37].

[36] https://www.gazeta.uz/news/2022
[37] https://www.gazeta.uz/news/2022

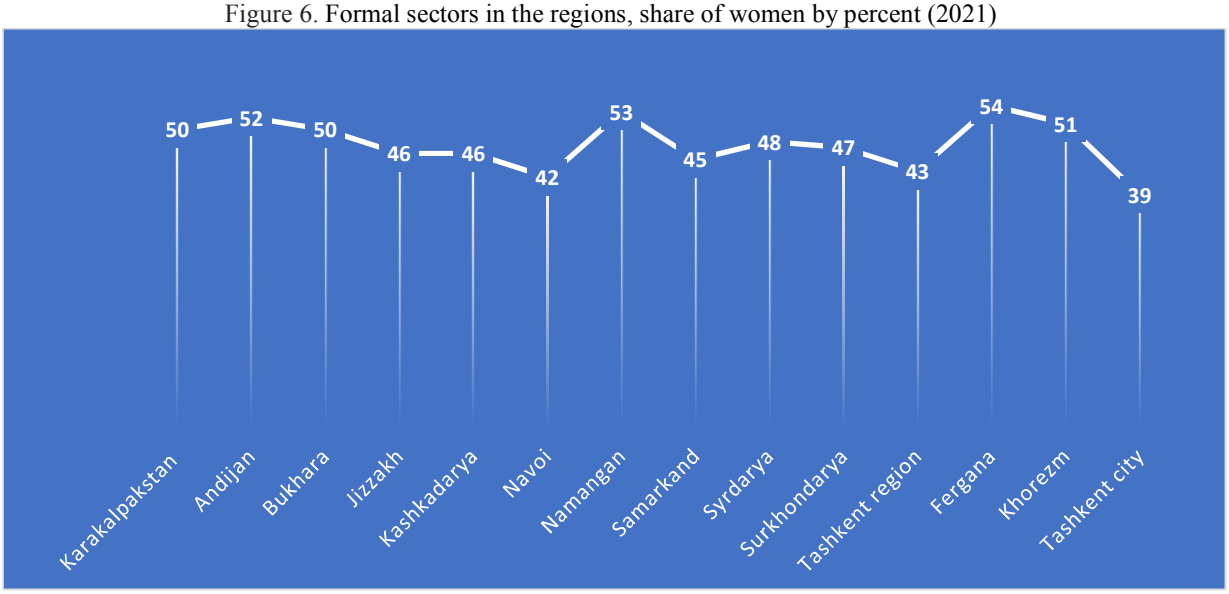
Figure 5: Number of women working in the formal sectors in the regions (2021)



Source: Unified National Labor System of Uzbekistan (2022)

Figure 5 shows that Tashkent (the capital city) has the highest number of women on the official state list. The lowest rates are in Syrdarya and Jizzakh provinces. The rest of the provinces show almost no significant differences. (See map for a clear view of the province, Figure 4). In terms of percentages, the opposite is true (see figure 6). In other words, Fergana, Namangan, and Khorezm provinces have the highest percentages, but Tashkent (capital city) has the lowest, and others have almost similar percentages. For a more accurate analysis of the distribution of women's work (or gender equality in workplaces), a comparison of the figures (figure 1) shows that in Fergana and Andijan provinces, gender equality in the workplace is (well) balanced. However, a somewhat negative distribution can be seen in the capital city. But this does not mean a clear conclusion. Factors and data need to be analyzed in more depth and close.

Figure 6. Formal sectors in the regions, share of women by percent (2021)



Source: Unified National Labor System of Uzbekistan (2022)

According to officials, the work on "decent employment of women" is under constant control. Gender equality is considered (Media press release). As evidenced by Table 2, the proportion of working women in both the formal and retirement age is high. Men have a slight advantage in working while studying. However, in a statement, the President of the Republic stressed that "the most difficult issues" reached with women and/or gender equality (Institute for strategic and regional studies under the President of the Republic of Uzbekistan)[38]. Hence, there is an incompatibility in the execution of reforms with such official data.

Table 2: Total formal workers in Uzbekistan 2021

Formal workers		
Categories	**Number**	**Percentage**
Total workers until retire age	**4473763**	**100**
Men	2390911	44
Women	2082852	46
Total workers with retired status	**327781**	**100**
Men	123887	37.8
Women	203894	62.2
Total part-time workers while study	**103347**	**100**
Men	57958	56.1
Women	45389	43.9

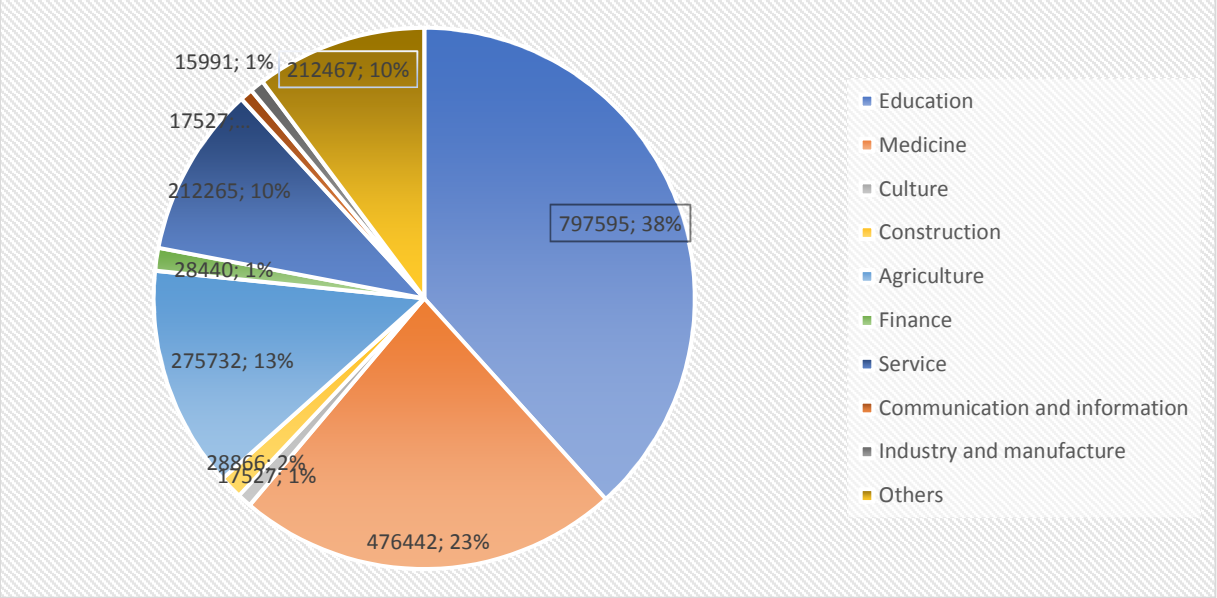
Source: Unified National Labor System of Uzbekistan, 2021

According to the National Statistics Committee of Uzbekistan, the share of women in the sectors is high in education (38%). The share of women in preschool education, secondary schools, and higher and secondary special education is also high worldwide. The figure shows that the lowest share of women is in the finance (1%) and information and communication (1%) sectors. It was

[38] http://isrs.uz/en/ozbekiston

not possible to find accurate data on the number of female leaders in the sectors. National statistical analysis does not provide spatial data for certain categories (figure 7).

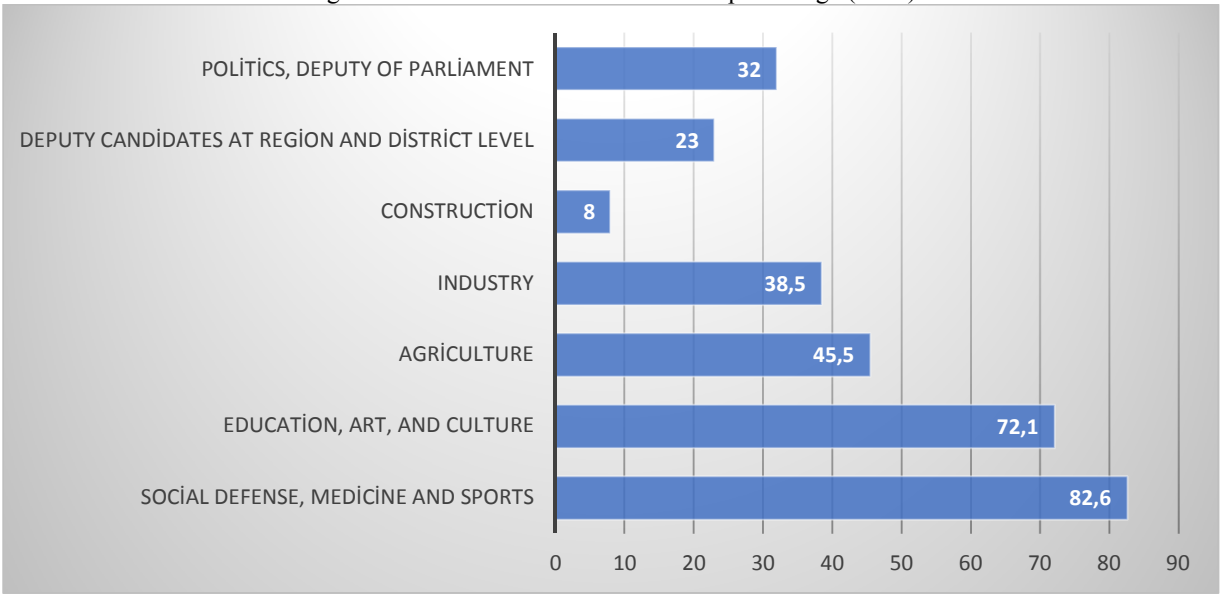
Figure 7: Sectors where women work



Source: National Statistics Committee website

According to official reports, 82.6% in Uzbekistan in the field of health, social protection, and sports, 72.1% in science, education, arts and culture, agriculture 45.5 percent, 38.5 percent in industry, and 8 percent in construction are women (figure 8). There are a total of 17 million women in the country, and even if the number of women with children and disabled persons is deducted, only 1,405 out of those 1 million women hold leadership positions. Besides, the share of women among the representative bodies and members of parliament is 32%. Thus, the analysis of the data on gender equality, in particular the data on the adequacy of women's leadership positions in the sectors, is not satisfactory. According to the analysis, in order to increase the participation of women in society, it is necessary to establish cooperation with various organizations, and political parties, train women socially active, improve their skills, and systematically recommend them for various leadership positions.

Figure 8. Sectors where women work in percentage (2021)



Source: Media release, official report during the interview. www.kun.uz

However, according to some official data, each region of the country has been studied, and a reserve of personnel (6,000 positions) for management positions has been formed (Academy of State and Social Construction under the President of Uzbekistan, 2022)[39].

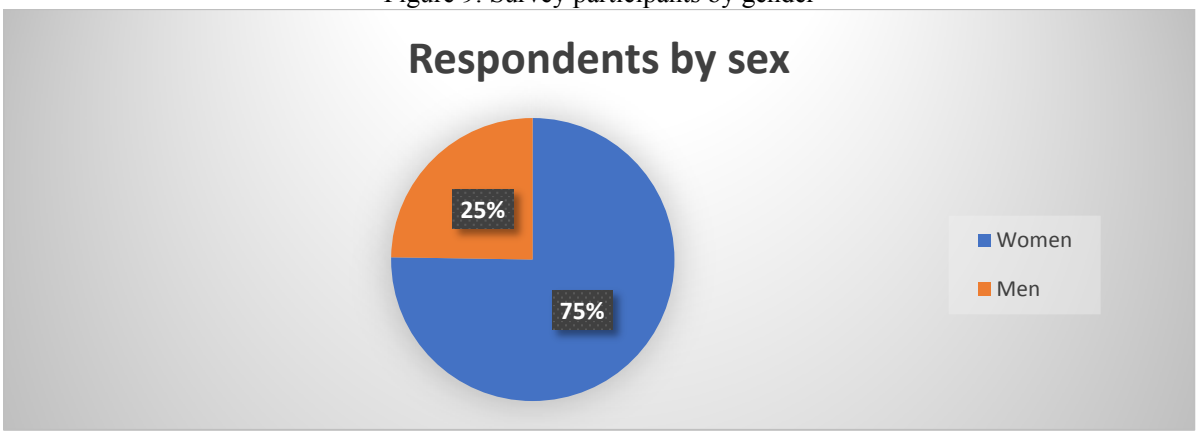
3.3.2. Obstacles to women's employment and promotion to leadership positions

1) Survey

I have an online survey[40] conducted by the Civil Service Development Agency[41] to study gender equality issues and their spatiality. The survey question is that "What prevents Uzbek women from advancing in their careers?" Thus, Uzbek women face certain difficulties in finding employment, mainly because of their household chores and low role in society. So, it was attended by 808 respondents. 74.5% of the participants are women and 25.5% are men (figure 9), of which 84.7% work in the public sector and 15.3% in the private sector (figure 10).

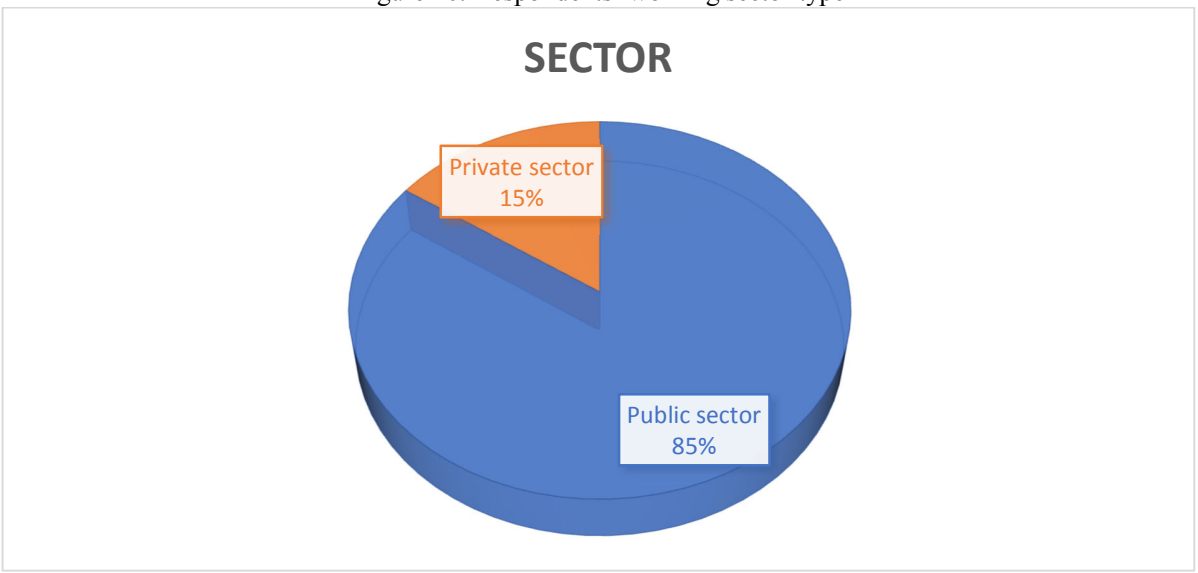
[39] https://dba.uz/admission?lang_is=set&lang_data=English
[40] It should be noted that the survey was presented in the form of written analytical information in the general context online. As the author, numbers, and data were selected from this informational text. The source of the survey is placed in the footnote as a link.
[41] https://argos.uz/en/services

Figure 9. Survey participants by gender



Made by the author of this study by the figures of survey

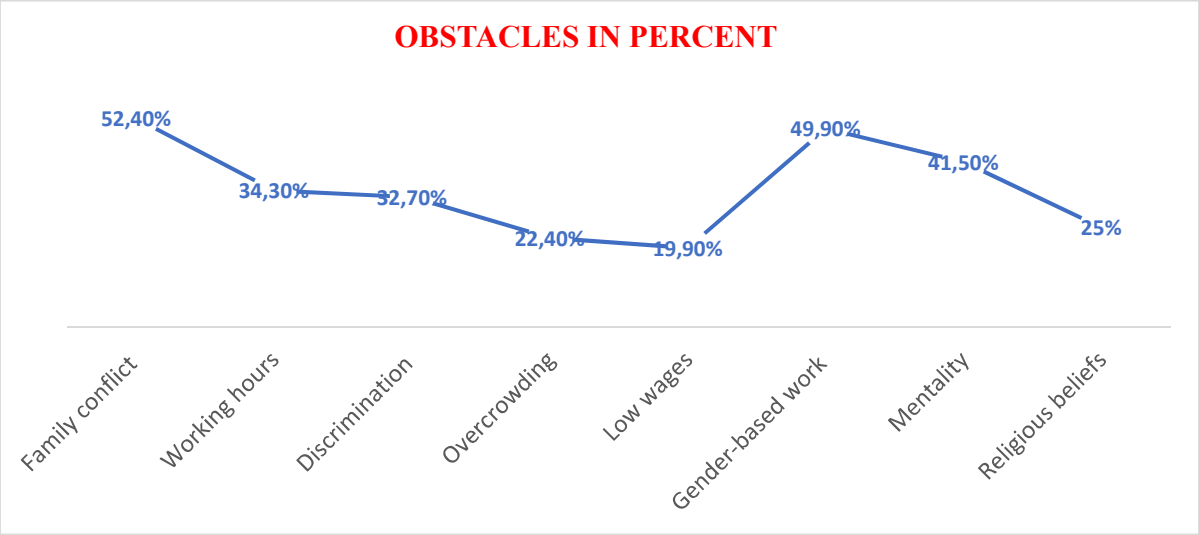
Figure 10. Respondents' working sector type



Made by the author of this study by the figures of survey

According to the survey, several obstacles were cited (figure 11). 52.4 percent of respondents replied that marital status was the reason why women did not enter the civil service. For example, stuck with child-caring, or spouses did not allow to work. In other words, they have family conflicts. 34.3 percent of participants said they were not satisfied with their working hours, while 32.7 percent said they had experienced gender discrimination in advancement toward a leadership position. At the same time, the options cited as overcrowding and low wages were 22.4 percent and 19.9 percent, respectively. Also, 49.9 percent of respondents among the problems of ensuring gender equality in the civil service divided occupations and positions into positions where it is considered appropriate to work for "men" and "women". 41.5 percent of respondents cited "national mentality" and 25 percent cited "religious beliefs".

Figure 11. Obstacles to women's activeness



Made by the author of this study by the figures of survey

2) Assumptions regarding the obstacles

Based on the above survey analysis and our data analysis, our research hypotheses are as follows. While the proportion of women in the workplace is almost equal, some concepts reflect a negative attitude towards women in administration. In other words, while Uzbekistan supports women's (political) leadership, this is expected to be slow in practice.

The first assumption is that "there is a negative attitude toward women in leadership positions". Assumptions from the survey analysis suggest that national culture and/or mentality can be assumed to be one of the main factors preventing women from advancing to the top of the career hierarchy. It can be said that women do not have support in organizational work. Because there is still a perception that women play a key role in the household instead of in higher positions.

The second assumption is that women are being prevented from becoming active in society under the pretext of religious beliefs.

The third assumption is that there is a negative and significant link between national culture and the attitude of people towards women in leading administrative positions.

The last assumption is that there are gender stereotypes about women. Such stereotypes increase the negative attitude of women towards leading positions.

4. Conclusion

The protection of women's rights and gender equality policy is being actively pursued in New Uzbekistan. But the practicality of the reforms is not yet complete.

The numbers are not transparent for data analysis.

Due to the change in the country's policy, a special database of government organizations working with women has not yet been formed.

Although gender equality reflects a positive attitude towards employment in the sectors, the election of women to leadership positions still faces certain obstacles.

Analyzing the theoretical views, we can say that feminism in Uzbekistan has no political support but is being replaced by a fair gender equality reform. Despite efforts to ensure human rights and gender equality in line with international standards, the distribution of women in the workplace remains disproportionate. The New Uzbekistan gender equality policy is commendable, with women outnumbering men in the labor sector. According to our data analysis, gender equality in the formal employment sectors in the provinces was found to be moderate (45%). But the distribution of women in leadership positions remains very weak (7-8%). Thus, the appointment of women leaders within the government and gender equality is better ensured than in other sectors. It was also found that the field of education is a field with a large number of female leaders and a high number of female workers. The root of the existing obstacles and causes, on the one hand, is rooted in the national-cultural mentality, on the other hand, indicates that the political will in many is not yet fully formed.

It is clear that in New Uzbekistan, the laws on equal rights and opportunities for men and women need to be fully implemented on the ground. We believe that it is important to continue the policy of expanding the recommendation of women for elected and appointed positions in all spheres and levels of social and public life. Because discrimination against women is caused by gender stereotypes in society. We believe that women's participation in decision-making processes at all levels is important. Our study and analysis show that while there have been government reforms in the nomination of women for leadership positions, the inclusion of women in leadership positions at all levels has not been enough yet. So, we believe that public control, along with law enforcement agencies, is important in guaranteeing that the law works in this regard. Topics such as the analysis of differences in women's wages and the analysis of the situation of women in the informal labor market are also worthy of further study in the future.

Acknowledgment

I thank my dear Professor Dr. Ahn Young-Jin (South Korea, Chonnam National University, Graduate School of Social Sciences) as my supervisor and for providing insight that greatly assisted the study. Professor Dr. Ahn was never tired of sharing his pearls of wisdom with me. And his comment and feedback greatly improved the manuscript. Any errors in the study are my own and I wish those shortcomings should not tarnish the reputation of my professor.

우즈베키스탄의 남녀평등 공간성과 여성지도자 지위

주라에브 주히르띤* • 안영진**

*대학원생, 사회학대학 지리학과, 전남대학교
**교수, 사회학대학 지리학과, 전남대학교

요약

이 연구는 사회지리학에서 과학 연구의 수에 기여하기 위한 것이다. “새로운 우즈베키스탄”이라는 용어 아래 공간분석을 바탕으로 개혁이 양성평등에 미치는 영향을 살펴본다. 우리의 연구는 이번이 처음이기 때문에,

우리는 먼저 이론적 배경을 더 자세히 다루었다. 우리는 페미니즘 및 성별 연구 분야를 우리의 연구와 과학적으로 연계시켰다. 우리는 여성 리더십 포지션 문제에 대한 양성평등과 여성주의적 접근법이 미치는 영향을 분석했다. 본 논문의 목적은 성별 연구의 맥락에서 이루어져야 할 작업을 분석하고 향후 이루어져야 할 연구를 파악하는 것이다. 또한 중앙아시아, 특히 우즈베키스탄에서는 페미니스트 지리학과 젠더 연구 작업에 대한 출처가 제한적이다. 이어 "새로운" 우즈베키스탄의 직장 내 여성 비율과 여성에 대한 리더십 포지션의 공간성을 조사했다. 결론 부분에서, 우리는 우리의 의견과 제안을 제공했습니다.

키워드: 성 연구, 여성 리더십 위치, 공간성, 페미니스트 지리, 성 평등.

승인:

안영진 교수님께서 저의 지도자로서 그리고 연구에 큰 도움을 주신 것에 감사드립니다. 안영진 교수는 자신의 지혜의 진주를 저와 나누는 것에 지치지 않았습니다. 그리고 교수님의 논평과 피드백은 원고를 크게 향상시켰습니다. 연구에서 어떤 오류도 저만의 것이고 그러한 단점들이 교수님의 명성을 더럽히지 않기를 바랍니다.

References

Aculai, E., A. Bulgac, V. Gryga, O. Krasovksa, and O. Linchevskaya (2009) *"Researching the Gender Aspect in Business Development (the Case of Moldova, Ukraine, and Belarus)."* Lex ET Scientia International Journal 16(1).

Alexander Gallas (2010) *"Dichotomy, Dualism, Duality: An investigation into Marxist conceptualizations of structure and agency"*., Academia.edu p.86.

Alice H.Eagly, and Linda L.Carli (2003) *"The female leadership advantage: An evaluation of the evidence"* Elsevier, Leadership Quarterly, volume 14, issue 6, pp., 807-834.

Amy L. Kotsenas (2017) *"Women in Leadership: Why So Few and What to Do About It"* Journal of the American College of Radiology, Volume 14, Issue 4, Pages 450.

Bartold, Vasily (2020) *"Turks. 12 Lectures on the History of the Turks of Central Asia"*. Moscow: Yurayt Publishing house. p. 181.

Chandra Talpade Mohanty, Ann Russo, and Lourdes Torres (1991) "*Third World Women and the Politics of Feminism"*., Paperback. Indiana University Press, Article n3, pp.251, 275.

Chandra Talpade Mohanty (1984) *"Under Western Eyes: Feminist Scholarship and Colonial Discourses"*., JSTOR., pp.335-341.

Friederike Welter, David Smallbone (2010) *"Institutional Perspectives on Entrepreneurial Behavior in Challenging Environments"*.

Jiekuan Zhang et al., (2020) *"Tourism and gender equality: An Asian perspective"*.

Judith Lorber (2006) *"Shifting Paradigms and Challenging Categories"* vol.53, p.462.

Lorenzo Guarcello, Fabrizia Mealli, and Furio Camillo Rosati (2010) *"Household Vulnerability and Child Labor: The Effect of Shocks, Credit Rationing and Insurance"* January edition, Journal of Population Economics 23(1): pp. 169-198

Hamza Yusuf Hanson, John Taylor Gatto, Nabila Hanson, and Dorothy Sayers (2003) *"Educating Your Child in Modern Times: How to Raise an Intelligent, Sovereign & Ethical Human Being*"., Paperback., pp 82-93

Helen Pluckrose et al., (2020) *"Cynical Theories*"., pp.139-147., USA.

Manju Rani, Sekhar Bonu (2008) *"Attitudes Toward Wife Beating: A Cross-Country Study in Asia"*., https://doi.org/10.1177/0886260508322182

McCarthy, Cormac (2006) *"The Road"* New York: Vintage. [Google Scholar], p.22

Naomi Ellemers, Floor Rink, Belle Derks, Michelle K. Ryan (2012) *"Women in high places: When and why promoting women into top positions can harm them individually or as a group (and how to prevent this)"* Journal name: Research in Organizational Behavior., Elsevier., Volume 32, Pages 163-187.

Pauline Couper (2015) *"A Student's Introduction to Geographical Thought",* chapter 6, feminist geography.

Rosemarie Tong, Tina Fernandes Botts (2017) *"Feminist Thought A More Comprehensive Introduction"*, Routledge publication house. Chapters 1-3.

Sattar, Sarosh (2012) *"Addressing the Gender Gap in Europe and Central Asia. Europe and Central Asia knowledge brief"*; issue no. 54. World Bank, Washington, DC.

Stefan Schutte (2014) *"Living with patriarchy and poverty: women's agency and the spatialities of gender relations in Afghanistan"* Gender Place and Culture A journal of Feminist geography (21)

Stevi Jackson (2001) *"Why a Materialist Feminism Is (Still) Possible—and Necessary"* Women's Studies International Forum 24, no.3–4.

Tim Cresswell (2012) Geographic Thought: A Critical Introduction., Feminist geography., pp. 147-165., Wiley-Blackwell edition.

A Study of The STEAM Toy Pedagogy Plan Evaluation for Elementary School

Wen-Te Chang, Yun-Hsin Pai

***Abstract*—** Purpose: Based on the interdisciplinary of lower grade Elementary School with the integration of STEAM concept, related wooden toy and pedagogy plans were developed and evaluated. The research goal was to benefit elementary school education. Design/methodology/approach: The subjects were teachers from two primary school teachers and students from the department of design of universities in Taipei. Amount of 103participants (Male: 34, Female: 69) were invited to participate in the research. The research tools are "STEAM toy design" and "questionnaire of STEAM toy Pedagogy plan." The STEAM toy pedagogy plans were evaluated after the activity of "The interdisciplinary literacy discipline guiding study program--STEAM wooden workshop," Finding/results: The study results: (1) As factors analyzing of the questionnaire indicated the percentage on the major factors were cognition teaching 68.61%, affection 80.18% and technique 80.14%, with α=.936 of validity. The assessment tools were proved to be valid for STEAM pedagogy plan evaluation; (2) The analysis of the questionnaires investigation confirmed that the main effect of the teaching factors was not significant (affection = technique = cognition); however, the interaction between STEAM factors revealed to be significant ($F(8, 1164) = 5.51$, $p < .01$); (3) The main effect of the six pedagogy plans was significant (climbing toy > bird toy = gondola toy > frog castanets > train toy > balancing toy), and an interactive effect between STEAM factors also reached a significant level, ($F(8, 1164) = 5.51$, $p < .01$), especially on the artistic (A/ Art) aspect. Originality/value: The main achievement of research:

(1) A pedagogy plan evaluation was successfully developed.

(2) The interactive effect between the STEAM and the teaching factors reached a significant level.

(3) An interactive effect between the STEAM factors and the pedagogy plans reached a significant level too.

***Keywords*—** STEAM, toy design, pedagogy plans, evaluation.

Wen-Te Chang is with the National Taipei University of Education, Taiwan (e-mail: wtchang@mail.ntue.edu.tw).

Finnish Teachers Promote Students' Phenomenon-based Learning through Situational Interest Teaching Strategies

Wen-Hua Chen

***Abstract*—** To develop the competencies needed for the 21st century, the City of Helsinki encouraged schools to adopt a phenomenon-based learning approach under the 2016 new core curriculum. Phenomenon-based learning motivates teachers and students to use real-world phenomena as themes or issues to encourage students to explore phenomena with a more holistic perspective. How the new learning method promotes and maintains students' interest and how teachers and students operate in school are all the issues that this research focuses on. Given that interest has an important place in the learning process and is also an essential topic for research, this study explores how schools can promote students' phenomenon-based learning through situational interest teaching strategies. The two main issues of this research include exploring the design elements of case schools that promote students' phenomenon-based learning through situational interest teaching strategies, and the main factors that influence students' phenomenon-based learning through situational interest teaching strategies. The case study approach is beneficial to employ when there is a need to obtain an in-depth appreciation of an issue, event, or phenomenon of interest in its natural, real-life context. This case study was conducted from mid-February to early June 2018 in a public school in Helsinki, applying observation, interviews, and document analysis. Several key findings emerged from this study. First, the design of phenomenon-based learning in the case school mainly started with the concept "Change" as a theme and then planned toward maintain situational interest. Second, the main factors that promote phenomenon-based learning are thematic interest, textual interest, task interest, personal characteristics, and process interest. Although the phenomenon-based learning implemented by the case school is only one week, it allows students to experience a diverse and positive learning environment and creates opportunities and space to integrate knowledge, skills, emotions, attitudes, and values in various fields. The above results are mainly related to the school's active use of situational interest teaching strategies to promote students' phenomenon learning. This study enriches the understanding of interest structure and indicates that teacher and student co-empowerment helps all learners learn through situational interest.

***Keywords*—** Finland education, phenomenon-based learning, situational interest, individual interest.

Wen-Hua Chen is with the National Taipei University of Education, Taiwan (e-mail: chenwenhua@mail.ntue.edu.tw).

The Increasing Importance of the Role of AI in Higher Education

Joshefina Bengoechea Fernandez, Alex Bell

***Abstract*—** In its 2021 guidance for policy makers, the UNESCO has proposed 4 areas where AI can be applied in educational settings: These are: 1) Education management and delivery; 2) Learning and assessment; 3) Empowering teachers and facilitating teaching, and 4) Providing lifelong learning possibilities (UNESCO, 2021).

Like with wblockchain technologies, AI will automate the management of educational institutions. These include, but are not limited to admissions, timetables, attendance, and homework monitoring. Furthermore, AI will be used to select relevant learning content across learning platforms for each student, based on his or her personalized needs. A problem educators face is the "one-size-fits-all" approach that does not work with a diverse student population.

The purpose of this paper is to illustrate if the implementation of Technology is the solution to the Problems faced in Higher Education. The paper builds upon a constructivist approach, combining a literature review and research on key publications and academic reports.

***Keywords*—** artificial intelligence, learning platforms, students personalised needs, life- long learning, privacy, ethics.

Josefina Bengoechea is with the Geneva Business School, Spain (e-mail: jbengoechea@gbsge.com).

Roles Currently Played by Educational Middle Leaders

Elaine Marta Pereira Aaltonen

***Abstract*—** Effective school leadership materialised in educational settings through the high standard professional performance of senior and middle leaders, has increasingly become an education policy priority around the world due to a wide recognition that schools need knowledgeable, skilled, and committed leaders, along with great teachers, in order to ensure outstanding education at all levels of schooling. The scope of this paper is the work of middle leaders, whose direct influence on teachers and classroom teaching, thus, on student learning outcomes, is a key component for successful school systems. It particularly aims at sharing some of the findings obtained through an academic study recently carried out by the same researcher, which was focused on enhancing understanding about aspects related to the professional performance of educational middle leaders, applied to the context of the lower elementary school division of a private mainstream school located in Brazil. The master´s dissertation findings included identifying the roles performed by a team of educational middle leaders throughout the year of 2021, as well as gaining insights on their perceptions about the roles performed, both through an electronic questionnaire and individual face-to-face interviews. Not only the roles of the middle leaders who participated in the research have been identified through the qualitative case study undertaken, but additional research finding lying within the sphere of the categorisation of such roles, based upon coherent domains of practice, has possibly been made. Hence, the main purpose of this paper is to outline the findings concerning the current roles played by educational middle leaders.

***Keywords*—** roles, middle leaders, educational leadership, school leadership, and management.

Elaine Aaltonen is with the University of London, Brazil (e-mail: aaltonenelaine@gmail.com).

Performance Management System for Specialization Course Coordinators

Ana Paula Novaes, Olga Guilhermina Farah, Carlos Alexandre Negri, Everton Teles Prazeres, Vivian Prieto Ferreira

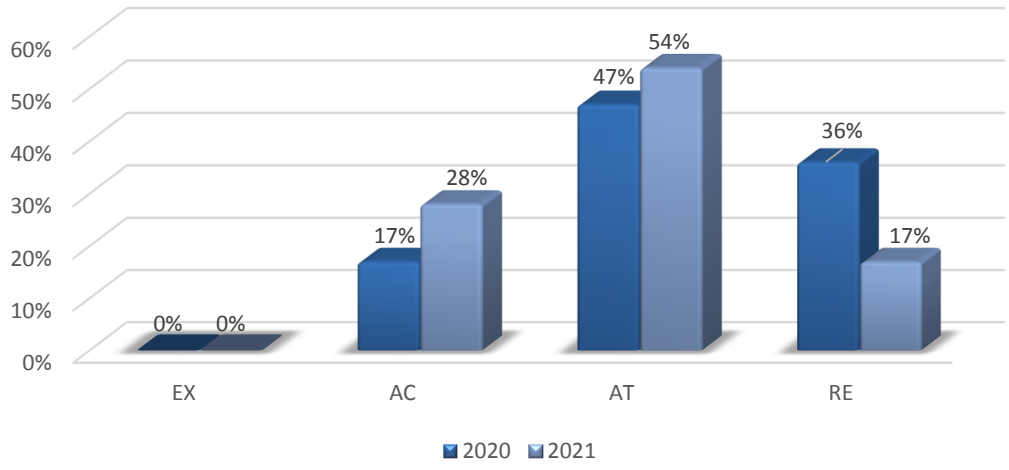
Vivian Prieto Ferreira is with the Albert Einstein Education and Research Center, Brazil (e-mail: vivian.pf@einstein.br).

Introduction: The role of the course coordinator in specialization courses is fundamental, as it involves academic, administrative and control actions, establishing an important link between students, teachers and the Institution. The Specialization Program, in alignment with the strategies of the Teaching Board of the Teaching and Research Institute of the Albert Einstein Hospital, seeks the constant improvement of the performance of course coordinators through their broad engagement, knowledge and academic development, thus ensuring a process of continuous improvement inserted in a culture of quality and guarantee of learning. Objective: This study aims to describe the performance evaluation processes of specialization course coordinators and relationship strategies with the groups related to each of the performance levels. Method: This is a quantitative, descriptive, prospective study. The population consisted of 150 course coordinators who account for approximately 145 classes. All active coordinators are evaluated at least once a year, subsidized by the Relationship Program with specialization course coordinators. The pillars, which are composed of a set of indicators, are organized in weights that reflect the level of relevance of each of the indicators in the context of the Program, configuring a performance measurement system. Using statistical description calculation techniques, the results obtained relate to the number of points achieved for each evaluation dimension. The composition of the results provides for the distribution of coordinators in the following performance ranges, as annex 1. Partial and final results are published on a unique electronic platform, preserving the privacy and reliability required by access control. In this presentation, indicators are stratified and related to their scores. A view of the overall score obtained is also presented, collaborating with the process of continuous improvement. In addition to feedback, the coordinators with underperforming performance are invited by the Specialization Program Management and the Department of Teaching Quality for a meeting, seeking to explore the results and formulation of the performance improvement plan. Results: The project started in 2020 where, the following stratification was observed: 17% of the coordinators placed at level d and performance above expected, 47% in the range within the expected and 36% below expected. Improvement actions were implemented, which were implemented, which showed the most performance fragility. In 2021, there was a significant improvement in distribution, obtaining 28% of the coordinators placed at a performance level above expected, 54% in the range within the expected and 17% below expected. The graph in Annex 2 consolidates the results, showing the gains obtained.

Conclusions: The results obtained contribute to the achievement of a Relationship Program based on recognition and obtaining benefits linked to each of the performance levels of the year. Additionally, there were additional gains aimed at greater group engagement, improvement in the management of the teaching-learning process, greater satisfaction of coordinators and teachers and strengthening of the culture of excellence and management by data and results.

Anexo1

Anexo 2



Bibliografia

1. Oluseye O, Tairat T, Emmanuel 0. Customer Relationship Management Approach and Student Satisfaction in Higher Education Marketing. Journal of Competitiveness [Internet]. 2014; 6(3); 49-62. Disponível em https://www.cjournal.cz/files/175.pdf

2. Timothy L, Lerzan A, Cooil B, Andreassen TW. Linking Customer Loyalty to Growth. MIT Sloan Management Review [Internet]. 2008; 49(4);51-57. Disponível em https://www.researchgate.net/publication/230557994_Linking_Customer_Loyalty_to_Growth
3. Novaes AP, Pavan FR. Fundamentos para Melhoria da Qualidade em Serviços de Saúde. 1a ed. São Paulo. Editora Cedet; 2018.
4. FONSECA M; FONSECA DM. A gestão acadêmica da pós-graduação lato sensu: papel do coordenador para a qualidade dos cursos lato sensu. Educ. Pesqui., São Paulo, v. 42, n. 1, p. 151-164, jan./mar. 2016

A Case Study of the Role of Environmental NGOs and Their Access to Court in Finland

Terhi Raikas

***Abstract*—** The Aarhus Convention gives Environmental NGOs a right to access to justice in its member countries. This means that these NGOs can appeal to court in cases environmental law is not complied for the reason or another.

In my article, it is researched in which roles environmental NGOs have been in the Finnish case law and in the case law of the European Court of Human Rights. This means that the legal cases are in six different groups:

1) Cases in which the main question has been whether the environmental NGO has even had the right to make a complaint;

2) Cases in which the NGOs have litigated and the actual theme of the litigation has been to research the questions the NGOs have complained about;

3) Cases in which it has been discussed whether the NGO has had a proper right to access to information;

4) Cases in which the organization has been used as an expert;

5) Cases in which NGOs have given a note; and

6) Cases that are related to Environmental Criminal Law.

The methods of my article are based on the traditional legal dogmatics and the use of interviews. For the research, both the officials and people who are volunteering in the Finnish Environmental NGOs are interviewed. In these interviews, it is researched whether the opinions about the role of the NGOs are varying between the officials and the NGOs.

It is clear that in most of the cases environmental NGOs have had the opportunity to make litigate without any significant problems. However, in some cases it has been necessary to research questions that have not been linked to the questions the NGOs have complained about and, instead of them, in these cases the main question has been linked to the NGOs' possibilities to litigate.

About the six different groups that were researched in this article, the most complicated one was the last one. In 1998, it was stated by the Council of Europe in the Convention on the Protection of the Environment through Criminal Law that each Party may, in accordance with domestic law, grant any group, foundation or association which, according to its statues, aims at the protection of the environment, the right to participate in criminal proceedings concerning offences established in accordance with this Convention. However, this is extremely rare.

As a conclusion, the requirements of the Aarhus Convention are taken good care of in the Finnish legal system. The role of the NGOs is strong especially in cases related to Administrative Law, but it could be made better in Criminal Law and this could help to protect the Finnish environment.

***Keywords*—** Aarhus Convention, access to justice, environmental organizations, Finland, NGOs.

Terhi Raikas is with the University of Vaasa, Finland (e-mail: talvikki.r@windowslive.com).

An Analysis of the Philippines' Legal Transition from Open Dumpsites to Solid Waste Management Facilities

Mary Elenor Adagio, John Roben Ambas, Ramilyn Bertolano, Julie Ann Garcia

***Abstract*—** Ecological Solid Waste Management has been a long-time concern in both national and international spheres. The exponential growth of waste generation is not properly matched with a waste management system that is cost-effective. As a result, governments and their communities within inevitably resort to the old ways of opening dumpsites to serve as a giant garbage bin. However, due to the environmental and public health problems these unmanaged dumpsites caused, countries like the Philippines mandated the closure of these dumpsites and converted them into or opened new sanitary landfills. This study aims to determine how the transition from open dumpsites to Solid Waste Management Facilities improve the implementation of the Solid Waste Management Framework of the government pursuant to Republic Act 9003. To test the hypothesis that the mandatory closure of dumpsites is better in the management of wastes in local government units, a review of related literature on analysis reports, news, and case studies was conducted. The results suggest that advocating for the transition of dumpsites to sanitary landfills would not only prevent environmental risks caused by pollution but also reduce problems regarding public health. Although this transition can be effective, data also show that with a lack of funding and resources, many local government units still find it difficult to provide their solid waste management plans and to adapt to the transition to sanitary landfills.

***Keywords*—** solid waste management, environmental law, solid waste management facilities, open dumpsites.

John Roben Ambas is with the Arellano University School of Law, Philippines (e-mail: ambasroben@gmail.com).

Plastic Pollution: Analysis of the Current Legal Framework and Perspectives on Future Governance

Giorgia Carratta

***Abstract*—** Since the beginning of mass production, plastic items have been crucial in our daily lives. Thanks to their physical and chemical properties, plastic materials have proven almost irreplaceable in a number of economic sectors such as packaging, automotive, building and construction, textile, and many others. At the same time, the disruptive consequences of plastic pollution have been progressively brought to light in all environmental compartments. The overaccumulation of plastics in the environment, and its adverse effects on habitats, wildlife, and (most likely) human health, represents a call for action to decision-makers around the globe. From a regulatory perspective, plastic production is an unprecedented challenge at all levels of governance. At the international level, the design of new legal instruments, the amendment of existing ones, and the coordination among the several relevant policy areas requires considerable effort. Under the pressure of both increasing scientific evidence and a concerned public opinion, countries seem to slowly move towards the discussion of a new international 'plastic treaty.' However, whether, how, and with which scopes such instrument would be adopted is still to be seen. Additionally, governments are establishing regional-basedstrategies, prone to consider the specificities of the plastic issue in a certain geographical area. Thanks to the new Circular Economy Action Plan, approved in March 2020 by the European Commission, EU countries are slowly but steadily shifting to a carbon neutral, circular economy in the attempt to reduce the pressure on natural resources and, parallelly, facilitate sustainable economic growth. In this context, the EU Plastic Strategy is promising to change the way plastic is designed, produced, used, and treated after consumption. In fact, only in the EU27 Member States, almost 26 million tons of plastic waste are generated herein every year, whose 24,9% is still destined to landfill. Positive effects of the Strategy also include a more effective protection of our environment, especially the marine one, the reduction of greenhouse gas emissions, a reduced need for imported fossil energy sources, more sustainable production and consumption patterns. As promising as it may sound, the road ahead is still long. The need to implement these measures in domestic legislations makes their outcome difficult to predict at the moment. An analysis of the current international and European Union legal framework on plastic pollution, binding, and voluntary instruments included, could serve to detect 'blind spots' in the current governance as well as to facilitate the development of policy interventions along the plastic value chain, where it appears more needed.

***Keywords*—** environmental law, European union, governance, plastic pollution, sustainability.

G. Carratta is a PhD Candidate at HHL Leipzig Graduate School of Management, Germany (g.carratta@hhl.de). Moreover, she is also an Early Stage Researcher within the EU-founded "LimnoPlast Project" (Grant Agreement No. 860720).

All in the Family: Parenting Style and Risk Behavior in College Students

Randall E. Osborne

***Abstract*—** Since pioneering work by Baumrind in the 1970's, researchers, educators and practitioners have attempted to delineate the impact of parenting style on children. Parenting style may contribute to family dysfunction and family dysfunction has been shown to relate to risk-behavior in college students. The current study assessed parenting style by asking participants to self-report of how often each parenting style – authoritarian, permissive and authoritative – was used in the home after definitions for each parenting style were provided. In addition, elements of family dysfunction were assessed (using the Edinburgh Family Scale) and risk behavior in college students was measured (utilizing the 2017 CDC Youth Risk Behavior Survey which assesses 9 categories of risk behavior). As predicted, students who rated their parent(s) as authoritarian in style, also reported significantly higher levels of several elements of family dysfunction and reported higher levels of risk behavior in their first year of college than students reporting a more authoritative style of parenting. These effects differed somewhat for males and females. The implications for parenting style on risk behavior in college students is discussed.

***Keywords*—**Authoratative parenting; authoritarian parenting; permissive parenting; parenting style; risk behavior.

Randall E. Osborne is with Texas State University, 601 University Drive, San Marcos, TX, USA (phone: 512-245-8236; fax: 512-245-3153; e-mail: ro10@txstate.edu).

Mindfulness and Motivational Based Intervention for Pregnant Women with Tobacco Dependency: Pilot Study

Ilona Krone

***Abstract*—** Maternal smoking during pregnancy increases the risk of perinatal/postnatal negative health outcomes; however, only 1 in 5 pregnant smokers quit smoking. That is a clinical and public health problem. Pregnant smokers have negative paternal support, and higher levels of perceived stress than non-smokers and quitters return to smoking in a stressful situation. A crisis like the COVID-19 outbreak causes significant uncertainty and stress. For pregnant women, additional stress may increase due to concerns for their fetus. Strategies targeting maternal stress and isolation may be particularly useful to prevent negative outcomes for women and their fetuses. Within the post-doctoral study, cooperating with leading specialists, an innovative program for pregnant smokers will be developed. Feasibility for reducing craving, distress intolerance, Covid 19 related stress, and fear in pregnant women in Latvia will be assessed.

***Keywords*—** COVID 19, mindfulness, motivation, pregnancy, smoking cessation.

Ilona Krone is with the University of Latvia, Latvia (e-mail: ilonakrone@gmail.com).

Design, Emotion, and the Ghost of Jay Doblin

D. J. Richter-O'Connell

***Abstract*—**The purpose of this paper is to examine the role of human emotion and emotional responses in design processes, methodologies and outcomes for both students and practitioners of Industrial Design. It takes a meandering personal path through academic and professional experiences but concludes with focused recommendations for more emotionally aware and inclusive strategies during the design research process.

***Keywords*—** Design Research, Design Methods, Emotion, Jay Doblin

I. Introduction

In the late 1980's, I had the great fortune of having Jay Doblin as one of my undergraduate professors at the Institute of Design in Chicago. At that point I do not think I fully appreciated his impact on the Industrial Design profession as a practitioner and theoretician, but I fully understood his great influence as an Educator. A quiet, caring, and kind-hearted elder who patiently nurtured head, hand, and heart skills; told great Loewy life-in-the-trenches stories; and had the uncanny ability to deconstruct complex concepts for his students into simple and manageable modules. Jay's health was failing in those years and we lost him in May of 1989. His spirit has guided me throughout my career, standing over my shoulder and gently prodding me to find the emotion-filled humanistic heart of any problem; empathize deeply; search fearlessly and exhaustively for the unexpected; tell stories through my visualizations; and edit down to the simplest and strongest systems and solutions I could uncover.

Jay's ability to turn complexity into understandable components was embodied in *A Short, Grandiose Theory of Design* [1]. Jay argued that the methodological design process was evolving from a simple, single-step sequence model for discrete products- into a multi-phase process of more complex uni-systems, multi-systems and eventually eco-systems of products, people, services, experiences, philosophies, information and applications.

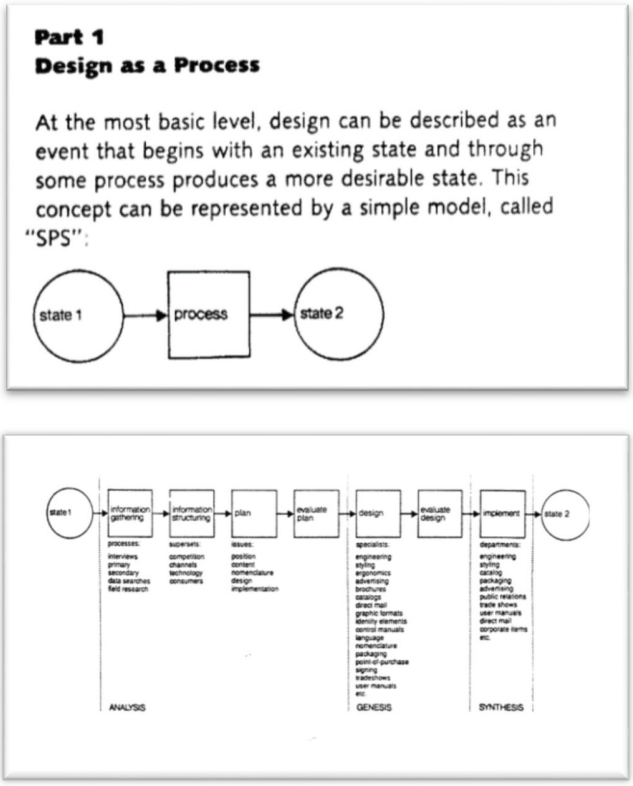


Fig. 1 Jay Doblin's simple SPS Design Process Model and the more complex Multi-Phase Model.

Figure 1 illustrated Doblin's discussion of the use of the 'SPS' design process to move from a less desirable 'current' state 1 to more desirable 'future' state 2. He argued that this simple model still applied to simple problems. In the second model, the end 'states remained, but the process evolved into a multi-phased sequence required of more complex problems and system-solving. It features multiple activities organized in three distinctive steps.

1. Analysis (Design Research)- Information Gathering, Structuring, Planning, and Evaluation
2. Genesis (Creative Exploration)- Design and Evaluation
3. Synthesis (Commercialization)- Implementation

D.J. Richter-O'Connell is with the Industrial Design program in the School of Art & Design at the University of Wisconsin-Stout, Menomonie, WI 54751 USA 715-232-2507 richter-oconned@uwstout.edu

Doblin's model laid the groundwork for Design Process models to this day and was open-ended enough to encompass the complex contemporary Societal, Cultural, Economic, Ecological, Technological, Competitive, Cooperative, Brand Building forces and factors that Industrial Designers balance every day. Jay's three step process heavily influenced me as a design student, and I have modeled and remodeled my own process map over the last 30 some years using the enduring Doblin sequence as an underlay.

II. Consumer Behavior 101

As a graduate student in Design Studies, I re-encountered Jay's simple 3 step SPS model in an unexpected place. In a Consumer Behavior course, the text was seeking to model the pattern of consumer 'need recognition' and buying behavior in a simple 'state' model [2].

1. Consumer's start in a Current or 'Actual State' that includes some level of awareness of a human need, desire, and/or a deficiency or dissatisfaction with existing conditions.
2. A process of researching, shopping, marketing/ messaging (consciously or subconsciously), evaluating, budgeting, and buying allows consumers to make a choice and make a purchase.
3. The final step is the hoped for 'Desired State' with a satisfied consumer experience and outcome.

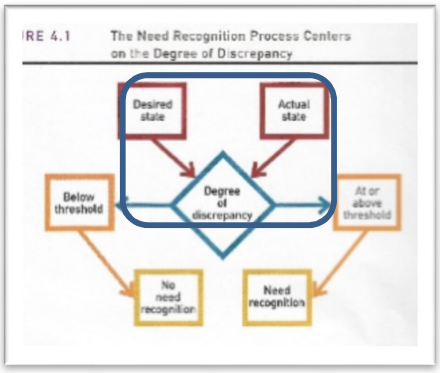


Fig. 2 Consumer Need Recognition and Purchase Decision making Model

This new perspective, from a consumer's point of view, offered a parallel pathway on the flip side of the design process. The Industrial Designer's Analysis/Research phase identifying consumer needs, expectations, pain and pleasure points, observations, ethnography, usability testing, surveys and interviews used to help understand consumer needs directly overlaps the Consumer's decision-making experience before, during and after the purchase of a new product or service.

The awareness of the consumer side of the story not only adds insights and benefits to Industrial Designers during the new product development process, but all members of the development and supply chain- marketers, engineers, manufacturers, sales, suppliers, distributors and retailers.

III. Ergonomics 101

My teaching duties include the fundamentals of Ergonomics/ Human Factors. I frame the class around identifying, understanding, and addressing human centric needs, desires, and experiences in the three realms of Physical; Cognitive; and Behavioral/ Emotional/ Psychological Ergonomics, shown in Figure 3.

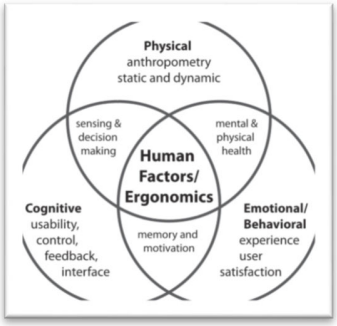


Fig. 3 Author's Teaching Tool- Ergonomics/ Human Factors

The Emotional/ Behavioral/ Psychological portion of the model is often the most difficult one to portray and predict for students. To better tell this part of the story, I created a short power point that walks through the process of 'human experience' including physical, cognitive and emotional components to put the three realms into the bigger human process and perspective. Figure 4 is one of the sequence of slides from that presentation. Please note: The use of the Jay Doblin person is not accidental.

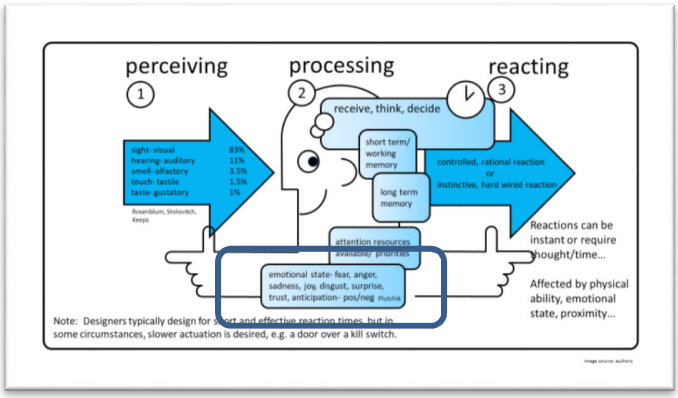


Fig. 4 Author's Teaching Tool- Human Experience Model- Perception, Processing and Reaction/Feedback

Near the bottom center of Figure 4 is a text panel that introduces 'emotional state' as a key filter in the information processing and decision-making portion of the experiential model. The diagram shows the Plutchik model of recognizing eight distinct human emotional states or states of feeling- joy, trust, fear, surprise, sadness, anticipation, anger, and disgust [3].

Plutchik argues that these eight are deeply rooted in our evolutionary past and remain vital and valid today. He created a wheel, in figure 5, that shows the basic emotional states in varying degrees from very low levels at the outer points of the wheel, increasing to extremes at the very top/center.

Fig. 5 Plutchik's Emotional Response Wheel

Over time researchers have challenged and edited this list down to 6 states- anger, happiness, surprise, disgust, sadness, and fear- based on relationships between subjects reported emotional states and corresponding facial expressions [4]. A recent study from the University of Glasgow suggests only 4 states- happy, sad, fear/surprise, disgust/anger [5]. All research teams agree, regardless of counts of core modalities, a wide spectrum of each emotion's intensity and combinations of emotions conspire to create a broad spectrum of emotional responses in humans.

It may be argued that, in all models, there are emotional states that are largely considered more positive than others. Happiness (Joy), Surprise (with safe and positive outcomes), Anticipation (again with safe and positive outcomes). Trust might be generally considered a positive emotional state that consumers and designers (and brands) may desire moving toward. Fear, Sadness, Surprise (with threatening outcomes), Anticipation (with threatening outcomes), Anger, and Disgust all might generally be considered negative emotional states that would want to be avoided or guarded against.

Conversely, some consumers and designers may really want to be scared and surprised on Thrill Rides, in Haunted Houses, or disgusted with Garbage Pail Kids toys and cakes filled with gummy worms. The current 'unboxing' trend in toys is seemingly based on the best aspects of Anticipation and Surprise- LOL Dolls, Hatchimals, Kinder Joy and Boxie Girls.

So where do Industrial Designers go with all of this? Is it methodologically applicable in the design process? If we agree that a component of our profession (whether from the designer's process point of view, or the consumer's behavioral/ choice point of view) is to help people transition from less desirable physical, cognitive and emotional 'states' to more desirable ones, then it can be argued that there is value in understanding how these emotions occur, develop, rise and fall and manifest in consumers choices as they solve their problems via acquiring designed object, systems, services and experiences.

Sometimes it is beneficial to observe the extremes…

IV. PSYCHOLOGY, PSYCHIATRY, AND BEHAVIORAL HEALTH THERAPIES

The Social and Medical sciences have developed professional fields and practices that help people with episodic or chronic behavioral/ emotional challenges through analysis, dialog, discussion, drugs, medical procedures, or combinations of the above. All mental and behavioral health professionals use assessment tools to identify negative and detrimental emotional trajectories and develop plans and regimens to move patients into more production and positive patterns, lifestyles, and 'states'. What do these assessment tools look like? How do they work? And could they be adapted and/ or integrated into classic design research methodologies via surveys, interviews, observational and ethnographic research, empathetic studies, shadowing, etc.?

The Diagnostic and Statistical Manual of Mental Disorders (DSM-5) represents something of a bible for these professionals and practitioners [6]. It contains an abundance of assessment tools and treatment scenarios for a variety of mental and mood disorders. The tools are geared towards profound and pathological patients but give clues about normal behavior and emotional levels as well. The surveys share some common attributes that yield both quantitative and qualitative information about a patient.

- Each asks a series of categorized questions, asking similar emotionally/ behaviorally focused questions multiple times in slightly different ways to make sure that a topic has been covered thoroughly and addresses compound or cross-over emotional maladies.
- Each answer is requested in an emotional time frame. Are you feeling this now? Today? This week? This month? This year? These answers are extremely helpful in determining what should be considered a 'state' vs. a 'trait'. A state is more temporal and episodic and may be handled through 'talking' therapies, a 'trait' has become entrenched in a patient's personality and requires more aggressive treatments. (A 'state' condition is considered closer to normal behavior)
- The third element is a gauge of emotional intensity. Are these feelings low level or crisis level? Rate them on a scale of 1 - 5 or 1 - 10. Are these thoughts irritating or aggravating? Are they passing thoughts or impulses that have been acted upon?

Again, these are tools designed to identify and address emotional extremes but may add insights into understanding and planning for emotional shifts and trajectories in everyday users and consumers. At their core they reinforce some important 'best practices' in investigative research techniques- multiple pathways of inquiry, allowances for quantitative data collection, and for qualitative reflection.

V. BACK TO DESIGN METHODS

Don Norman, Cognitive Scientist, Educator and Design Analyst, who has devoted years to understanding the connections between design, decision making and human emotion, says that "Emotion is always passing judgements, presenting you with immediate information about the world: here is a present danger, there is potential comfort; this is nice, that is bad". He describes the 'emotional state' as it relates to the design of objects/ systems in three levels- visceral, behavioral, and reflective [7].

- *Visceral* a subconscious reaction to what we perceive through the senses. We either like it or we do not based on innate inherited human emotional factors- a sense of beauty, proportion, scale, form, or… something is wrong, off balance, out of harmony, distorted.
- *Behavioral* is again subconsciously driven but based on learned behavior about how things should work, where controls should be, what I expect to happen, whether I react positively or negatively to these usability factors.
- *Reflective* is the conscious review of all the inputs, memories, judgements, and emotional influencers along the experiential pathway. It is analogous to the entire experiential model fig 4.

Ironically, Norman talks about the fact that we can understand the physical, physiological, intellectual, and emotional model of user/consumer experience but still cannot get inside consumer's heads and predict or anticipate emotional outcomes. The implementation of this knowledge, by designers, during the design process, is somehow hidden in unwritten chapters of his books- and can only be realized when we internalize, empathize, mature our human intuition and master the art of understanding human need.

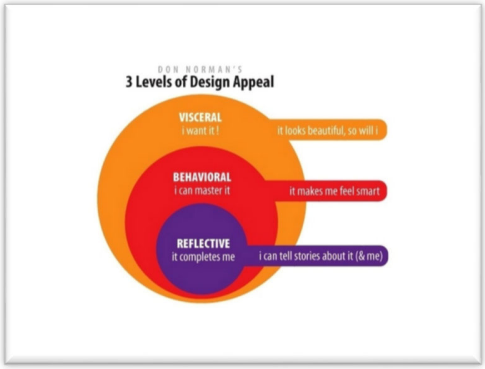


Fig. 6 Norman's Levels of Emotional Design

VI. DESIGN RESEARCH METHODS

Design Research is the specialized sequence of inquiry and analysis of trends and trajectories affecting a particular design and development project. It is a core foundational component of all Industrial Design curriculums and lays the critical and creative thinking foundation for design thinking and decision-making throughout a designer's career. In its simplest form it is a guide for designers to ask and answer a series of questions covering the What? Why? Who? When? Where? and How? of a designed product or system striving to be new, novel, addressing contemporary and anticipated human needs, desires, and issues [8].

In Human-Centered Design strategies that focus heavily on addressing human/consumer needs and desires, the 'Who?' question is often the most important one to answer. If we do not have a customer and an unmet need, we do not have a product. Some of the most valuable tools in this endeavor are those that allow Designer and Consumer to meet and discuss human-centered issues and opportunities. These are typically observations/ethnography/shadowing, interviews, surveys, and usability testing.

All of these tools can be used early in the research process when issues are still being searched for and discovered, or later in the process as issues, opportunities, and options need consumer verification.

Tools used early in the process as designers are hunting and gathering information are generally more qualitative, open ended, conversational, involve observation and listening, allowing researchers to pursue unexpected paths of questioning. Tools used later in the process are often more quantitative and are looking to verify forms, functions, fit and feel.

Observational Research/ Ethnography/ Shadowing

Observational research can be active or passive but always involves watching users perform tasks in their natural environments- places where they are familiar, comfortable, stress-free, and able to act in ways that are normal and

indicative of their everyday experiences. Ethnographic Research and Shadowing usually involve no real time interaction. Active Observational Research allows the researcher to ask questions on the fly as activities are happening to clarify and verify. Skilled researchers probe the physical activities and decisions, but also ask the 'how do you feel about what is happening?' questions.

Using the 4 emotional states model, perhaps the 'feel' questions can be informed and formulated to address these directly. Does this make you feel Happy? Sad? Fearful/Surprised? Disgusted/Angered? …and why? What was it about what just happened that triggered this emotion? Is it a low level or high-level emotion? Does this happen to you often? Has it happened in the past with a similar product or activity? Do these emotions rise to the level of you wanting to take action? Return product to the store? Buy something new? Spend more money to get something better? Hire someone else to do it?

Interviews/ Surveys

Interviews and Surveys are often not in a consumer's natural surroundings. Researchers find that consumers may 'talk the talk' in an interview setting but not 'walk the walk' in a real-world scenario, so outcomes can be unreliable. Interviews, however, can still be extremely useful in front end research when designers are still searching for features and benefits that will generate a new 'product experience' with unexpected and meaningful advantages. They can be powerful downstream tools as well, as designers plumb the depths of reactions to formal and functional features being formulated.

Surveys are beneficial for quantifiable data, rankings and needs hierarchies but suffer some from the inability to probe further into unexpected areas. Good surveys include discrete multiple choices, value or intensity ranking, and room for 'other' responses that allow respondents to explain issues at a deeper level. On-line survey tools like Qualtrics and Survey Monkey are readily available to design students and professionals alike and often offer formatting choices that represent best practices in inquiry techniques.

Mood/Image/Inspiration Boards

Mood boards are strong visual tools that help designers tell the stories and show the inspirations for their formal and functional design objectives. Perhaps a shift from 'Mood' to 'Emotional State' boards could help designer students communicate the behavioral shifts they hope to help consumers make through the experience of their solutions. A movement from Sad to Happy, + good Surprise, + good Anticipation.

Usability Studies

Usability Studies focus on the preferences and frustrations of likely users/ consumers throughout the design process. Early uses tend to focus on establishing the landscape of the marketplace and do head-to-head analyses between key competitive players. Who am I competing against? What are the cost/ price objectives? What are the feature and benefit sets that are a key to entry into this marketspace?

Downstream Usability Studies are very important in the refinement stages of design when major and minor differences in form, function, fit, and feel can have major market and consumer consequences. Unpleasant, unsafe, uncomfortable attributes are potential showstoppers and must be addressed before final design and a great deal of effort is made to identify and correct negative perceptions and potential usability issues in the final design phases. These 'pain points' may show up as physical issues but may not register as emotional issues if not asked in that way. 'Could the discomfort issues you have identified lead to emotional responses of Sadness?, Fear/unwanted Surprise?, Disgust/Anger?'

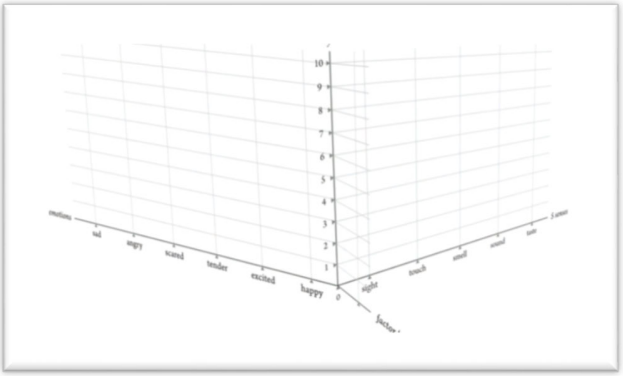


Fig. 7 Lee's Design for the Senses + Emotions Model

Usability studies may tend towards the sensorial experiences of a product and less the feelings and emotions they may illicit. What if we first framed usability studies to address all five primary senses in the order of importance- visual, auditory, olfactory, tactile, and gustatory? Jinsop Lee, designer, educator and researcher argues product experiences can be more meaningful and memorable if addressing all five senses [9]. In figure 7, he also argues for matrixing the 5 senses against the 8/6/4 emotional states and a potentially new emotionally informed methodological tool for design students and professionals.

VII. CONCLUSIONS

If we agree, as posited earlier, that…

- a significant part of our role and responsibility as Industrial Designers- students, educators and professionals is to help users/ consumers move from Doblin's and Blackwell's 'less desirable state' (state 1) to a 'more desirable state' (state 2)…

- and that we navigate this transition through the research, design, and development of objects, experiences, interactions, interfaces, applications, services, and systems…
- and that those states are a complex cocktail of physical, cognitive, and emotional/behavioral attributes…

Then a sensitivity to the inclusion and integration of emotion-centric research, observation, questioning, inquiry, verbiage, and evaluation are required in the teaching, development, and application of design methodologies.

The worlds of Psychology, Psychiatry and Behavioral Therapy can suggest models for paths of inquiry that elevate awareness of a range of normal-to-extreme root emotions both positive and negative.

Specific Design Research methodologies can be re-written with emotionally inclusive discussion paths and trajectories. These more qualitative results may require the higher levels of maturity, empathy, experience, and worldliness that Norman argues are hidden in the unwritten chapters of his books.

The author suggests that these kinds of emotionally sensitive lines of inquiry and research are essential to the best practices of Industrial Designers and, like the lessons of Jay Doblin, are best established early in Design Education experiences.

VIII. THE GHOST OF JAY DOBLIN

In the last visit Jay paid to us in the spring of 1988, he had been in and out of the hospital several times, and his fatigue and frailties were showing. We asked if there were things we could do to help with his healing. He smiled his fatherly smile and pulled out his Walkman and ejected a cassette. It was Richard Pryor Live on the Sunset Strip. He asked us for comedy tapes! Laughing was the one thing that helped him endure the discomfort of his treatments and the best thing we could do for him was to help him laugh.

Later I could see how Doblin-ian that request was. His Walkman was a well-designed, human-centric, object that existed in an 'experiential' uni-system of delivering audio content in a very personal, private and portable format and, in turn, co-existed within a multi-system of creative content providers operating in a multi-faceted entertainment and information industry… which included comedy albums.

I might add that a very real emotional design exchange happened that day as Jay's (and our) experiential eco-system expanded to include the intangible attributes of love, care, respect, honor, empathy, and our desire to stay connected with Jay through his last journey. There had been a tangible shift of mood in the room from a state of 'sadness' to a state of 'happiness/joy'

I can still see the image of him smiling that afternoon.

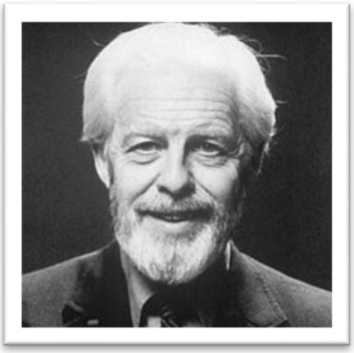


Fig. 8 Jay Doblin 1920-1989

REFERENCES

[1] Doblin, J., *A Short Grandiose Theory of Design*, STA Design Journal, 1987, republished 1990 Doblin, Keeley, Malin, Stamos

[2] Blackwell, R.D., Miniard, P.W., & Engel, J.F., *Consumer Behavior*, Thomson Southwestern, 2006, Mason, OH.

[3] Plutchik, R., *Basic Emotions*, 1980, PersonalityReseacrh.org

[4] Eckman, P.,*Emotions Revealed, Recognizing Faces and Feelings to Improve Communication and Emotional Life*, Henry Holt and Co. LLC, 2003, New York, NY.

[5] Jack, R.E., Garrod O.G.B, Schyns, P.G., *Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals over Time*, Current Biology, 2014, ScienceDirect.com

[6] American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Arlington, VA: American Psychiatric Association, 2013.

[7] Norman, D., *Emotional Design, Why We Love (or Hate) Everyday Things,* Basic Books, 2004, Cambridge, MA.

[8] Milton, A., Rodgers, P., *Research Methods for Product Design*, Laurence King Publishing Ltd., 2013, London, UK.

[9] Lee, J., *Design for all 5 Senses*, presentation, and video, 2013, TED2013.

FIGURES

Figure 1. Graphic representations of simple and complex Industrial Design Processes. Reprinted from Doblin, J. (1987) *A Short Grandiose Theory of Design*, STA Design Journal, republished 1990 Doblin, Keeley, Malin, Stamos

Figure 2. Graphic representation of the Consumer's Decision-Making Process. Reprinted from Blackwell, R.D., Minicar, P.W., & Engel, J.F. (2006), *Consumer Behavior*, Thomson Southwestern, Mason, OH.

Figures 3 & 4. Author's graphic content depicting Teaching Model; Experiential Design & Decision-Making Teaching Process Model. 1988-2020.

Figure 5. Graphic representation of Plutchik's Emotional Wheel. Reprinted from Plutchik, R. (1980) *Basic Emotions*, PersonalityReseacrh.org

Figure 6. Graphic representation of Norman's 3 Levels of Emotional Design

Figure 7. Graphic representation of Jesop Lee's 5 Senses and 8 Emotions Evaluation Tool. Reprinted from Lee, J., (2013) *Design for all 5 Senses*, presentation, and video, TED2013.

Figure 8. Portrait of Jay Doblin reprinted from Wikimedia.

Adolescents' and Young Adults' Well-Being, Health, and Loneliness during the COVID-19 Pandemic

Jessica Hemberg, Amanda Sundqvist, Yulia Korzhina, Lillemor Östman, Sofia Gylfe, Frida Gädda, Lisbet Nyström, Henrik Groundstroem, Pia Nyman-Kurkiala

***Abstract*—**Purpose: There are large gaps in the literature on COVID-19 pandemic-related mental health outcomes and after-effects specific to adolescents and young adults. The study's aim was to explore adolescents' and young adults' experiences of well-being, health, and loneliness during the COVID-19 pandemic. Method: A qualitative exploratory design with qualitative content analysis was used. Twenty-three participants (aged 19-27; four men and 19 women) were interviewed. Results: Four themes emerged: Changed social networks – fewer and closer contacts, changed mental and physical health, increased physical and social loneliness, well-being, internal growth, and need for support. Conclusion: Adolescents' and young adults' experiences of well-being, health, and loneliness are subtle and complex. Participants experienced changed social networks, mental and physical health, and well-being. Also, internal growth, need for support, and increased loneliness were seen. Clear information on how to seek help and support from professionals should be made available.

***Keywords*—**adolescents, COVID-19 pandemic, health, interviews, loneliness, qualitative, well-being, young adults.

Jessica Hemberg is with the Åbo Akademi University, Finland (e-mail: jessica.hemberg@abo.fi).

The distribution of *HLA-DQA1*01:02* and *HLA-DQB1*06:02* frequencies in Thais and Asians: Genetics Database insight for COVID-19 severity

Jinu Phonamontham[1], Patompong Satapornpong, Ph.D.[2,3]

1. Khon Kaen Wittayayon school, KhonKaen Thailand
2. The division of general pharmacy practice, department of pharmaceutical care, College of Pharmacy, Rangsit University, Pathum Thani, Thailand
3. Excellence Pharmacogenomics and Precision Medicine Centre, College of Pharmacy, Rangsit University, Pathum Thani, Thailand

Abstract

Introduction: Coronavirus, also referred to as COVID-19, is a virus caused by the SARS-Cov-2 virus. The pandemic has caused over 10 million cases and 500,000 deaths worldwide through the end of June 2020. In a previous study, *HLA-DQA1*01:02* allele was associated with COVID-19 disease (*p*-value = 0.0121). Furthermore, there was a statistical significance between *HLA- DQB1*06:02* and COVID-19 in the Italian population by Bonferroni's correction (*p*-value = 0.0016). Nevertheless, there is no data describing the distribution of *HLA* alleles as a valid marker for prediction of COVID-19 in the Thai population.

Objective: We want to investigate the prevalence of *HLA-DQA1*01:02* and *HLA-DQB1*06:02* alleles that are associated with severe COVID-19 in the Thai population.

Materials and Methods: In this study, we recruited 200 healthy Thai individuals. Genomic DNA samples were isolated from EDTA blood using Genomic DNA Mini Kit. *HLA* genotyping was conducted using the Lifecodes HLA SSO typing kits (Immucor, West Avenue, Stamford, USA).

Results: The frequency of *HLA-DQA1* alleles in Thai population, consisting of *HLA-DQA1*01:01* (27.75%), *HLA-DQA1*01:02* (24.50%), *HLA-DQA1*03:03* (13.00%), *HLA-DQA1*06:01* (10.25%) and *HLA-DQA1*02:01* (6.75%). Furthermore, the distributions of *HLA-DQB1* alleles were *HLA-DQB1*05:02* (21.50%), *HLA-DQB1*03:01* (15.75%), *HLA-DQB1*05:01* (14.50%), *HLA-DQB1*03:03* (11.00%) and *HLA-DQB1*02:02* (8.25%). Particularly, *HLA- DQA1*01:02* (29.00%) allele was the highest frequency in the NorthEast group, but there was not significant difference when compared with the other regions in Thais (*p*-value = 0.4202). *HLA-DQB1*06:02* allele was similarly distributed in Thai population and there was no significant difference between Thais and China (3.8%) and South Korea (6.36%) and Japan (8.2.%) with *p*-value > 0.05. Whereas, South Africa (15.7%) has a significance with Thais by *p*-value of 0.0013.

Conclusions: This study supports the specific genotyping of the .*HLA-DQA1*01:02* and *HLA-DQB1*06:02* alleles to screen severe COVID-19 in Thai and many populations.

Keywords: *HLA-DQA1*01:02, HLA-DQB1*06:02,* Asian , Thai population

Influence of Board Composition and Organizational Culture on the Performance of Faith Based Hospitals in Kenya

Authors: [1&]Josephine Ndege, [1]Wanja Tenambergen, [1]Kezia Njoroge

[1]Department of Health Systems Management, Kenya Methodist University

[&]Corresponding author: Email: joendege2000@gmail.com ; Cellphone: +254 721214938

ABSTRACT

The world over, there has been the proliferation of faith-based organizations focusing on all aspects of society and the economy. In Kenya, faith-based hospitals provide 40% of health care services to the medium and low-income populations in the rural areas. It has also been observed that some health facilities are failing to a level of total collapse while others have registered extremely poor organizational performance in terms of various outputs to a level that they merely exist and are run with little sustainable growth. Some scholars have attributed some of these failures to overlooking the adoption of corporate governance practices where for instance, boards of management of health facilities are largely made up of religious officials, therefore making the facilities lack proper corporate culture when compared to their more privately-owned health facilities counterparts. Scholars and experts in health systems have consequently argued that it is important to stem the trend of poor performance or the imminent collapse of these hospitals. It is against this background that this study sought to examine the manner in which faith-based hospitals have infused corporate governance practices if any, and how these practices influence the hospitals' overall organizational performance. The study examined the influence of board composition and organisational culture on the performance of faith-based hospitals in Kenya. A cross-sectional survey design was adopted, whereby mixed methods data collection strategies were employed. For quantitative data, questionnaires were used with senior managers of Christian faith-based hospitals in addition to nine purposively selected informants for an in-depth interview. For quantitative data, descriptive and inferential analysis techniques were used, while for qualitative data, content analysis was done. From the regression results, corporate governance practices of board composition and organizational culture had significant and positive influences on the performance of faith-based hospitals in Kenya. The study recommends an enhanced criterion for choosing board members to ensure there is no bias in the selection process, and board members should be properly inducted and continuously trained on governance to ensure they execute their mandate effectively.

INTRODUCTION

Faith-based health facilities in Kenya form a key pillar in provision of health services. They are particularly relevant in provision of health care in rural and marginalized areas of the country where public health service provision is low, failed or non-existent (Basu, et al., 2012). Together with other private health facilities, they complement public health service provision in a major way. However, with proliferation and increase in the number of health care provision ranging from public to private for-profit and private not-for-profits (where faith-based facilities fall) health providers, it has become pertinent for faith-based health facilities that rely on their self-funding models to evaluate the sustainability of their organization performance. When compared with other private facilities which are for-profit, faith-based health facilities present a bleak outlook from an organizational performance and sustainability perspective.

Studies shows that the private run hospitals provide the ideal model for sustainably running and managing health facilities because they have been able to embrace the best corporate governance practices and rules (Basu, et al, 2012; Kinyanjui, et al., 2015). By adopting good corporate governance, private health facilities are able to attract diverse skills, competence and knowledge by incorporating diverse set of membership in their management boards thus improving quality decision making by way cultivating accountability, ensuring inclusivity by engaging stakeholders, enabling setting up of shared direction, as well as ensuring good stewardship of resources in more prudent ways. These four elements thus become the key cogs in ensuring that for-profit private health providers remain vibrant, sustainable and high performing health systems.

In contrast however, most of the not-for-profits private health providers like faith-based health facilities are yet to embrace such good corporate governance practices, therefore exposing the institutions to the risks of continued reliance on boards of managements that are strictly constituted and run by religious leaders. The same religious are also responsible to specialist decisions making affecting the performance competitiveness of these health facilities. Therefore, under these kinds of situations, corporate governance practices in these institutions have been thought to be unclear, short of transparency, lacking in accountability, lacking complete disclosures, failing, and approaching their resultant collapse (Kiliko, 2016).

In the light of these observations, there are clear gaps in the current understanding and knowledge on how good corporate governance practices and the values they bring to running of health facilities influences the overall organizational performance of Faith-based health facilities. It is not therefore clear how much the governance of not-for-profits faith-based health facilities are guided by the corporate governance practices (Malesi & Njeru, 2017). This study therefore sought to systematically examine the manner in which corporate governance practices influences the organizational performance of the not-for-profits faith-based hospital that have been in operation in Kenya up to the year 2020. In more particular terms, the study focused on five key corporate governance practices of board composition, cultivating accountability, organisational culture, setting of shared direction for a hospital and stewardship of resources which have been argued to be influential in setting up high performing health systems.

METHODS

Research design

The study sought to answer the questions by providing an analysis of the evidence from the data collected. The study primarily adopted a cross sectional descriptive survey design to collect data to help in the achievement of the research objectives and answer the research questions. The research design presents a snap shot of manifestation of variables in a large number of subjects at one point in time (cooper and Schindler, 2011). It is used to describe characteristics of variables, analyse their frequency, distribution, features and observable phenomena of the target population.

Study Site

Because the focus of this this study was the faith-based health facilities in Kenya, it then implies that the entire country where these facilities are located was the site for study. In Kenya, there are 47 counties and it has about 1,072 faith-based health facilities (Goosby, 2014). Faith based health facilities are thus spread across the entire country. However according to CHAK and KCCB reports, in 2019, faith-based hospitals are distributed in 34 of the 47 counties in Kenya.

Target Population

The study was carried out as a cross sectional survey of the entire not for profit faith-based hospitals in Kenya. According to the (ministry of health Kenya, 2016) master list, there are 9,722 facilities in Kenya with 55,233 inpatient beds and 7,495 baby cots. Of these 3613facilities with 11, 213 beds and 1,638 cots are private for-profit facilities while 1,388 facilities with 12,490 beds and 1,534 cots are operated by the not for profit faith-based organizations. Hence, while there is a large number of Faith-based health facilities such as hospitals, health centers and dispensaries in Kenya providing healthcare services, this study targets all faith-based hospital across the country. Hospitals are major contributors in healthcare services and have well established governance structures from strategic (top) level to the lower level management that were likely to influence organizational performance. The target population was the total population of all faith-based hospitals in the country, spread across 34 out of the 47 counties in the country; 62 hospitals from Kenya Conference of Catholic Bishops (KCCB) and 26 from Christian Health Association of Kenya (CHAK). The study targeted 4 strategic level managers in all the 88 hospitals purposely selected considering the position they hold in the hospital, ability and willingness to respond to questions concerning variables under study. These were; one hospital board member (community representatives), CEOs/Administrators, finance managers, and medical superintendents thus making a total of 352 respondents.

Sample size determination

A sample is a subset of the population used to determine the truth about a population (Field 2009) a small portion of the population is selected for observation and analysis to save time and money. However, in this study a census method was used where all faith-based hospital listed in the KCCB and CHAK report of 2020 was analysed. This method is useful when case intensive study is required providing intensive and in-depth information covering many

facets of the problems. Field (2009) noted that the method provides a more accurate and reliable information since every item is put into consideration. The census method provided in-depth information concerning corporate governance practices, an area that has not been intensively studied in faith-based hospitals. A total of 352 respondents were included in this study. Furthermore, 9 key informants were also purposely selected whereby 4 were from CHAK and KCCB secretariats and five board members purposely selected from level 5 and 6 faiths-based hospitals. The two secretariats coordinate faith-based health facilities; they monitor their performance and also appoint the hospital's board members.

Sampling Procedure

A sampling frame constructed from the list of all hospitals affiliated to CHAK and KCCB was made. The latest list (See appendix 6) from the two agencies secretariats 2020 annual report indicates there are 88 hospitals. From each hospital four strategic level managers (board member, CEO/administrator, finance manager and medical superintendents) were the respondents. These purposively selected persons were better suited to respond to questions on board composition, accountability, shared strategic direction, organization culture and stewardship of resources which yielded to a total sample size of 352. The key informants were purposely selected from Level 5 and 6 hospitals that have active board of directors with 9-13 members and have community representation in the board. More specifically, 2 national executive secretaries and 2 chairmen of both CHAK and KCCB secretariats were selected. Then, 5 board members (community representatives) purposely selected from level 5 and 6 hospitals of PCEA Kijabe, Our lady of Lourdes Mwea hospital, Maua Methodist, Consolata hospital Nyeri and St. Mary's Mission Kakamega. These were well placed to enrich and compliment the data from the senior managers in response to the questions on corporate governance practices.

Data collection method and Instrument

The study used a structured questionnaire. Questionnaires provide a relatively cheap, quick and efficient way of obtaining large amounts of information from a large sample of people as in 352 senior managers (respondents) in entire populace of faith-based hospitals in Kenya. In this case, data was collected relatively quickly because the researcher would not need to be present when the questionnaires were completed. The researcher distributed the questionnaires through hand delivery and email. The study also used key Informant Interviews. These are qualitative in-depth interviews with people who know what is going on in the community and the facilities' performance. In this study the key informant interview was conducted with 9 purposively selected informants; 4 from both KCCB and CHAK secretariats and 5 purposely selected board members from level 5 and 6 faith-based hospital. The two secretariats coordinate health services in faith-based hospitals, evaluate their performance and appoint the board members thus most suitable in enriching and complimenting the data from the senior managers.

Quantitative data was collected by administering structured questionnaires to senior hospital managers while qualitative data was collected via key informant interviews with the selected hospitals. The researcher distributed the questionnaires through hand delivery and email. The questionnaires had opening remarks explaining their purpose, instructions for filling and giving an assurance to the respondents that the responses given would be treated with confidentiality. After completion of data collection exercise all questionnaires were adequately checked for data quality. The questionnaire was then systematically coded and categorized for the purposes of transcribing the findings into a computer program. The process consisted of data cleaning and initial data analysis. Data cleaning was done to inspect and correct erroneous entries. The data was then entered and analysed as per the research objectives using statistical package for social sciences (SPSS) verson 21.0. Qualitative approach was used to help overcome the weaknesses associated with only one method. Quantitative approach on the other hand was used to help in establishing the relationships between the predictor variable and the dependent variables.

Ethical Considerations

This research was conducted in a safe, ethical manner that ensured that the humanitarian imperative of Do-No-Harm was strictly followed. The study also respected the ethical principles of confidentiality and informed consent and took serious steps to avoid raising expectations. The research was reviewed and approved by the Kenya Methodist University Science, Ethics, and Research Committee (SERC) and the National Commission of Science and Technology (NACOSTI). The researcher sought clearance from CHAK and KCCB secretariats and hospital CEOs/Administrators. Informed consent was also obtained from all the respondents. Furthermore, the researcher assured the informants and the respondents about the utmost observance of confidentiality of use of their information purely for research purposes only and that just in case of need for transmission of data other than for

this thesis, their consent will be sought before it is used. Only the stamped consent forms were used for recruitment of respondents.

Data Analysis and presentation

The collected data was first keyed in to statistical software to aid in the analysis. In particular, the Statistical package for social sciences commonly known as IBM SPSS was used in the analysis. The procedure involved creating unique codes for all the variables, then keying in the requisite data as collected from the survey. The data was then cleaned to recode again to facilitate statistical analysis. The coded data was then analyzed using descriptive and inferential statistics. Descriptive statistics involved the use of frequencies, percentages and cross tabulations. Inferential statistics involved the use of the chi-square test of significance and the logistic regression. Chi- square test was used as a preliminary test to show whether each independent variable is significantly associated with the dependent variable. On its part, logistic regression was used to determine the cause and effect relationship existing among variables. Logistic regression is used when the dependent variable is dichotomous or binary in nature. Put in a simpler manner, the dependent variable must possess two mutually exclusive outcomes such as bad or good, success or failure and so on. In this study the dependent variable which is access to healthcare services was coded into two mutually exclusive outcomes which are; 'accessible' and 'not accessible'.

Data presentation was done using tables, diagrams, bar graphs and pie charts. The presentations sought to display relationships and reveal patterns between the particular predictor variables and the independent variables with the aim of deriving important conclusions from which recommendations can be made. Quantification of the qualitative data is done to help in describing certain characteristics and enable representation of values in terms of quantities or numerical values. The data was also categorized in a way that makes it easy for the mind to absorb and understand which ultimately help in describing the situation through numbers and representing it through tools like tables' histograms, bar charts, pie charts and frequency curves.

RESULTS

Board Composition

Majority hospital boards were not satisfied with the composition of hospital board 215(71.2%), existing skills 238 (78.8) and relevant academic qualification 236 (78.1%). However, a considerable number of boards had members with adequate levels of experience109 (36.9%). On the stake holder's representation, the respondents disagreed that the governance's framework recognizes the rights of stakeholder's in the management of their facilities 224 (74.1%), inclusion of members with leadership expertize 204 (67.5%), that stakeholders have access to relevant information 221 (73.2%), stakeholders are often represented in key decision making 220 (68.5%). However, on stakeholder's forums aspect, majority of the respondents agreed that their health facility always has at least one stakeholders' forum in a year 207(68.5%). A majority also showed dissatisfaction with the frequency of the stakeholder forums 216 (72%). The results are shown in the table below.

Statements	**Dissatisfied (%)**	**Satisfied (%)**
Board composition		
I am always satisfied with the composition of our health facility board	215(71.2)	87 (28.8)
I am always satisfied with the variety of experience of the health facility board members.	193(63.9)	109(36.1)
I am always satisfied with the existing skills exhibited by the health facility board of directors.	238(78.8)	64(21.2)
The health facility's sponsor often considers relevant academic qualification when appointing the health facility board members.	236(78.1)	66 (21.9)
stakeholder's representation		
The governance framework recognizes the rights of the stakeholders in the management of the health facilities.	224(74.1)	78(25.9)
The health facility board composition includes members who are recognized for leadership and/or areas of expertise and are representative of stakeholders including program participants.	204(67.5)	98(32.5)
The stakeholders always have access to relevant information	221(73.2)	81(26.8)
The stakeholders are often represented in key decision making related to the	220(72.8)	82(27.2)

organization's governance.		
Stakeholder's forums		
The health facility always has at least one stakeholders' forum in a year.	95(31.5)	207(68.5)
I am always satisfied with the frequency of the stakeholder forums	209(69.2)	93(30.8)

This study reveals that the composition, skills exhibited by the health facility board of directors, and qualifications are associated with perceived health facility good performance. In addition, having at least one stakeholders' forum was also associated with perceived good performance. However, the latter four significantly influenced perceived health facility good performance. Having skilled board of directors is key to prosperity of health facilities since they are key element of corporate governance. Moreover, health facilities that consider academic qualifications of board members to be selected boost health facility performance. This study has established that stakeholder's forums enhance health facility performance, compared to those who do not. An active engagement of stakeholders can thus be considered both as a condition for and a consequence of stakeholder approach to corporate governance.

Organizational culture

The respondents were requested to indicate their level of agreement with the statements on organizational culture in faith-based hospitals in Kenya. The results are shown in the table below.

Statements	**Disagree n(%)**	**Agree n(%)**
Organizational Norms		
The hospital board always**…**		
…ensure that the organizational culture is aligned to the mission and vision of the hospital.	112(37.1)	190(68.3)
… uphold high work ethics in hospital. encourage sharing, collaboration, and mutuality in this organization	224(74.2)	78(25.8)
… ensures that the hospital respects cultural diversity, language, and religious beliefs.	231(76.5)	71(23.5)
… ensures that the hospital promotes a non-discriminating work environment.	219(72.5)	83(27.5)
… hospital CEOs have activities to reinforce the core values of the hospital among the staff.	249(82.5)	53(17.5)
… ensures that employees of the hospital are aware of the hospitals core values.	217(71.9)	85(28.1)
On boarding process of new staff		
The hospital has an orientation program that is consistently applicable across departments to on board new staff.	129(42.7)	173(57.3)
The on boarding process for new staff is comprehensive (staff are taken through all departments)	209(69.2)	93(30.8)
The on boarding process takes more than 2 weeks.	211(69.7)	91(30.3)
New staff members are always assigned a coach during the on boarding process.	66(21.9)	236(78.1)

Majority of the respondents agreed that the hospital board always ensure that the organizational culture is aligned to the mission and vision of the hospital 190(68.3%), that the board has an orientation program for board members 173(57.3%) and that New staff members are always assigned a coach during the on boarding process 236(78.1%). However majority disagreed board members uphold high work ethics in hospital. Encourage sharing, collaboration, and mutuality in this organization 224(74.2%), ensures that the hospital respects cultural diversity, language, and religious beliefs. 231(76.5%), and ensures that the hospital promotes a non-discriminating work environment.219 (72.5%), hospital CEOs have activities to reinforce the core values of the hospital among the staff. 249(82.5%), ensures that employees of the hospital are aware of the hospitals core values 217(71.9%) the on boarding process for new staff is comprehensive (staff are taken through all departments) 209(69.2%) and the on boarding process takes more than 2 weeks 211(69.7%).

DISCUSSION

Board Composition

This study reveals that the composition, skills exhibited by the health facility board of directors, and qualifications are associated with perceived health facility good performance. In addition, having at least one stakeholders' forum was also associated with perceived good performance. The findings above concur with both CHAK, KCCB and price (2018) who notes that boards with diversity of skills, knowledge base and core competences bring about various kinds of expertise and perspectives out well in favor of the overall improved performance. KCCB (2018) also supports the idea of proactive culture of commitment and engagement by the board, this ultimately translate to high effective boards and drives the organization towards high performance. According to CHAK governance policy guidelines, the board shall be balanced considering factors such as age, gender, profession and stakeholder's representation (medical/nursing, clergy, management, accountants, lawyers, provincial administration, local authorities and community leaders. Boards should, as a matter of necessity, initiate and oversee a process of stakeholder engagement forums in a manner that will make its stakeholders satisfied with the services they are rendered (Cornforth, 2003). Accelerating stakeholder consciousness and connectivity makes imperative for corporate leaders to master the art of organizational metamorphosis—of transforming their organizations from nearsighted, shareholder-centric systems to organizations designed to serve all of their stakeholders all the time (Katwesigye, 2015). KCCB (2018) supports the idea of proactive culture of commitment and engagement by the board, this ultimately translate to high effective boards and drives the organization towards high performance the manners in which boards are composed have an impact in organization's performance (Valeur &Humzah, 2017). The results of this study indicate that the standards set by various scholars, KCCB and CHAK corporate governance policy guidelines on board composition for effective performance of organizations have not been met in faith-based health facilities. This is in line with Emisile (2007) who noted that key challenges to governance in developing countries mixed health systems include a weak infrastructure for gathering information about providers and other stakeholders in fragmented health systems and underdeveloped policy consultations and policy analysis mechanism. Galer et al (2005) argued that effective boards engage with their stakeholders in order to enhance the sense of inclusion participation and collaboration in decision making which is lacking in majority of faith-based hospitals.

Organizational Culture

Organizational culture and norm are key in enhancing the effect of other important cooperate governance elements towards faith-based health facility performance improvement. Therefore, should be considered when designing intervention to enhance performance of faith-based health facilities. The results concur with the study by Prybil et al (2008) who found that only a handful of boards have assigned clear responsibility for board orientation, education and development to a standing board committee which means that it is possible that this important responsibility is either being performed informally or in an ad hoc manner.

CONCLUSION

Board composition, skills exhibited by the health facility board of directors, and having at least one stakeholders' forum have significant influence on faith-based health facility performance. Moreover, the hospital boards' accountability has also been demonstrated to enhance the performance of faith-based hospitals in Kenya. Within the transparency and disclosure context in the board's accountability, it was found that health facility boards that include recognized leaders and/or experts, involve stakeholders in key decision making, provide strategic reports of the facility to the stakeholders at least once a year, often explains the department policies and procedures to the departmental staff, and provide comprehensive report to stakeholders are associated with good health facility performance.

Organizational culture and norm are key in enhancing the effect of other important cooperate governance elements towards faith-based health facility performance improvement. Therefore, should be considered when designing intervention to enhance performance of faith-based health facilities.

REFERENCES

Basu, S., Andrews, J., Kishore, S., Panjabi, R., & Stuckler, D. (2012). Comparative performance of private and public healthcare systems in low- and middle-income countries: a systematic review. PLoS med, 9(6), e1001244.

CHAK. (2019). Governance and Sustainability in Faith based Health care Providers. CHAK Secretariat. Nairobi, Kenya.

Emslie SV (2007). Exploring the Factors that Measure the Performance of Boards of Directors of NHS Foundation Trusts and Their Association Between Board and Organisational Performance. Unpublished MSc dissertation, Birkbeck, University of London

Galer, J. B., Sylvia, V., & Ellis, A. (2005). Managers who lead: a handbook for improving health services. Cambridge, MA: MSH.

Kiliko, J. M. (2016). Strategic Planning in a Turbulent Environment: A Conceptual View. DBA Africa Management Review Journal, 3(1), 73-89.

Kinyanjui, GK, Gachanja, PM, Muchai, JM. (2015). Technical efficiency of hospitals owned by faith based organizations in Kenya. The Journal of Pan African Studies, 8(6):45-63.

Malesi, N. and Njeru, W. (2017). Corporate Governance Practices and Performance of Commercial Banks Listed At the Nairobi Securities Exchange, Kenya. IJBMP, 1(1).

Prybil, L., Levey, S., Peterson, R., Heinrich, D., Brezinski, P., Price, J., & Zamba, G. (2008). in Nonprofit Community Health Systems.

Rao, C. A., Rao, B. P., & Sivaramakrishna, K. (2008). Strategic management and business policy : texts and cases. Excel. Rudani, R. B. (2013). Principles of management. McGraw Hill Education (India).

Valeur, C. and Humzah, D. (2017).Board composition is the beating heart of good governance and high performance. McGraw-Hill.https://www.boardapprentice.com /2017/11/28/board-composition/

Understanding Retail Benefits Trade-Offs of Dynamic Expiration Dates (DED) Associated with Food Waste

Junzhang Wu [a], Yifeng Zou [b *], Alessandro Manzardo [a,], Antonio Scipioni [a]

[a] CESQA (Quality and Environmental Research Centre Department of Industrial Engineering), University of

Padova, Via Marzolo 9, 35131, Padova, Italy

[b] Research Centre for Cold Chain and Standardization, Guangzhou University, 510006, Guangzhou, China

Abstract: Dynamic expiration dates (DEDs) play an essential role in reducing food waste in the context of the sustainable cold chain and food system. However, it is unknown for the trades-off in retail benefits when setting an expiration date on fresh food products. This study aims to develop a multi-dimensional decision-making model that integrates DEDs with food waste based on wireless sensor network technology. The model considers the initial quality of fresh food and the change rate of food quality with the storage temperature as cross-independent variables to identify the potential impacts of food waste in retail by applying s DEDs system. The results show that retail benefits from the DEDs system depend on each scenario despite its advanced technology. In the DEDs, the storage temperature of the retail shelf leads to the food waste rate, followed by the change rate of food quality and the initial quality of food products. We found that the DEDs system could reduce food waste when food products are stored at lower temperature areas. Besides, the potential of food savings in an extended replenishment cycle is significantly more advantageous than the fixed expiration dates (FEDs). On the other hand, the information-sharing approach of the DEDs system is relatively limited in improving sustainable assessment performance of food waste in retail and even misleads consumers' choices. The research provides a comprehensive understanding to support the techno-economic choice of the DEDs associated with food waste in retail.

Keywords: Dynamic expiry dates (DEDs); Food waste; Retail benefits

1. Introduction

Food waste has attracted increasing attention from both governments and the scientific community worldwide in recent years. Globally, approximately one-third of human food produced is wasted or lost along the food supply chain, amounting to a total loss of 1.3 billion metric tons per year (Blakeney, 2019). It was also estimated that about 931 million tons of food waste are generated around the world in 2019, among which 61% came from households, 26% from food service, and 13% from retail (United Nations Environment Programme, 2021). Therefore, the Sustainable Development Goal (SDG) 12.3 of the United Nations calls for halving food waste and reduce food loss at the retail and consumer levers by 2030

(Nations United, 2015).

One of the major policy decisions in retail is setting an appropriate expiration date to manage perishable foods by a way of information technology (Gaukler et al., 2017). This study aimed to address retail benefits trade-offs between fixed expiration dates (FEDs) and dynamic expiration dates (DEDs) to manage fresh foods. In fact, the choice of expiration dates for fresh foods represents the trade-offs, that is, the choice of expiration date is governed by the relative costs of outdating good foods and selling perished foods. Relative costs include healthy, economic, and environmental trade-offs. Retailers may prematurely discard edible fresh foods when the expiration date is set too short. Instead, the retailer may inadvertently sell perished foods to the final consumer at risk of human safety and health if the expiration date is set too long. However, in current practice, the actual remaining life shelf of fresh foods is typically unknown at retail. This is a particularly big issue for retailers since it is challenging to determine the perishing stage of fresh foods through visual or olfactory inspection.

The combined radio frequency identification (RFID), wireless sensor network (WSN), and the related internet of things (IoT) technologies with the theory of temperature-time-tolerance (TTT) in managing perishable food products presents detailed environmental conditions as food flows supply chain, which in turn could be translated into an estimated remaining shelf life of each food products at retail (Jevinger et al., 2014). This technology-enabled mode of setting expiration dates incorporating TTT theory is defined as *dynamic expiration dates* (DEDs). It is vital to address the asymmetry of shelf-life information by integrating dynamic environmental conditions in the food supply chain to facilitate accurate prediction of the remaining shelf life of food products, and thus minimize food loss or waste. In fact, many researchers have proved that it is feasible for first-expire, first-out (FEFO), and least-safe, first-out (LSFO) applied in the food transportation or storage. For example, (Jedermann et al., 2014) used the FEFO system instead of the traditional first-in, first-out (FIFO) to examine the food loss of meats. The application of FEFO reduces the food loss of porks from 16% to 8%, and the food loss of fishes from 15% to 5% during transportation. Besides, (do Nascimento Nunes et al., 2014) also reduce the food loss of strawberries from 37% to 23% by the FEFO system using in the distribution center. However, it is still unknown to quantify the benefits of the DEDs associated with food waste in the retail sector. Some authors pointed out that the DEDs system is more advantageous than the FEDs system, and they only consider specific food type, storage temperature, and fixed replenishment cycle. But interestingly, (Salinas Segura & Thiesse, 2017) doubted that DEDs using sensors may not better than FEDs with a fixed expiration date. Therefore, the potential benefits of the DEDs system in reducing food waste at retail is needed to be further studied.

The aim of the study is to understand the retail benefits trade-offs of the DEDs system in multi-dimensional scenarios based on WSN technology, making it possible to identify the potential improvement in food waste at retail. This work focuses on the retail that sells random shelf-life products under periodic review. The retail also sets expiration dates dynamically based on the TTT theory. Significantly, this study compares the DEDs system with current FEDs practice, in which the expiration date of each food product is predicated on worst case conditions regardless of the uncertain and stochastic environment along the food supply chain. The model considers 27 complicated scenarios using code

programming methods, including different initial quality, change rate, storage temperature, replenishment cycle, in managing fresh foods. Our main contribution is the assessment of the trade-offs in retail benefits to the operating procedure from FDEs to DEDs practices. In addition, this study is seeking to discuss the specific characteristics and applicable conditions of the proposed DEDs system, making it possible to support the integrated expiration dates and food waste coordinated decision-making system for the retail sector.

2. Method

Our modelling vision regarding intelligent food supply chains associated with both the shelf life and food waste could be visualized as Fig. 1. At the product level, fresh foods are pre-cooled immediately after being picked, and then transported to the cold storage. They will be then transported to retail outlets or e-commerce express delivery for sale. Finally, consumers will buy fresh foods and place them in the household refrigerator before eating. At the information level, the WSN system could collect or record the remaining shelf life of all fresh foods by sensing their temperature, humidity, and odor etc. in real time based on RFID technologies. After loading them to the data center, the algorithm was used for the analysis of food quality to dynamically adjust the shelf life of perishable foods based on the TTT theory. The device level concerns cold storage, refrigerated trucks, retail shelf, and household refrigerators. The relations between chain actors in food supply chain network are input to the food expiration date model. The simulation model was programmed using MATLAB platform.

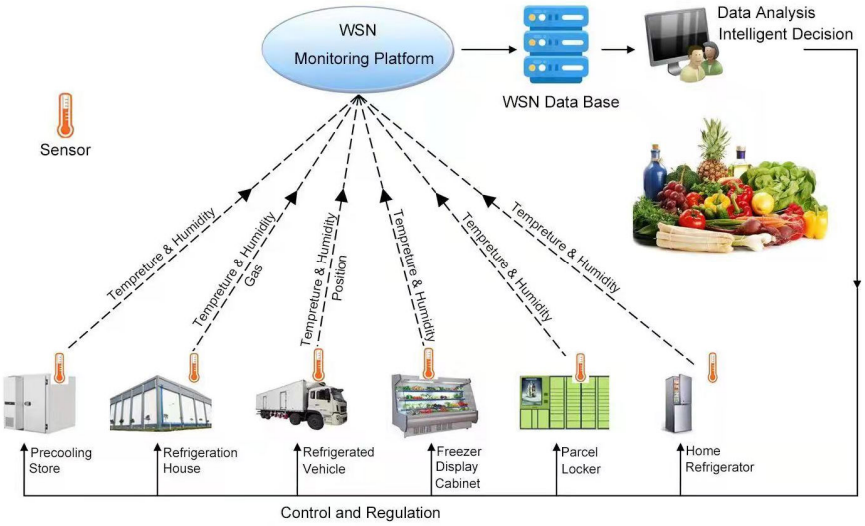


Fig. 1. The general WSN monitoring system for perishable foods.

2.1. Modelling supply chain

One 'average' food supply chain was simulated as Fig. 3. Fresh foods are generally flowed from suppliers, retailers, and consumers along the supply chain.

2.1.1. Supplier

Directly after ordering received from the retailer, new product packages were produced to meet the

retail's replenishment level. After production, suppliers would directly deliver these fresh foods to the distribution center of the retail outlet within a specified period. It is worth noting that each delivery food quality could not consistent since fresh foods always go through a period of logistics operation before they are delivered to retail outlets. Moreover, even the same batch of fresh foods has individual differences in food quality due to temperature sensitivities. We assumed here that it complies with a normal distribution.

2.1.2. Retail outlet and products

In current practices, retail outlets involve food purchased acceptance, sales, order policy, and inventory management. After shipment from the supplier's warehouse, the food batches were transported to the retail outlet. If the percentage of the initial quality of fresh food was less than 30%, it will be rejected. Even qualified fresh foods were sold on the shelf, the food temperature varied with time and shelf device or location (As mentioned in Section 2.2). There are three states of fresh food in retail outlet: on shelf, sold, and wasted. The WSN system may automatically monitor the remaining shelf life of all fresh foods by sensing their temperature and humidity. In general, the remaining expiration date may be changed or modified in the DEDs system if fresh foods may not perish; Otherwise, it would be discarded since it was below a threshold quality limit which the fresh foods are no longer salable. Perishing means reaching a quality state below the quality threshold for salable products here. The detailed explanation is shown in section 3.4. We assumed the replenishment cycle from the supplier to the retailer's distribution center was 1 – 2 days (order policy). Preliminary simulation results showed that when the sales cycle was 10, the distribution interval in each cycle is 10 days and 20 days respectively; thus, the total time range was 100 days and 200 days respectively. From now, the sustainability performance of the expiration date model on retail benefits is not known on forehand, which is determined by the expiration date model itself, consumer demand, order policy, food waste et al. Fresh foods are either sold or thrown away.

2.1.3. Consumer demand

The system is triggered by consumers demanding for fresh foods (Fig. 4). All fresh foods were divided into individual items that could be numbered separately. We assumed the inter-internal time between subsequent incoming consumers to be exponentially distributed. As a result, the visiting consumer each day complies with a truncated normal distribution. In the current work, it was assumed that consumers will simply pick up the freshest products with a longer shelf life. The number of orders received by the retail outlet one time could meet the customer needs for several days; If consumers were faced with an empty retail shelf, which was counted as out of stock. The impact of prices is not included in this study. Many retailer outlets tend to be more cautious about it in China because retailers may bear huge safety risks for the price-cutting promotion policies for expired foods.

2.1.4. Food waste

In the FEDs model, it will be counted as food waste once a product exceeded the indicated shelf life; while it will be removed from the shelf as the waste treatment if the actual food state was below a

threshold quality in the DEDs system. The detailed performance indicators are explained and shown in Section 3.3.

2.1.5. Order policy

Retail outlets are allowed to order every day. A fixed-time period of replenishment with a certain service level was used to model the effect of ordering policies of the retail outlet on the FEDs and DEDs model. At fixed times (such as 22:00 h), the shelf stock will be checked, and orders were made by subtracting the actual stock from a fixed replenishment level. Due to uncertainty of the replenishment and service level, the sensitivity analysis was also performed to understand the retail benefits when applied different expiration date models in this case.

2.2. The expiration date model

In view of the diversity and complexity of the shelf-life prediction on perishable foods, the parameter Q_{10} was often introduced to indicate the increased change rate K in quality of perishable foods when the temperature rose by 10 □ under the condition of the reference temperature T. In general, Q_{10} was set to be 2-4, which corresponds to the food category and reference temperature. The average value of Q_{10} in this study was 3 due to the complexity and uncertainty of quality changes of perishable foods (Jedermann et al.2013). The Q_{10} for temperature dependence is shown in Eq. (1). Although the model varies in the level of its predictive accuracy, its error was still acceptable in a smaller temperature range (Yi et al., 2010).

$$Q_{10} = \frac{k(T+10)}{k(T)} \quad (1)$$

Where, $K(T)$ is the change rate in quality of perishable foods at the reference temperature, and $K(T + 10)$ means the change rate in quality at a temperature of 10 higher than the reference temperature. Assuming that the change rate in quality of the perishable food follows the zero-order reaction equation (such as most fruits and vegetables), the increased change rate in quality of perishable foods caused by the temperature increase in $\Delta T \in [0,10]$ within a certain temperature range is expressed in Eq. (2).

$$Q_{\Delta T} = \frac{k(T+\Delta T)}{k(T)} = （Q_{10} - 1） \cdot \frac{\Delta T}{10} \quad (2)$$

Then, the change rate in quality of perishable foods at any temperatures of $T + \Delta T$ is obtained according to the eq. (3).

$$k(T + \Delta T) = [1 + (Q_{10} - 1) \cdot \frac{\Delta T}{10}] \times k(T) \quad (3)$$

Thus, after a certain time t, the remaining shelf life of the perishable food within a certain temperature is calculated as Eq. (4).

$$SLR = SL - \left[1 + (Q_{10} - 1) \cdot \frac{\Delta T}{10}\right] \times k(T) \times t \qquad (4)$$

In addition, retailers operate the food supply chain under the constraint of service level (α) to reduce the occurrence of out-of-stocks. The service level here refers to the level of guarantee related to out-of-stock. It is expressed in Eq. (5).

$$Q = \lambda \cdot LT + F_\alpha \cdot \sqrt{LT} \cdot SD \qquad (5)$$

Table 2. Parameter's definition.

Symbol	Definition	Unit
k	Change rate in quality of products	%/d
T	Food temperature	□
ΔT	Temperature change	□
α	Service levels	%
λ	Average demand	Consumer unit
L	Simulated sales time	d
M	Total replenishment times	
N	Numbers of actual demand in one replenishment cycle	Consumer unit
Q	Order amount	Consumer unit
LT	Order lead time (or replenishment cycle)	d
SD	Demand variance	Consumer unit
Q_{10}	Quality change	
SQ_i	Out of stock on day *i*	Consumer unit
DQ_i	Demand on day *i*	Consumer unit

2.3. Performance assessment indicators

Practitioners have a growing interest in managing the food quality (safety) and fresh food waste in terms of DEDs practices. In addition, the loss sales of having zero stock while consumers are willing to buy (out of stock) were included in this study. These performance indicators are assessed and quantified as follows.

3.3.1. Average Remaining Shelf life

For every single product sold in a sales cycle, the average remaining shelf life (SLR_{av}) is calculated in Eq. (6).

$$SLR_{\mathrm{av}} = \frac{1}{N}\sum_{i=1}^{N} SLR_i \qquad (6)$$

3.3.2. Food Waste Rate

In this work, the food waste rate (FWR) was used to measure variation in the number of food waste at retail in order to quantify the impacts between the DEDs and the FEDs. The food rejected at the time of incoming by retailers are not included in this study. The food waste rate (FWR) is given in Eq. (7).

$$FWR = \frac{N_{fw}}{M \times N} \quad (7)$$

Where N_{fw} means the total amout of wasted food.

3.3.3. Demand Satisfaction Rate

In general, the higher the demand satisfaction rate (DSR) of customers, the fewer food comsumer units are out of stock. This way, retailers may gain better benefits by implementing the DEDs model along the food supply chain. The DSR is calculated in Eq. (8).

$$DSR = \frac{1}{L}\sum_{i=1}^{L}(1 - \frac{SQ_i}{DQ_i}) \quad (8)$$

3.3.4. Food Safety Rate

In the current study, we defined the expired food sold rate (EFSR) of the DEDs model as 100%. The EFSR is calculated in Eq. (9).

$$\mathrm{EFSR} = \frac{\mathrm{N_{safe}}}{\mathrm{N_{total\_sale}}} \quad (9)$$

Where, $\mathrm{N_{safe}}$ is the number of safe foods sold, and $\mathrm{N_{total\_sale}}$ is the total number of foods sold.

2.4. Scenarios

To better understand the retail benefits from different expiration dates system, multiple scenarios were simulated. Two reference scenarios were considered: one reference scenario with a FED while logistics follow the description in Section 3.1, and the second reference scenario which has the same logistics but combined with a DED system as described in Section 3.2. This study focuses on the different factors of initial quality, quality change rate, and storage temperature of perishable food to examine the impact of sustainability performance indicators related to food safety and waste (as described in section 3.3) on expiration date management systems. The detailed scenarios were considered as shown in Table 3.

(1) To investigate the retail benefits of the FEDs and the DEDs of fresh foods, the storage temperature of the retail shelf is classified into 3 zones: low-temperature (LT, 0 - 8), middle-temperature (MT, 12 - 18), and high-temperature (HT, 20 - 28). Many fruits & vegetables, and most daily foods, such as milk, yogurt, are best stored in the recommended LT; some fruits and vegetables, such as

bananas, are suitable for storing in MT areas; and HT areas are suitable for foods that are not sensitive to temperature.

(2) This study is not limited to a single category of fresh food but divided into three categories according to the change rate in quality: easily perishable, medium perishable and not easily perishable, with three benchmarks of 0 □, 10 □, and 20 □. The preliminary results showed that the simulated shelf-life ranges from 2 to 60 days, which covered the shelf life of most fresh foods.

(3) Even for the same batch of fresh food, the initial quality of the fresh food entering the retail point was different. They are divided into three categories according to the average quality and quality variance: 1) high initial quality means that the initial quality of this kind of food is good, and the difference between individuals is small; 2) medium initial quality means that the initial average quality and variance of this kind of fresh food is average; 3) poor initial quality means that this kind of food The average quality of fresh food is poor, and individual differences are large.

3. Results

As mentioned in Section 2.4, this study focuses on the retail benefits trade-off between two reference scenarios, e.g., the FED mode and the DED mode, through which key factors are analyzed to identify its potential improvement and applicable conditions. In the current work, four system performance indicators in 27 different scenarios of combining fresh food initial quality, quality change rate, and storage temperature conditions (LT, MT, and HT) are quantified to assess the actual retail benefits when DEDs mode and FEDs mode is applied into the practice considering food security and food waste. In addition, two fixed replenishment cycles (RC) of 1 day and 2 days were considered to compare the system performance.

3.1. Shelf life

In this study, the average remaining shelf life (SLR_{AV}) of the fresh food is used to evaluate system performance. The longer SLR_{AV} generally represents the higher quality of food product since it is usable for a certain period after the sale if properly stored. Thus, these products have the requisite quality consume-time that the customer expects. In overall, the SLR_{AV} of fresh food sold in retail in the DEDs mode is slightly better than FEDs mode for multiple scenarios regardless of replenishment cycles (Fig. 2). Especially, the SLR_{AV} of DEDs mode is obviously advantageous from S.10 to S.27 when the replenishment cycle is 2 days. As we know, rational consumers tend to choose the product with the longest SLR_{AV}, whose quality on the retail shelf is randomly distributed. The SLR_{AV} of the DEDs varies per minute while it is still unknown for consumers in the FEDs mode. As a result, the shelf-life information obtained by consumers is asymmetrical. For example, the longest SLR_{AV} of the food indicated on the package may not be freshest in the FED mode at the higher storage temperature of the shelf (as shown in S.10 to S.27, RC=2), which in turn the SLR_{AV} of fresh foods are reduced. However, the SLR_{AV} of different modes are basically the same when the shelf storage temperature is relatively lower (such as S.1 to S. 9 in Fig. 2). Because in the FED mode, the actual SLR_{AV} of fresh foods selected by the consumer based on the marked date is often higher than the market value despite different SLR_{AV} units between DEDs mode and FEDs mode. Therefore, the FEDs mode is still recommended for fresh food with good storage conditions considering the difficulty of technology and economic cost of the DED

pattern.

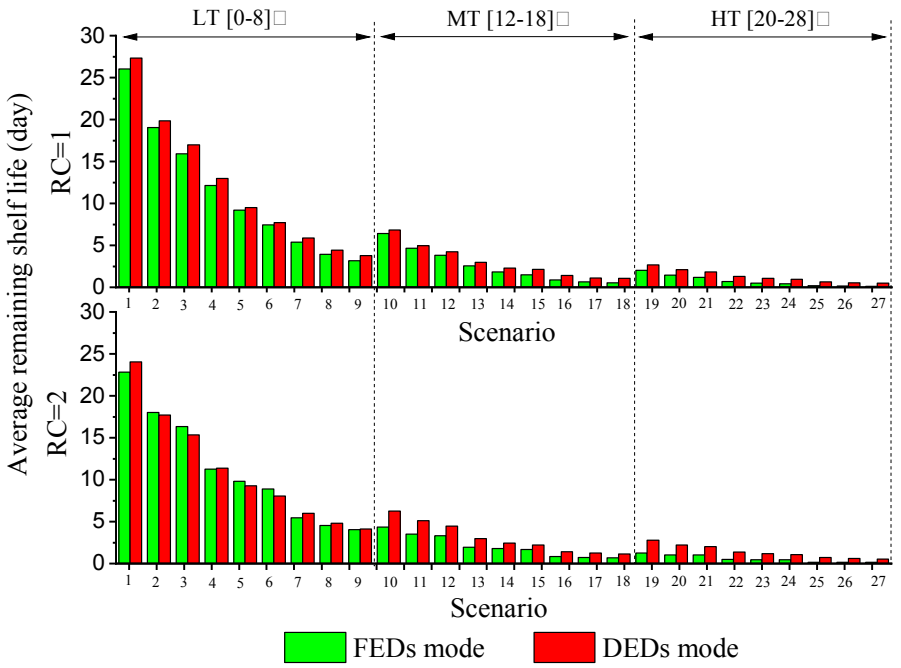


Fig. 2. The average remaining shelf life (SLR_{AV}) under different replenishment cycles.

3.2. Food waste

The food waste rate (FWR) is used to measure variation in the number of food waste at retail to quantify the impacts of the DEDs and the FEDs mode on system performances, which is calculated as food products that are removed from the shelf due to its expiration date. As shown in Fig. 3 (RC=1), the FWR of the FEDs mode is more advantageous than that of the DEDs, except for S.1 to S.5. Because the DEDs mode has a lower change in the quality rate at a lower shelf storage temperature. It is also possible that the backlog of products may be purchased by the consumer (as Section 4.1 explains). Moreover, the FWR of S.1-9 (LT), S.10-18 (MT), and S.19-27 (HT) shows a reciprocating upward trend in the DEDs mode, which indicates that the system performance is influenced in the order of importance by shelf storage temperature, quality change rate, and initial quality of fresh foods. While these factors do not make significant impacts on the FWR of the FEDs mode. However, the FWR of the FEDs mode is significantly increases when the RC is 2 days (Fig. 3). For example, the FWR of S.1 is increased from 0 to 11.62% and from 0.03% to 13.82% in S.10 compared with the 1 day' RC. On the other hand, the FWR change trend of the DEDs mode is not affected by the extension of the RC because consumers choose products at retail based on the actual remaining shelf life instead of the marked date as made in the FEDs mode. Therefore, the use of the DEDs mode may be suggested in reducing food waste on the condition that the perishable foods with a longer RC and the shelf storage temperature could be maintained within an appropriate temperature range.

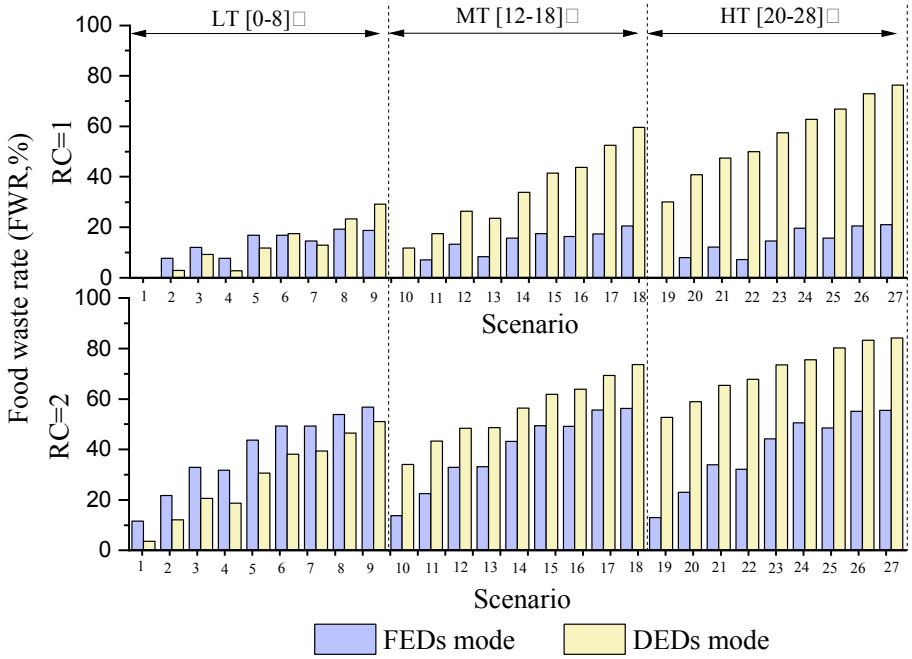


Fig. 3. Food waste rate (FWR) under different replenishment cycles.

3.3. Demand satisfaction of customers

The demand satisfaction rate (DSR) of customers is one of the important characteristics to measure the economic benefits of retail stores. As shown in Fig. 4 (RC=1), the DSR of the DED mode could be up to 90% (S.1-7 and S.10) on the condition that only both lower shelf storage temperature and lower quality change rate, while all 100% for the FED mode when the RC is 1 day. In fact, the DSR varies inversely with the FWR. The increased shelf storage temperature causes more food waste in the DEDs mode, while the FWR of the FEDs mode is still to be at a low level when the RC is 1 day (Fig. 7). Thus, the DSR of the DEDs mode is lower than that of the FEDs mode. In addition, the use of the DEDs mode could reduce the food waste at the appropriate shelf storage temperature range, which in turn improves the DSR when the RC is 2 days. However, the DSR of the FEDs mode is decreased due to the increase in the FWR.

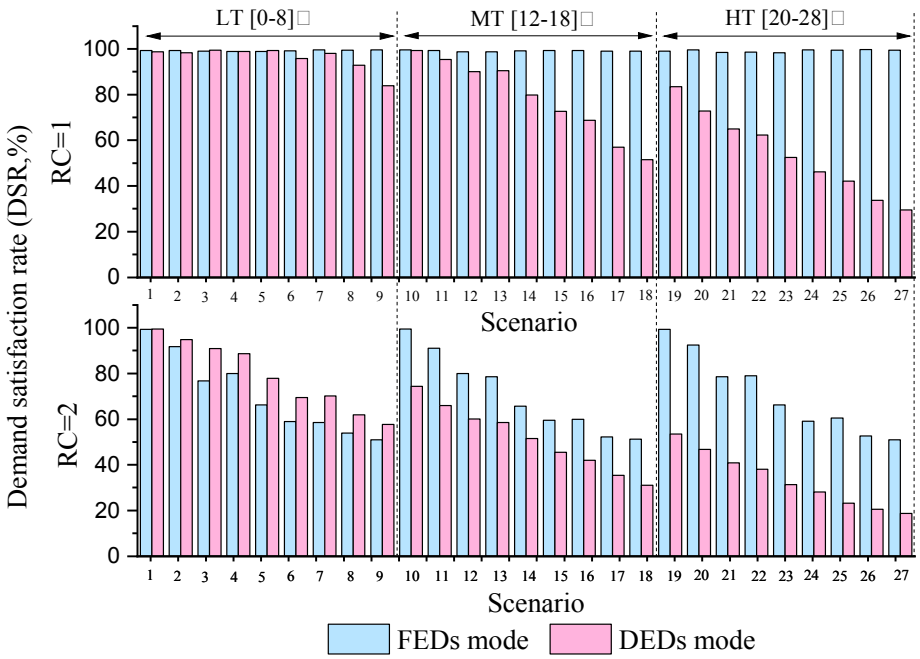


Fig. 4. Demand satisfaction rate (DSR) of customers under different replenishment cycles.

3.4. Food safety

As for the FEDs mode, there is still uncertainty whether the expired food will be sold. As we can see from Fig. 5 (RC=1) that the EFSR increases from 3.97% to 69.8% corresponding to S.6 and S.27, which leads to a great food safety risk. In fact, the shelf storage temperature significantly impacts the performance of the EFSR. For example, the EFSR from S.1 to S.7 (LT group) is close to 0 when the RC is 2 days, while the probability of selling expired food in S.19 - S.27 (HT group) reaches up to 40%. But among the HT group, the EFSR rises rapidly from 14.48% in S.19 to 34.05% in S.21 (RC=1), which indicates that the initial quality of food has a negative influence on the EFSR. This is consistent with the analysis of Section 4.2. In addition, the extension of the RC increases the EFSR of the HT group, for example, the EFSR increased from 14.48% to 44.97% (S.19). However, for food retail with medium shelf storage temperature and high-quality change rate (such as S.17 - S.18), the extension of the RC, in turn, reduces the EFSR. This may be because the quality of the medium shelf temperature group decreases more slowly than the highest shelf temperature group; thus extended the replenishment cycle is conducive to the sale of impending food in the FEDs mode.

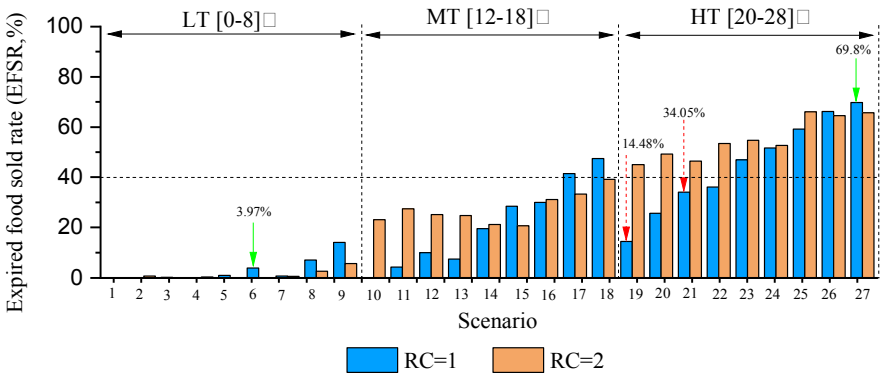


Fig. 5. The expired food sold rate (EFSR) of the FED mode under different replenishment cycles.

4. Conclusion

In this study, the retail benefits from the DEDs system depend on each scenario despite its advanced technology. In the DEDs, the storage temperature of the retail shelf leads to food waste rate, followed by the change rate of food quality and the initial quality of food products. We found that the DEDs system could reduce food waste when food products are stored at lower temperature areas. Besides, the potential of food savings in an extended replenishment cycle is significantly advantageous than the fixed expiration dates (FEDs). On the other hand, the information-sharing approach of the DEDs system is relatively limited in improving sustainable assessment performance of food waste in retail and even mislead consumers' choices.

Reference

Blakeney, M. (2019). Food loss and food waste: Causes and solutions. In *Food Loss and Food Waste: Causes and Solutions*. https://doi.org/10.4337/9781788975391

do Nascimento Nunes, M. C., Nicometo, M., Emond, J. P., Melis, R. B., & Uysal, I. (2014). Improvement in fresh fruit and vegetable logistics quality: Berry logistics field studies. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *372*(2017). https://doi.org/10.1098/rsta.2013.0307

Gaukler, G., Ketzenberg, M., & Salin, V. (2017). Establishing dynamic expiration dates for perishables: An application of RFID and sensor technology. *International Journal of Production Economics*, *193*(November 2016), 617–632. https://doi.org/10.1016/j.ijpe.2017.07.019

Jedermann, R., Nicometo, M., Uysal, I., & Lang, W. (2014). Reducing food losses by intelligent food logistics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *372*(2017). https://doi.org/10.1098/rsta.2013.0302

Jevinger, Å., Båth, K., & Göransson, M. (2014). A Field Test Study on a Dynamic Shelf Life Service for Perishables. *In Proceedings of the Nordic Logistics Research Network (NOFOMA) 2014 Conference*.

Nations United. (2015). *Sustainable Development Goals 2030*. Department of Economic and Social Affairs Sustainable Development.

Salinas Segura, A., & Thiesse, F. (2017). A comparison of sensor-based issuing policies in the perishables supply chain. *International Journal of RF Technologies: Research and Applications*, *8*(3), 123–141. https://doi.org/10.3233/RFT-171672

United Nations Environment Programme. (2021). Food Waste Index Report 2021. In *Unep*.

Yi, T., Jing, X., & Hong, X. (2010). Prediction model of shelf life of Trichiurus haumela using an electric nose. *Transactions of the CSAE*, *26*(2), 356–360. https://doi.org/doi：10.3969/j.issn.1002-6819.2010.02.062

When Food Cultures Meet: The Fur Trade Era on the North American Plains

C. Thomas Shay

***Abstract*—** When cultures meet, so do their foods. Beginning in the seventeenth century, European explorers, missionaries and fur traders entered the North American Great Plains, bringing with them deadly weapons, metal tools and a host of trade goods. Over time, they also brought barrels of their favorite comestibles—even candied ginger. While Indigenous groups actively bartered for the material goods, there was limited interest in European foods, mainly because they possessed a rich cuisine of their own.

***Keywords*—** native Americans, europeans, great plains, fur trade, food.

Tom Shay is with the University of Manitoba, United Kingdom (e-mail: tomshay666@gmail.com).

Illness Experience Without Illness: A Qualitative Study on the Lived Experience of Young Adults During the COVID-19 Pandemic

Gemma Postil, Claire Zanin, Michael Halpin, Caroline Ritter

***Abstract*—** Illness experience research typically focuses on people that are living with a medical condition; however, the broad consequences of the COVID-19 pandemic are impacting those without the virus itself, as many experienced extensive lockdowns, social isolation, and distress. Drawing on conceptual work in the illness experience literature, we argue that policy and social changes tied to COVID-19 produce biographical disruptions. In this sense, we argue that the COVID-19 pandemic produces illness experience without illness, as the pandemic comprehensively impacts health and biography. This paper draws on 30 in-depth interviews with young adults living in Prince Edward Island (PEI), which were conducted as part of a larger project to understand how young adults navigate compliance with the COVID-19 pandemic. We then inductively analyzed the interviews with a constructivist grounded theory approach. Specifically, we demonstrate that young adults living in PEI during the COVID-19 pandemic experienced biographical disruptions throughout the pandemic despite not contracting the virus. First, we detail how some participants experience biographical acceleration, with the pandemic accelerating relationships, home buying, and career planning. Second, we demonstrate biographical stagnation, wherein participants report being unable to pursue major life milestones. Lastly, we describe biographical regression, wherein participants feel they are losing ground during the pandemic and are actively falling behind their peers. These findings provide the novel application of illness experience concepts to the context of the COVID-19 pandemic, contribute to work on illness experience and ambiguity, and extend Bury's conceptualization of biographical disruption. In conclusion, we demonstrate that young adults experienced the biographical disruption expected from having COVID-19 without having an illness, highlighting the depth to which the pandemic affected young adults.

***Keywords*—** illness experience, lived experience, biographical disruption, COVID-19, young adults.

Gemma Postill is with the University of Toronto, Canada (e-mail: gemma.postill@mail.utoronto.ca).

Effects of High-Intensity Microwaves on the Physical and Chemical Characteristics of Fast-Growing Wood

Wang Zhenyu, Xu Enguang, Lin Lanying, Fu Feng

***Abstract*—** Mongolian scotch pine (Pinus sylvestris var. mongolica Litv.) is one of the dominant fast-growing species in northern China. Before the solid utilization of fast-growing wood, an appropriate functional modification is necessary to improve their natural defects.High-intensity microwave (HMW) treatment can be used as an efficient and green way to significantly increase the wood permeability, which benefits the modification process and the applications of renewable fast-growing wood resources. However, researches on the changes that occurred at the scale of macromolecular after HMW treatment were insufficient, resulting in a lack of full understanding of HMW treatment mechanism. In this study, the effects of HMW on the chemical and crystal structure characteristics of wood were investigated, mainly using Fourier transform infrared spectroscopy (FTIR) and X-ray diffraction (XRD) to detect the changes in chemical functional groups and crystallinity, respectively. During HMW treatment, both the microwave intensity and wood moisture content (MC) had a significant influence on wood macromolecular characteristics. From the FTIR spectrum, the intensity of carbohydrate peaks, especially for hemicellulose, decreased after HMW treatment, whereas that of lignin peaks was relatively stable. Although the length of the hydrogen bond didn't show great differences between each HMW groups, with the increase of microwave energy density, the total hydrogen bond energy showed a downward trend in 20% MC samples, while it first decreased and then increased in 30% and 40% MC samples. In XRD tests, the crystallinity and the width of the crystal zone of HMW-treated wood were determined. Compared with control samples, the crystallinity decreased obviously in 20 kWh/m3HMW treatment, followed by a slightly rebounding in 40 kWh/m3due to the crystallization of non-crystalline area in the wood. When the microwave intensity came to 60 kWh/m³, the crystallinity reached the maximum and decreased again in 80 kWh/m³mainly because of the degradation of cellulose.

***Keywords*—** wood, microwave, chemical structure, crystal structure.

Zhenyu Wang is with the China Academy of Forestry, China (e-mail: wangzhenyu1992@126.com).