Cognitive Sustainability

Dec 2023

Vol. 2 | No. 4 | ISSN 2939-5240

Cognitive Sustainability

Cognitive Sustainability (CogSust) is a double-blind peer-reviewed scientific journal published by CogSust Ltd.

(H1116 Budapest Putnok u 9.)

The person responsible for publishing: Mária Szalmáné Csete editor@cogsust.com

The person responsible for editing: Ádám Török info@cogsust.com

CogSust is an online quarterly journal, publication frequency: quarterly, by

March, June, September, December.

ISSN 2939-5240

The journal is indexed by:

Library of Hungarian Scientific Works

Repository of the Library of Hungarian Academy of Science

Google Scholar

Sherpa Romeo

RePEc (Research Papers in Economics)

Directory of Open Access Journals

Crossref

Semantic Scholar

ERIH PLUS

Index Copernicus International

Dimensions

Scilit

Lens

https://doi.org/10.55343/CogSust.85

Motorcycle Police Training and Service Vehicle Development in Hungary

Peter HOMONNAI[1,2]
0000-0003-0585-4657
[1] *Budapest University of Technology and Economics, Faculty of Transport Engineering and Vehicle Engineering*
Department of Transport Technology and Economics
[2] *Police Education and Training Centre*
Budapest, Hungary
homonnai.peter@kjk.bme.hu

Abstract

As a result, effective implementation of the service's mission system will depend on the training development that can change simultaneously with technical development and deliver systemic results in implementing missions. This article examines how police personnel authorised to drive specialised motorcycles use their vehicles and their main functions and analyses seasonal mileage of motorcycle usage. A questionnaire survey of motorcycle police officers enrolled in the Basic Motorcycle Training and License Renewal Training Programme of the Police Education and Training Centre in 2022 provided the data. According to aggregated data, the average annual mileage of the country's total of 218 service vehicles is 3,599 km. The main patrol category, with 104 people (officer, patrol leader, patrol commander, team commander), is significantly better with 6589 km/year mileage, while the 114-strong secondary patrol category (event scene and investigation, district officer, patrol commander, driver) has 2266 km mileage on average in 2022. The results of the survey highlight the low average mileage of motorcycle police. The main reason for this is that for more than 50 percent of trained officers, performing duties on a service motorcycle is only a secondary task. However, the training system for police motorcycle usage must be restructured.

Keywords
motorcycle traffic, questionnaire survey, occupancy study, average mileage, and distinctive signs.

1. Introduction

Motorcycle patrols have been prominent in traffic policing in the past century. The performance of professional tasks in public areas requires that the tasks using service motorcycles be carried out professionally by police personnel trained for this purpose (*Christie, 2020*). In order to achieve this, a flexible, innovative and system-oriented basic training and a further training system must be put in place. Therefore, the Police Education and Training Centre is key in building and managing the training system. This task is particularly difficult, similar to the relationship between schools and the labour market. The question is who to train and how to train them so that they can carry out their tasks as effectively as possible in the future. In order to give precise answers to these questions, it is necessary to explore the problems that arise during the training and to propose solutions. This research aims to reveal what the members of the Hungarian police force who are entitled to drive service motorcycles with distinctive markings do, i.e. what their "main task" is and approximately how many kilometres they drive in a year. The primary purpose of this research is to analyse police motorcycle training. This illustrates the high professional standards expected of motorcycle police officers.

2. Background

Training police officers on motorcycles is important for several reasons, as motorcycles play a unique and crucial role in law enforcement. Here are some key reasons why such training is emphasised:

Manoeuvrability: Motorcycles are more manoeuvrable than traditional police vehicles, allowing officers to navigate traffic more efficiently. This agility is particularly valuable in urban environments, where congestion is common.

Accessibility: Motorcycles can access areas that larger vehicles cannot, making them essential for patrolling parks, pedestrian zones, and other locations where cars cannot easily go. This enhances the police presence and response capability in diverse settings.

Quick Response: Motorcycles can respond quickly to emergencies and incidents, reducing response times. This rapid response is critical when immediate action is necessary, such as accidents, crimes in progress, or other emergencies.

https://doi.org/10.55343/CogSust.85

Cost-Efficiency: Motorcycles are generally more fuel-efficient and have lower maintenance costs than larger patrol vehicles. This can result in cost savings for law enforcement agencies, allowing them to allocate resources more effectively.

Specialised Roles: Some police departments use motorcycles for specialised roles, such as traffic enforcement, escort duty, and event management. Specialised training ensures that officers can perform these roles safely and effectively.

Skill Development: Riding a motorcycle requires specific skills, and training helps officers develop the expertise needed to operate the vehicle safely in various conditions. This includes training in defensive riding techniques and emergency manoeuvres.

Enhanced Pursuit Capabilities: While pursuits should always be conducted safely and within established guidelines, motorcycle officers can more effectively pursue suspects through congested or narrow areas where larger vehicles might struggle.

Adaptability: Motorcycle training equips officers to adapt to different road conditions and environments. This adaptability is crucial in diverse policing scenarios.

In summary, motorcycle training for police officers enhances their capabilities, responsiveness, and engagement with the community. It allows law enforcement agencies to deploy officers effectively in various situations, ultimately contributing to public safety.

2.1. Training conditions

Article 4 of the Instruction of Ministry of Interior18/2016 (VIII.16.) defines the categories of motorcycle training and the entitlements:

(d) Basic motorcycle training: 120 hours of training in preparation for driving a service vehicle using a distinctive sign with a category A licence;

(e) ‘renewal training for the licence for basic motorcycle training means compulsory periodic training of 20 hours of refresher training, repeated annually from the date of completion of the basic motorcycle training, to renew the qualification obtained based on the training referred to in (d);

Upon completing the initial and periodic training specified in points (d) to (e), a driver may ride a service motorcycle with a distinctive sign.

§ 7 (1) Conditions for participation in training and further training for drivers of motor vehicles for internal affairs authorities by category of motorcycle:

(d) basic motorcycle training:

(da) valid category A driving licence;

(db) successful completion of advanced driver training or advanced driver renewal training;

(dd) a valid Group 2 medical certificate;

(de) a valid driving psychology test (PAV I or equivalent as defined by the training body);

(df) passing a motorcycle pre-selection test; and

(dg) planning for a position involving the driving of a motorcycle;

(e) renewal training for the licence for basic training for motorcyclists:

(ea) a valid category A licence;

(eb) successful completion of basic motorcycle training;

(ec) at least six months of accident-free driving experience, without any fault on the motorcycle driver’s part, immediately prior to the start of the renewal training in basic motorcycle training;

ed) a valid Group 2 medical certificate; and

(ee) a valid driving psychology test (PAV I or equivalent as defined by the training body).

https://doi.org/10.55343/CogSust.85

The conditions of participation in such training show that, in the case of an application for motorcycle training, the applicant is already a person with significant previous training and driving experience. After completing advanced driver training or license renewal training, he/she has acquired the right to use distinctive signs in road traffic with a service vehicle (*Buda, 2022*). Driving a motorcycle and the training as well is a cognitive load, meanwhile the training ensure long term sustainability of motorcycle police in Hungary.

Motorcycle training is the most complex level of training from a technical point of view, which places a considerable demand on police officers in terms of theoretical and practical implementation. The execution of tasks in a two-wheeled vehicle is highly dangerous during the performance of official duties, and it is therefore vital that motorcycle training achieves its objectives in the light of the official duties to be performed in the future and that the participants in the training meet the conditions of competence for the outcome.

2.2. Basic motorcycle training

The primary aim of the basic motorcycle training is to train police officers employed by central, regional and local police forces, who are scheduled to drive service motorcycles, to perform their duties using distinctive signs. Furthermore, persons performing motorcycle duties in other bodies under the Ministry of Interior are also trained. The training program provides drivers of service motorcycles with the theoretical and practical knowledge required to handle and drive service motorcycles at a high standard and fully perform general and specific service tasks using motorcycles. Using a police motorcycle is an extreme cognitive load therefore training the usage of these vehicles and persons are essential.

Having completed the training program, the trainees must be able to react promptly and professionally in traffic with rapid situational awareness (*Goodwin et al., 2022*), drive and operate the service motorcycle safely, carry out a motorcyclist’s service tasks and complete the driving test course under time stress within the specified level time (complex professional report):

Table 1. Curricular units and curricular elements of the professional programme

Unit	Topic	Theory [lesson]	Practical [lesson]	Total [lesson]
1	Basic knowledge of training	4	–	4
2	Transport psychology	2	–	2
3	Healthcare, first aid	2	–	2
4	Vehicle physics, motorcycle controls	2	2	4
5	Vehicle stopping, vehicle tracking	4	4	8
6	Driving exercises on the training track	6	56	62
7	Professional practice with the police motorcycle subdivision	4	12	16
8	Driving practice on public roads	1	7	8
9	High-speed cornering techniques on closed race track	1	7	8
10	Practical report	–	2	2
11	Complex professional report (theory, practice)	2	2	4
Total		28	92	120

* Source: Instruction of Ministry of Interior18/2016 (VIII.16.) Training programme 120 hours of basic motorcycle training for service motorcycle drivers.

https://doi.org/10.55343/CogSust.85

2.3. Basic motorcycle training renewal training

This program offers theoretical and practical training for drivers of service motorcycles employed by bodies under the control of the Ministry of Interior. It provides drivers of service motorcycles with the theoretical and practical knowledge they have acquired to handle the vehicle with high proficiency and use it professionally. The training participants should be able to react promptly and professionally in traffic while assessing the situation quickly, driving and operating motorcycles safely, and carrying out their duties safely.

Table 2. Curriculum details

Curriculum element	Content	Theory [lesson]	Practical [lesson]	Total [lesson]
1. Theoretical training	• Knowledge of the law (Highway Code, use of distinctive signs) • Driving theory (motor vehicle physics)	2	–	2
2. Practical training	• Complex tasks based on driving exercises on a training track at low speed	-	8	8
3. Practical training	• Driving exercises on a training track at high speed (driving on an ideal curve)	-	8	8
4. Driving under time stress	• Completion of a set of exercises from a series of practised driving tasks within a limited time	–	2	2
Total		2	18	20

* Source: Instruction of Ministry of Interior18/2016 (VIII.16.) Mandatory annual 20-hour basic motorcycle training for service motorcycle drivers, licence renewal training programme

The successful completion of the training will provide a professional basis which will be further developed in the practical performance of the duties of the service, individually and in groups. The ever-changing and evolving road traffic environment, including the increasing number of powerful vehicles, encourages the police to keep pace with these technical developments. To this end, the fleet of service vehicles, including service motorcycles, has significantly developed recently. Between 2017 and 2021, 70 high-performance BMW 1200 and 1250 RTP motorcycles were delivered to police forces in Hungary, and the National Accident Prevention Committee (OBB) plans to purchase and deliver 50 more vehicles by 2023. It can be seen that there is a will for technical improvement and that the OBB is providing substantial financial resources for this.

At the same time, questions arise about whether the staff training is up to the new technical challenges and how effectively the trained motorcycle officers are driving the service motorcycle. Consideration should be given to the primary (those whose direct task is to provide service motorcycle support) and secondary (those whose non-primary task is to provide service motorcycle support) responsibilities associated with the post determining the motorcycle service’s composition.

3. Methodology and results

I conducted a questionnaire survey as a motorcycle instructor at the Police Education and Training Centre, with the research period running from the first motorcycle licence renewal training this year, 13 April 2022, to the last, 22 August 2022. The main purpose of the questionnaire was to accurately assess a trained police motorcyclist’s effectiveness in performing his/her duties. To do this, it is necessary to know the current national staffing situation, the trained personnel's primary duty assignments, and the service motorcycle's mileage. In the light of this data, we can draw appropriate conclusions for evaluating effectiveness. The survey lasted six months and involved 218 people who completed the questionnaire on paper.

Before presenting the results of the questionnaires, it is important to define what it means to be a motorcycle police officer today (2023) in Hungary. This depends on the department the policeman serves, his position, and his skills (*Sakashita et al., 2014*). During training, the examination system is uniform, but the performance of the duties varies considerably. This depends on the professional experience and skills of the motorcycle officer concerned.

https://doi.org/10.55343/CogSust.85

Table 3. Results of the questionnaire

Group I 104 people	Mileage [km/year]	Group II 93 people	Mileage [km/year]	Group III 21 people	Mileage [km/year]	Total 218 people	Mileage [km/year]
patrol	7314	accident scene investigator	1230				
patrol leader	7621	district officer	2026				
patrol commander	7923	service control	4030				
team leader	3500						
Average	**6589**	**Average**	**2428**	**Average**	**1789**	**Average**	**3599**
Average for those riding less than 5000 km/year (42 people)	**120**	Average for those riding less than 5000 km/year (75 people)	**62**	Average for those riding less than 5000 km/year (11 people)	**740**	Average for those riding less than 5000 km/year (128 people)	**307**

Source: own compilation

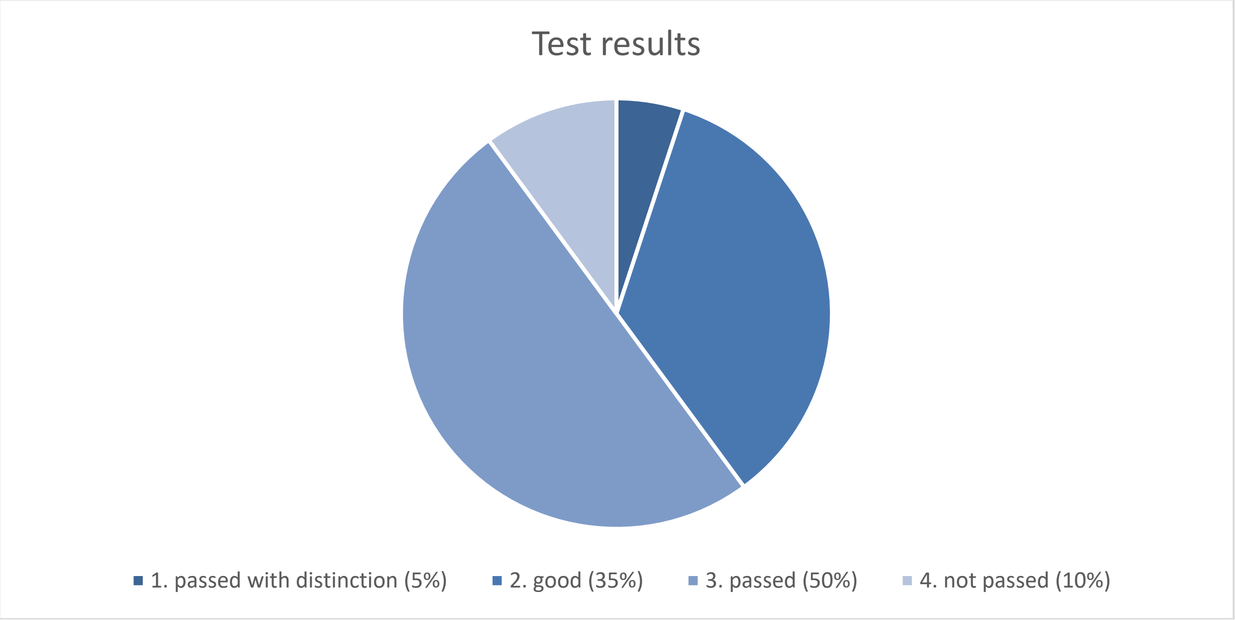


Figure 1. Driving test results
Source: own compilation

4. Analysis & discussion

Based on the evaluation of the questionnaires and the average mileage data, it can be concluded that the average use of service motorcycles by trained police personnel (218 persons) is low (3,599 km/year), of which the average use of service motorcycles by those under 5,000 km/year (128 persons) is critically low (307 km/year). Although the number of service motorcycles (265) is adequate for the number of police officers, and these vehicles are constantly being upgraded and replaced by the OBB, the lack of training and the drastic drop in the level of training cause significant problems in the practical implementation of tasks due to the low mileage.

More than 50 % of the 218 police personnel entitled to drive a service motorcycle with a distinctive sign have an average mileage of 307 km/year, which is extremely low. Based on my professional experience, at least 5000–10,000 km of annual service motorcycle driving is required to maintain the routine and develop the driver's skills. A good basis for comparison would be the mileage rate of civilian motorcyclists in Hungary. This is put at 6000 kilometres by several civilian trade journals, but this is more likely to be based on the service interval of motorcycles. Authoritative sources such as the Central Statistical Office (KSH) or the Public Roads Nonprofit Ltd. database do not include the mileage of motorcycles, only the number of units or cross-sectional passages.

https://doi.org/10.55343/CogSust.85

Driving a service motorcycle requires highly professional experience, so the training is based on the basics. The training programme will not achieve its purpose if the required professional experience cannot be achieved. I therefore propose a compulsory minimum of 5,000 km/year of service motorcycle driving, which can be completed in a specific format, in the home area, regionally or in a group. It is also necessary to consider the regional services' workload in other directions, but the level of qualification can only be maintained and improved if regular driving practice is implemented.

5. Conclusion

For training police officers who can use the distinctive marking on their motorcycles, 20 hours of closed course training per year can be considered a low number of hours (Table 2). With the amendment of the Ministry of Interior instruction in 2021, the 80 hours of basic motorcycle training was converted into 120 hours by including professional practice in the increased number of hours that could lead to more sustainable training.

This made it possible for the personnel involved in the practical training to perform joint service with the Motorcycle Subdivision of the Budapest Police Headquarters in the inner area of Budapest and its surrounding area. The results of the basic training are encouraging, with the police officers who participated in the new type of training completing the complex professional report with significantly greater efficiency and greater tolerance on cognitive load.

Following this improvement practice, I propose to increase the length of the motorcycle licence renewal training from two days (20 hours) to three days (30 hours). The increased number of hours (10 lessons) would also create joint professional practice opportunities.

The research results highlight that the average mileage of a police officer on a motorcycle is low. The aim is to ensure that the performance of the duties of the service, in particular police escorting and traffic policing (vehicle enforcement) tasks, are carried out at a high standard. In order to achieve this, the current closed-course training system needs to be adapted to include at least 10 hours of practical training in addition to the basic tasks set out in the training programme. This would also mean extending the training without placing an undue burden on regional and central bodies. The implementation of practical training in public places could also have a reinforcing effect on the police.

Another important factor is the significant technical improvements made in recent years. Police officers must update their technical knowledge to operate the 120 high-performance service motorcycles acquired by the police between 2017 and 2023. In other words, technical development implies the need for training development.

The development of further training in the renewal of motorcycle licences and the introduction of hours of practical training in public areas is justified. Raising the level of qualification is a prerequisite for increasing the driving experience of motorcyclists, and I therefore consider it necessary to set a minimum mileage. The professional programme aims to ensure a high level of service.

References

Buda K. (2022). Driving training in law enforcement today and tomorrow [in Hungarian: Vezetéstechnikai képzések a rendvédelemben ma és holnap]. Faculty Student Conference, Faculty of Law Enforcement, National University of Public Service [in Hungarian: Kari Tudományos Diákköri Konferencia. Nemzeti Közszolgálati Egyetem, Rendészettudományi Kar]. Budapest.

Christie, N. (2020). Managing the safety of police pursuits: A mixed method case study of the Metropolitan Police Service, London. *Safety Science*. 129, 104848. DOI: https://doi.org/gkdvs4

Goodwin, A. H., Wang, Y. C., Foss, R. D., Kirley, B. (2022). The role of inexperience in motorcycle crashes among novice and returning motorcycle riders. *Journal of Safety Research*. 82, 371–375. DOI: https://doi.org/k34c

18/2016 (VIII.16.). Ministry of Interior Instruction. Laying down the training requirements for drivers and instructors of service vehicles and service vessels used by the Ministry of Interior [in Hungarian: A belügyi szerv által használt szolgálati gépjárművet és szolgálati hajót vezető és a gépjárművezetést oktató személy képzési követelményeinek megállapításáról szóló 18/2016. (VIII. 16.) Belügyminisztérumi utasítás.] URL: https://net.jogtar.hu/jogszabaly?docid=A16U0018.BM&txtreferer=00000003.TXT

Sakashita, C., Senserrick, T., Lo, S., Boufous, S., de Rome, L., Ivers, R. (2014). The Motorcycle Rider Behavior Questionnaire: Psychometric properties and application amongst novice riders in Australia. *Transportation Research Part F: Traffic Psychology and Behaviour*. 22, 126–139. DOI: https://doi.org/f5rzz5

https://doi.org/10.55343/CogSust.91

Production of Light Naphtha by Flash Distillation of Crude Oil

N. Manuel
0000-0002-4872-3123
Instituto Superior Politécnico de Tecnologias e Ciências (ISPTEC), Departamento de Engenharias e Tecnologias (DET) - Av. Luanda Sul, Rua Lateral Via S10, Talatona – Município de Belas, Luanda – Angola

A. A. C. Barros
0000-0001-5922-5368
Instituto Superior Politécnico de Tecnologias e Ciências (ISPTEC), Departamento de Engenharias e Tecnologias (DET) - Av. Luanda Sul, Rua Lateral Via S10, Talatona – Município de Belas, Luanda – Angola
E-mail: chivanga.barros@isptec.co.ao

Abstract:
The escalating global demand for sustainable, renewable energy sources has catalyzed research into advanced refining processes, particularly focusing on the generation of essential raw materials like light naphtha. This paper delves into the production of light naphtha via the flash distillation of crude oil, underlining its critical role in fulfilling the petrochemical industry's requirements. The research delineates the fundamentals of flash distillation, incorporating both physical and mathematical models, with a specific emphasis on Raoult's Law, to demystify the hydrocarbon separation process in crude oil. The mathematical framework encompasses material and energy balances, coupled with the application of liquid-vapor equilibrium equations, thereby shedding light on the thermodynamic principles steering this procedure. The progression of this study involved analyzing the distillation curve of the referenced crude oil and identifying the pseudo-components representative of the involved hydrocarbons. Utilizing the DWSIM commercial simulator, the flash distillation process of Nemba crude oil was simulated, taking into account diverse operational conditions and thermodynamic models. The outcomes were juxtaposed with industrial datasets, demonstrating a significant alignment and affirming the simulation's predictive efficacy. A parametric sensitivity analysis was conducted to refine light naphtha yield, elucidating the effect of variables like the temperature of crude oil feed and naphtha separator feed on the process efficiency. The response surface methodology underscored the feasibility of augmenting naphtha production under certain conditions. Furthermore, this study assessed the impacts of employing two distinct thermodynamic models – the Soave-Redlich-Kwong (SRK) and Peng-Robinson (PR) equations of state. This analysis revealed minimal discrepancies in physical parameters, confirming the simulation's robustness and the reliability of both models. The findings of this paper validate the efficacy of flash distillation in light naphtha production, with numerical simulation emerging as a potent tool for enhancing process optimization and predictive analysis. These insights not only contribute to a deeper understanding of environmental sustainability but also offer strategies for maximizing light naphtha production in the oil industry. In conclusion, the results from this investigation significantly advance our comprehension of phase equilibrium profiles and their relationship with the applied thermodynamic state equations, thereby enriching the quality of the results.

Keywords:
Light Naphtha, Flash Distillation, Crude Oil, Simulation, DWSIM.

1. Introduction

From the perspective of energy production, the quest for cleaner and more sustainable fuels has remained a persistent challenge for governments and researchers worldwide. This endeavor stems from a dual concern: firstly, the pursuit of alternative energy sources is driven by the imperative to address global warming, reduce the carbon footprint, and mitigate environmental pollution associated with the utilization of fossil fuels. Secondly, there is a recognition of petroleum as a finite and non-renewable resource (*Handogo, 2021; Noriler et al, 2009, Wolf-Maciel et al, 2001*).

Another pivotal facet of this endeavor lies in the identification of exceptionally valuable raw materials that facilitate the development of novel industrial processes. In this context, the transformation of crude oil into naphtha to meet the exigencies of the petrochemical sector assumes paramount significance in fostering and sustaining a diverse array of industrial activities.

Naphtha, as a constituent of crude oil, predominantly comprises hydrocarbons featuring carbon chains ranging from 6 to 12 carbons in length. It serves as a foundational feedstock for the production of an extensive spectrum of chemical products, including plastics, most notably polyethylene and polypropylene, which are prominent commodities in the global market. Naphtha is procured through a direct distillation process and encompasses hydrocarbons with boiling points within the temperature range of 38°C to 200°C. Its versatility extends to its use in gasoline production and as a fundamental raw material for the petrochemical industry. The latter application involves its conversion into ethylene, propene, butadiene, benzene, toluene, and xylenes through processes such as thermal cracking and catalytic reforming. The specific classification of

https://doi.org/10.55343/CogSust.91

naphtha into light or heavy variants in refineries is contingent upon the cut-off point determined within the atmospheric distillation column (*ABIQUIM, 2007*).

This dynamic landscape has spurred investments in scientific research geared towards developing processes that render naphtha a viable source for the production of butadiene and aromatics. Consequently, the quality of naphtha is predicated upon its hydrocarbon composition, boiling point characteristics, and the presence of impurities. As delineated by *Rahimpour et al (2013)* and *Prauchner et al (2022)*, the distribution of paraffins, olefins, naphthenes, and aromatics serves as a decisive determinant of naphtha quality. This is particularly pertinent in the context of catalytic reforming processes, where naphthenes undergo selective transformation into aromatics.

Numerous research efforts are being directed towards finding sustainable solutions for liquid fuels, with a significant focus on renewable feedstock-based alternatives. Among these, the development of biodiesel has a well-established history, as evidenced by studies from *Lulic et al. (1998)*, *Alahmer et al. (2022)*, and *Jeyakumar et al. (2022)*. Another major area of investigation is the potential use of alcohols as diesel substitutes, which has been intensively researched (e.g., *Tutak et al., 2015*). Additionally, waste-based fuels, including pyrolytic feedstocks, have garnered attention for their potential in sustainable fuel production (*Kondor et al., 2022*).

While exploring alternatives to traditional fossil fuels, researchers have also examined options like OME (oxymethylene ethers), which, despite their fossil origin, present as viable alternative liquid fuels (*Virt and Arnuld, 2022*). Beyond liquid fuels, there's a growing interest in gaseous fuels like CNG (compressed natural gas), as highlighted in the work of *Matijosius et al. (2022).*

A notable contribution in this field is the research by *Goldbach et al. (2022)* and *Da Silva Mateus et al (2022)*, which demonstrates that fast pyrolysis can be an effective method for large-scale biofuel production from waste fish oil, yielding products with properties akin to conventional petroleum fuels. While there have been practices of blending diesel oil with vegetable or waste oil as an alternative, these blended fuels often fail to meet the stringent fuel quality standards set by governments and engine manufacturers. Issues such as corrosion and polymerization are common medium-term consequences of using these blends. However, biofuels produced through pyrolysis show promise due to their compatibility with current engine standards, addressing these concerns.

Furthermore, naphtha has emerged as a precursor in the synthesis of an extensive spectrum of chemical products, with notable applications in the realm of resins, solvents, and pharmaceuticals. Additionally, it assumes a pivotal role as a key constituent in the formulation of high-octane gasoline formulations.

1.1. Principle of Flash Distillation

Flash distillation of petroleum represents a widely employed and intricate separation process conducted within flash tanks, amenable to description through both physical and mathematical models, as illustrated in *Figure 1* :

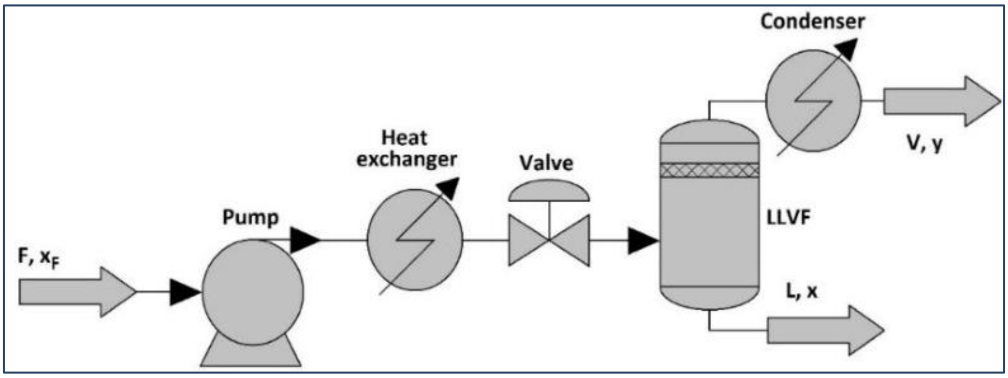


Figure 1 Flowchart of a flash distillation unit.

The flash tank, a single-stage distillation unit, plays a pivotal role in the separation of hydrocarbons within petroleum. It operates by receiving a preheated liquid hydrocarbon stream that passes through an expansion valve before entering the flash tank. Within this tank, a sudden pressure drop occurs, leading to the separation of the liquid and vapor phases. Given the diverse range of hydrocarbons present in the incoming crude oil, the vapor stream emerging from the top of the flash tank predominantly consists of lighter hydrocarbons, falling within the category of light naphtha. In contrast, the descending

https://doi.org/10.55343/CogSust.91

liquid stream contains hydrocarbons with higher carbon structures, necessitating further fractionation within an atmospheric distillation column (*Pasquini and Bueno, 2007*).

For the elucidation of this separation process, the physical model relies on Raoult's Law. This law expounds upon the principles of liquid-vapor equilibrium within ideal liquid mixtures, with a primary focus on vapor pressure and mole fraction of constituent components as the principal process variables. Raoult's Law affords insights into the phase equilibrium governing the various components of crude oil at varying temperatures, a fundamental aspect of flash distillation.

In this context, a comprehensive grasp of phase equilibrium principles underpins the predictability of vaporization rates during the depressurization process within flash distillation columns. This entails a profound understanding of the key physical properties of crude oil components, encompassing boiling temperatures and vapor pressures. Establishing a robust relationship between these properties and the principles of phase equilibrium is instrumental in enhancing vaporization rates and facilitating the recovery of lighter components.

Moreover, the physical model encompasses the tenets of heat and mass transfer, particularly concerning the preheating of crude oil and the ensuing vaporization. This entails the application of mass and energy balances, intertwined with the thermal exchanges inherent to the process.

Consequently, the mathematical model associated with this process emerges from material balances as the foundational framework for modeling component flow within a flash distillation process. These equations enable the determination of mass flow rates for both the distillate (vapor) and the bottom product streams (as represented by Equation 1) (*Luyben, 2006*) :

(1) $$F = V + L$$

Where F, V and L are the feed, top product and bottom product flow rates, respectively, in kmol/h.

For a given component, mass balances can be drawn up using Equation (2).

(2) $$Fx_F = Vy + Lx$$

Where x_F, y and x are the fractions of the components in the feed, at the top and at the bottom of the flash distillation column, respectively.

Energies balances are used to calculate the heat transfer rate and temperature changes in the distillation process and are essential for determining energy requirements and ensuring separation efficiency (Equation 3).

(3) $$Fh_F + Q = Lh_L + VH_V$$

Where hF, hL and HV are the feed, bottom product and top product enthalpies, respectively, in kj/h, and Q, in kj/h, is the energy added to the system that results in the heating of the crude oil stream to ensure phase separation in the flash tank.

On the other hand, liquid-vapor equilibrium equations, based on Raoult's Law or other phase equilibrium models, are used to predict the composition of the vapor phase at a given temperature and pressure. These equations make it possible to understand the distribution of the different components between the liquid and vapor phases during flash distillation (Equation 4) (*Reis et al, 2006 and Barros (2022)*) :

(4) $$y_i = \frac{x_i \gamma_i P_i^{sat}}{\phi_i P}$$

Where P^{sat} is the vapor pressure calculated based on the equation proposed by Antoine (Equation 5) and γi and ϕi are, respectively, the activity and fugacity coefficients associated with phase equilibrium in the liquid and vapor phases :

(5) $$\ln P^{vap}(KPa) = A_i - \frac{B_i}{T(°C) + C_I}$$

Characterized as a separation technique based on the principles of phase equilibrium, this process guarantees the effective recovery of the light components contained in oil, especially light naphtha. In addition to the equilibrium principles described by Raoult's law, the equilibrium equation (Equation (6)) proposed by *Lewis* can be used (*Perry, 2018; Kalvelage et al, 2017*).

(6) $$y_n = \frac{\alpha x_n}{1 + (\alpha - 1) * x_n}$$

Where α is the relative volatility that relates the vapor pressures of the light and heavy key components.

https://doi.org/10.55343/CogSust.91

Flash distillation calculations serve as a crucial tool for determining the requisite thermal (temperature) and mechanical (pressure) conditions necessary to achieve phase equilibrium between the exiting streams from the flash tank. These calculations are inherently iterative in nature and involve the simultaneous resolution of mass and energy balances, in conjunction with phase equilibrium equations. This integrated approach is indispensable for solving a system of differential equations, which is pivotal in evaluating the behavior of the flash distillation column (*Shuncheng and Bagajewicz, 2002*).

The resolution of flash distillation problems can be executed through the utilization of commercial simulators or the development of algorithms grounded in mathematical modeling, which rely on the formulation of mass and energy balances. Numerous computational tools are available for this purpose, including Aspen HYSYS, CHEMSEP, DWSIM, and others (*Leite et al, 2005*). These tools enable the simulation of flash distillation processes, as well as more complex separation processes, providing valuable insights into the separation of crude oil into various fractions, including naphtha. Additionally, alternative thermodynamic models may be explored to enhance process performance in terms of separation capacity for specific crude oil fractions. Within this context, *Kister (1992)* and *Fraga (2010)* noted that flash distillation offers notable advantages such as reduced energy consumption, lowered capital costs, and heightened process sustainability when compared to conventional distillation methods.

Handogo (2021) has conducted studies demonstrating the widespread utilization of flash distillation in refineries and oil processing plants to segregate light components, such as naphtha, from the heavier constituents of crude oil. According to the author, flash distillation effectively achieves the swift separation of light naphtha from the heavier components by capitalizing on differences in relative volatility and vapor pressure. It is noteworthy, however, that the precise composition of the light naphtha can vary contingent upon the type of crude oil being processed and the operational conditions of the flash distillation column. Subsequent to the separation of streams within the flash tank, the light naphtha can undergo additional processing stages, particularly sulfur recovery and other refining processes, to meet the requisite quality specifications and fulfill the demands of the chemical and petrochemical industries that rely on this essential raw material.

In this specific context, a flash distillation study was undertaken on Nemba crude oil for the production of light naphtha, employing the support of the commercial simulator DWSIM. This investigation involved the exploration of optimal operating conditions and thermodynamic models to ensure peak process performance, with the resulting data subjected to comparison with existing literature data.

2. Data and methods

2.1 Software

The research conducted in this study facilitated a parametric evaluation of the light naphtha production process via flash distillation of Nemba crude oil. The study harnessed the capabilities of the commercial simulator DWSIM, version 8.6.1. The primary focus of this investigation was to systematically analyze the process parameters, with a specific emphasis on assessing operating conditions and thermodynamic models that could enhance process performance.

The utilization of process simulation provides a powerful means to gauge the sensitivity of various parameters in separation operations and other industrial processes. As elucidated by *Leite et al. (2005)*, who employed a commercial simulator to investigate physical and chemical gas absorption processes, the flexibility afforded by commercial simulators greatly expands the analytical capacity for industrial processes. It enables the identification of critical parameters that exert the most significant influence on the outcomes of these processes, thereby facilitating informed decision-making and optimization efforts.

2.2. Characteristics of Simulated Oil

The oil used in this simulation study was extracted in Angola, and is coded as Nemba. Its characteristics are as follows: Density @ 15ºC: 0.8394 g/cm³; API gravity @ 60ºC: 37.0 degrees API (American Petroleum Institute); Viscosity @ 10ºC: 13.3 cSt and Acidity: 0.24 mg KOH/g.

On the other hand, the distillation curve for Nemba crude oil is shown in Figure 2. For this configuration, *Prauchner et al. (2022)* point out that the constituent compounds of petroleum cover very wide boiling temperature ranges, as illustrated in the distillation curve in *Figure 2*, which describes the main way of characterizing a petroleum sample. Generally, paraffinic compounds are concentrated mainly in the low and medium boiling point fractions, while naphthenic and aromatic compounds are present in the heavier fractions, with the usual concomitant occurrence of two or more of the aforementioned classes of compounds, as illustrated in *Figure 2* :

https://doi.org/10.55343/CogSust.91

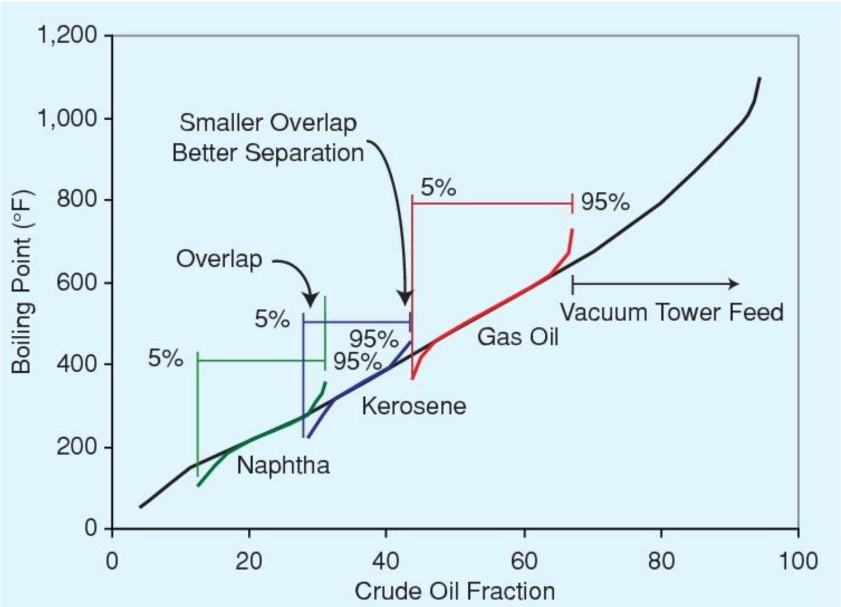


Figure 2 Main crude oil derivatives by boiling temperature range.

The naphtha from Namba Crude Oil was characterized by cutting it in two temperature ranges, i.e. 15-65°C and 15-80°C, respectively. Once the cuts had been made, the physicochemical characterization was carried out in TOTAL Energy's laboratory, the results of which are shown in *Table 1*:

Table 1 Result of cutting light naphtha at Crude Oil Namba.

Parameters	Unity	Results	
Cutting temperature range	°C	15-65	15-80
Yield	wt %	4.95	6.64
	vol %	6.40	8.41
Density at 15°C	kg/m^3	648.00	661.00
Sulphur content	wt %	0.0001	0.0001
Mercaptans content (RSH)	mg/kg	0.00	0.00
Naphthenic compound content	vol %	7.90	14.40
Aromatic compounds content	vol %	0.70	2.20
RON	-	74.20	71.50
MON	-	72.70	69.70

Where RON and MON stand for Research Octane Number and Motor Octane Number respectively. The data in Table 1 was compared with simulation data, based on the use of two cubic equations of state, which made it possible to assess the consistency of these models in analyzing the behavior of flash distillation columns.

2.3 Procedures Used for the Simulation

This work was carried out at the Instituto Superior Politécnico de Tecnologias e Ciências (ISPTEC), with the support of the physical and operational parameters of Nemba crude oil shown in Table 1. The simulation studies involved the following sequence of procedures:

i. **Definition of the operating variables:** the critical operating variables were established, covering parameters such as flow rates, temperatures, pressures and compositions. These variables play a fundamental role in configuring the simulation model.

https://doi.org/10.55343/CogSust.91

ii. **Pseudo-component selection:** in the simulation process, the initial steps involved using the distillation curve to define the pseudo-components of the crude oil mixture, the results of which were used in the simulation to represent the multi-component nature of the feedstock used in flash distillation.

iii. **Selection of the thermodynamic model:** the Peng Robinson model was selected for this process, as it enables the phase equilibrium behavior of substances in high pressure and temperature systems to be described.

iv. **Drawing up the flowchart:** DWSIM software was used to describe the process flowchart, which describes the sequence of equipment and respective operations in the flash distillation of crude oil for the production of light naphtha.

v. **Definition of operating conditions**: the feed conditions were then defined, with emphasis on the flow rate, temperature and pressure. Once these conditions had been established, the simulation was carried out, which essentially consisted of closing the mass and energy balances in order to meet the specifications previously established.

vi. **Evaluation of the results:** the results obtained in the simulation were evaluated to check the consistency and validation of the thermodynamic model and to identify any discrepancies or needs for improvement.

Therefore, the simulation reproduces the industrial processes of flash distillation of crude oil to produce light naphtha, the results of which can be evaluated in terms of consistency and performance by comparing the simulated data with industrial data.

2.4. Experimental planning

Flash distillation involves the intensive use of energy, a factor that affects operating costs in these industrial units. To this end, parametric sensitivity analysis makes it possible to explore the behavior of the operations involved in flash distillation columns. This condition resulted in the elaboration of the 22 factorial experimental design, used to evaluate the effect of two independent factors on a response variable. For this case, each factor has two configuration levels.

For this study, two independent variables were used (crude oil feed temperature – X_1 and naphtha separator feed temperature – X_2) and one dependent or response variable (flow rate of the light naphtha produced), a condition that guarantees the optimization of light naphtha production. Table 2 shows the experimental planning used in this study, based on the descriptions of the variables described above.

Table 2: Experimental design 22 used in this study

Experiment	X_1	X_2
1	-1	-1
2	1	-1
3	-1	1
4	1	1
5	0	0

In this way, 5 experiments were carried out to evaluate the variables described in this work and the responses evaluated in terms of comparative consistency with the experimental data.

3. Results and Discussion

The simulation first involved using experimental data from the distillation curve and API of the Nemba crude oil as input data for the DWSIM simulator and then choosing the operating conditions and thermodynamic model. Once this stage was completed, the characteristic pseudo-components of the crude were generated and used to carry out the numerical simulation study carried out in this work.

https://doi.org/10.55343/CogSust.91

Table 2 Pseudo-components generated from Nemba crude oil

Pseudo-components	Mass Fraction	Boiling Temp. (K)	SG	Molar weight (Kg/Kmol)
C_5181_NBP_-51	0.0846	222.444	0.5732	37.90
C_5181_NBP_18	0.0941	291.278	0.6691	62.24
C_5181_NBP_92	0.1771	364.875	0.7410	96.64
C_5181_NBP_141	0.2347	414.628	0.7787	125.00
C_5181_NBP_213	0.1217	486.550	0.8234	173.53
C_5181_NBP_282	0.1016	555.138	0.8583	228.39
C_5181_NBP_346	0.0800	618.848	0.8860	287.04
C_5181_NBP_412	0.0477	685.163	0.9112	356.12
C_5181_NBP_481	0.0315	753.674	0.9347	436.19
C_5181_NBP_550	0.0271	823.477	0.9567	526.86

Pseudo-components represent a complex mixture of hydrocarbons with a boiling temperature within an established temperature range. For this case, the pseudocomponents individually represent a discrete component, with an average boiling point temperature and average properties throughout the mixture (*Quirino, 2009*). In this way, the sum of the fractions of all the pseudocomponents constitutes the total sample of crude oil from the referenced crude.

3.1. Simulation

In this study, the mass and energy balances were carried out using numerical simulation, supported by the DWSIM simulator, whose procedure followed the one described in the methodology of this work and involved preheating, pre-flash distillation, stabilization and fractionation of the naphtha, in order to produce products with a quality similar to that produced in the industrial unit, referred to here in this work.

The operating conditions involved a feed flow rate of 1000 Kg/s of crude oil, with the Peng Robinson cubic equation of state used as the thermodynamic model for the gas phase. In this way, the crude oil was heated to 274°C (548.15 K) and a pressure of 4.2 bar, a condition that guarantees the total vaporization of the components associated with light naphtha and, partially, the components of heavy naphtha. Thus, in this operating condition, the stream enters the flash distillation column when there is a sudden reduction in pressure to approximately 1.2 bar, resulting in the separation of the descending liquid and ascending vapor phases in this industrial equipment.

Also in this process, the ascending vapour stream is cooled in a heat exchanger to 80°C (353.15 K), and then feeds the second flash distillation column, a condition that guarantees the separation of LPG as ascending top vapour and descending liquid naphtha in this industrial equipment. The second flash distillation column is therefore characterised as a stabilisation tower and is used to produce light naphtha (C4-C6), the results of which are shown in Table 3.

Based on the data in Table 3, it can be seen that it is possible to produce up to 231.72 tonnes/day of light naphtha, a condition that could increase the rate of petrol production through catalytic reforming or the use of light naphtha as a raw material for the petrochemical industry. On the other hand, the efficiency of light naphtha production from this simulation study is on a par with TOTAL ENERGY's experimental data (Table 3), with an average deviation of 3.06%.

Numerical simulation is therefore a fundamental tool for analysing the performance of industrial processes, capable of predicting the behaviour of separation equipment, with deviations obtained within the range established in the literature. For these cases, Fernandes (2018) states in his studies that deviations between experimental data of up to 5% are admissible because they are related to the characteristics of the mathematical and thermodynamic models used in these processes.

https://doi.org/10.55343/CogSust.91

Table 3: Results of the mass flow of products obtained in the simulation.

Product	Boiling Temperature (°C)	Mass Flow		Recovery Percentage (%)
		Feed (Kg/s)	Output (Kg/s)	
Crude oil	---	1000	---	---
LPG	15 – 80	---	6.473	0.647
Light Naphtha	15 – 80	---	64.366	6.437
Heavy Naphtha	80 –175	---	390.075	39.007
Atmospheric Current	175 –370	---	539.086	53.909
Balance (Input = Output)		1000 Kg/s = 1000 Kg/s		100%

The parametric simulation analysis based on the experimental planning made it possible to evaluate the effects of the crude oil feed temperature and the naphtha heating temperature before the flash separator or distiller. The results obtained were plotted on a response surface graph (Figure 5) and show that the increase in naphtha production rates is related to the maximisation of the experimental planning parameters mentioned here. Unlike the results in Table 3, Figure 5 defines the best parametric operating conditions and, as mentioned above, it can be seen that temperatures close to the maximum admissible boiling temperatures of the light naphtha components (C4-C6) guarantee higher yields in terms of flow rates or light naphtha production.

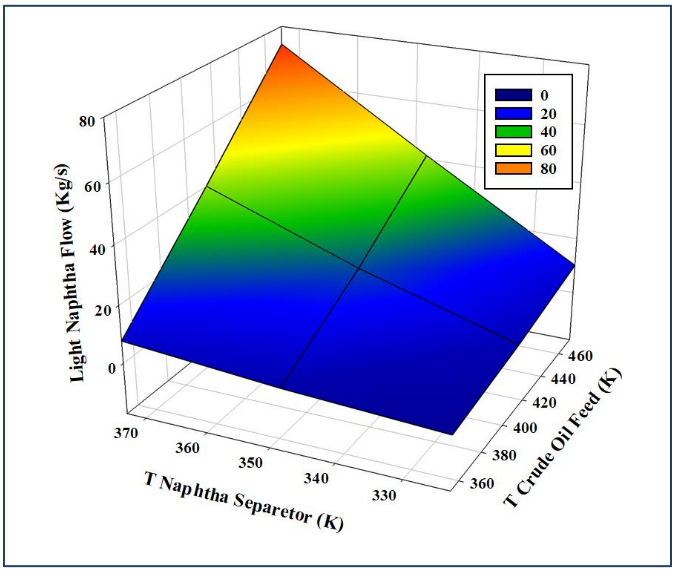


Figure 5: Response surface for optimising light naphtha production.

3.1.1 Analysis of thermodynamic parameters

To better understand the influence of the thermodynamic model in terms of response quality, the flash distillation was simulated using two cubic equations of state, Soave-Redlich-Kwong (SRK) and Peng-Robinson (PR) respectively. In order to analyse the data obtained, the physical parameters of the products of the ascending vapour stream of the stabilised flash distillation were compared, according to the results shown in Table 4. Overall, it can be seen that there is a numerical similarity between the parameters derived from the two equations of state, with minimal deviation between them, even considering that the LPG produced involves a multicomponent mixture, which interferes differently, based on the molecular interaction forces, characterised by repulsion and attraction of the molecules making up the mixture. For this analysis,

https://doi.org/10.55343/CogSust.91

thermal conductivity showed the greatest deviation, followed by specific mass, volumetric flow, internal energy and molar internal energy. This behaviour is due to the mathematical similarity between the two thermodynamic models, and because they consider the acentric factor which relates the size of the molecules when evaluating the molecular forces involved in the process.

Table 4 Results of the mass flow of products obtained in the simulation.

Parameters	Unity	SRK	PR	Variation (%)
Thermal Conductivity	W/[m.K]	0,01949	0,01947	0,15%
Specific Mass	kg/m^3	2,41777	2,41957	0,07%
Internal Energy	kJ/kg	67,28390	67,31500	0,05%
Molar Internal Energy	kJ/kmol	4874,56000	4876,82000	0,05%
Helmholtz Free Energy	kJ/kg	-120,65000	-120,65100	0,00%
Molar Gibbs Free Energy	kJ/kmol	-5704,68000	-5707,01000	0,04%
Specific Enthalpy	kJ/kg	109,19200	109,19200	0,00%
Molar Enthalpy	kJ/kmol	7910,72000	7910,72000	0,00%
Specific Entropy	kJ/[kg.K]	0,50359	0,50373	0,03%
Molar Entropy	kJ/[kmol.K]	36,48360	36,49390	0,03%
Gibbs Free Energy	kJ/kg	-78,74210	-78,77420	0,04%
Molar Helmholtz Free Energy	kJ/kmol	-8740,84000	-8740,91000	0,00%
Volumetric Flow @ T, P	m^3/s	26,62220	26,60240	0,07%

On the other hand, the analysis of light naphtha from the second stabilised flash distillation column requires the use of thermodynamic models for the liquid phase, which makes it possible to determine the physical parameters of this phase and consequently compare the data obtained with industry data, a condition that can favour an understanding of the phase equilibrium principles involved in this study.

4. Conclusion

Based on the results obtained from this work, it can be concluded that:

a) Flash distillation is an effective operation capable of predicting the operational performance of processes for producing light naphtha from crude oil, whose separation capacity depends on the boiling points of the constituent components of the mixture;

b) Process simulation makes it possible to predict the behaviour of each operation and to know the energy content involved, conditions that make it possible to evaluate various mathematical and thermodynamic models that fit the experimental data;

c) The simulation carried out using the DWSIM simulator made it possible to predict naphtha production rates and the results were compared with those of TOTAL ENERGY's industrial processes, with an average deviation of 3.06%;

d) Flash distillation, based on the principles of phase equilibrium, can be numerically optimised to guarantee knowledge of environmental sustainability, which is necessary for the national oil industry, maximising light naphtha production rates.

Acknowledgement

The authors of this paper express sincere appreciation for the collaboration of the Post-Graduation Coordination and the Undergraduate Coordination of the Chemical Engineering Program at ISPTEC, for their unwavering dedication and support in the development of this research.

https://doi.org/10.55343/CogSust.91

References

ABIQUIM (Brazilian Chemical Industry Association). Demand for petrochemical raw materials and probable origin up to 2020. 2007.

A. A. Chivanga Barros, Evaluation of extractive distillation using efficiency correlation and experimental data, Studies in Engineering and Exact Sciences. 3 (2022) 737–754, https://doi.org/10.54021/seesv3n4-012.

Alahmer, A., Rezk, H., Aladayleh, W., Mostafa, A. O., Abu-Zaid, M., Alahmer, H., ... és Ghoniem, R. M. (2022). Modeling and Optimization of a Compression Ignition Engine Fueled with Biodiesel Blends for Performance Improvement. Mathematics, 10(3), 420.8.

D. Noriler, H. F. Meier, A. A. C. Barros, M. R. Wolf-Maciel, Prediction of efficiencies through simultaneous momentum, mass and energy transfer analyses in a distillation sieve tray by CFD techniques, Computer Aided Chemical Engineering. 27 (2009) 1167-1172.

Fraga, M. K. New method for estimating the characterisation properties of light petroleum fractions. Dissertation (Master's in Engineering) - Postgraduate Programme in Chemical Engineering, Federal University of Rio Grande do Sul, Porto Alegre, 2010.

Goldbach, A., Meier, H. F., Wiggers, V. R., Chiarello, . L. M., & Chivanga Barros, A. A. (2022). Combustion Performance Of Bio-Gasoline Produced By Waste Fish Oil Pyrolysis: Scientific paper. Chemical Industry & Chemical Engineering Quarterly, 28(1), 1–8. https://doi.org/10.2298/CICEQ200810010G

Handogo, Renanto. Preliminary Design of Mini Oil Refinery Plant. Journal of Advanced Research in Fluid Mechanics and Thermal Sciences92 (2021). SSN: 2811-3950.

Handogo, Renanto. Optimization on the performance of crude distillation unit (CDU). Asia-Pacific Journal of Chemical Engineering7 (2012): S78-S87.https://doi.org/10.1002/apj.644.

Jeyakumar, N., Huang, Z., Balasubramanian, D., Le, A. T., Nguyen, X. P., Pandian, P. L., és Hoang, A. T. (2022). Experimental evaluation over the effects of natural antioxidants on oxidation stability of binary biodiesel blend. International Journal of Energy Research.

Ji Shuncheng and Miguel Bagajewicz. Design of crude distillation plants with vacuum units. I. Targeting. Industrial & engineering chemistry research41, no. 24 (2002): 6094-6099.https://doi.org/10.1021/ie011040u.

Kister, Henry Z. (1992). Distillation Design, 1st Edition, McGraw-Hill. ISBN 0-07- 034909-6.

Kondor, I. P., Zöldy, M., & Mihály, D. (2021). Experimental Investigation on the Performance and Emission Characteristics of a Compression Ignition Engine Using Waste-Based Tire Pyrolysis Fuel and Diesel Fuel Blends. Energies, 14(23), 7903.

L. T. Popoola, J. A. Adeniran, S. O. Akinola, Investigations into Optimization Models of Crude Oil Distillation Column in the Context of Feed Stock and Market Value, Advances in Chemical Engineering and Science. 2 (2012) 474-480. https://doi:10.4236/aces.2012.24058.

Luyben, William L., Distillation Design and Control Using Aspen Simulation. John Wiley & Sons, Inc (2006). https://doi.org/10.1002/0471785253.

Lulić, Z., Mavrin, I., és Mahalec, I. (1998). Aspects of Using Biological Regenerative Fuels in Internal Combustion Engines. Promet-Traffic&Transportation, 10(1-2), 75-80.

Leite, André Búrigo; Bertoli, Sávio Leandro and Chivanga Barros, António André; Absorção química de dióxido de nitrogênio (NO_2); Engenharia Sanitaria e Ambiental, (2005) 10, 49-57; https://doi.org/10.1590/S1413-41522005000100006.

Mateus, Adriano da Silva; Francisco, Arleth Prata Serafim; Francisco, Paulo; Manuel, Ngoma; Barros, António André Chivanga; (2022); Thermocatalytic Cracking of Frying Oil and Plastic Waste for Production of Organic Liquid Products and Derivatives; Chemical Engineering & Technology 45 (10), 1835-1841.

Matijošius, Jonas, Akvilė Juciūtė, Alfredas Rimkus, és Jurijus Zaranka. "Investigation of the Concentration of Particles Generated by Public Transport Gas (CNG) Buses". Cognitive Sustainability 1, no. 1 (March 30, 2022). Accessed October 1, 2022. https://cogsust.com/index.php/real/article/view/10.

Miria H. M. Reis, António A. C. Barros, Antonio J. A. Meirelles, Rubens Maciel Filho, and Maria Regina Wolf-Maciel, Application of Plate and Component Efficiency Correlations in Homogeneous Azeotropic Distillation Processes, *Industrial & Engineering Chemistry Research* **2006** *45* (16), 5755-5760, DOI: 10.1021/ie051182e

Mohammad Reza Rahimpour, Mitra Jafari and Davood Iranshahi; Progress in catalytic naphtha reforming process: A review; Applied Energy Volume 109, September 2013, Pages 79-93 ; https://doi.org/10.1016/j.apenergy.2013.03.080

Marcos J. Prauchner, Ruana D. Brandão, Antônio M. de Freitas Júnior and Silvia da C. Oliveira. Petroleum-based Fuels : Obtaining, Properties and Uses. Petroleum-based Fuels: Obtaining, Properties and Uses. Rev. Virtual Quim., 2022, in press, 1-18 ©2021 Sociedade Brasileira de Química.

M. R. Wolf-Maciel, C. Soares, A. A. C. Barros, Validations of the nonequilibrium stage model and of a new efficiency correlation for non-ideal distillation process through simulated and experimental data, Computer Aided Chemical Engineering. 9 (2001) 321-326.

Pasquini, C., Ferreira Bueno, A., Characterization of petroleum using near-infrared spectroscopy: Quantitative modeling for the true boiling point curve and specific gravity. Fuel, 86, pp. 1927–1934, 2007.

P. M. S. Kalvelage, A. A. Albuquerque, A. A. C. Barros, S. L. Bertoli, (Vapor + Liquid) Equilibrium for Mixtures Ethanol + Biodiesel from Soybean Oil and Frying Oil, International Journal of Thermodynamic 20 (2017) 159 – 164. https://doi.org/10.5541/eoguijt.337160.

Quirino, F. A. B. Simulation of the Heavy Oil Deasphalting Process to Obtain Asphalts and Lubricating Oil, 2009.

R. H. Perry, D.W. Green, M.Z. Southard, Perry's Chemical Engineers' Handbook. London (2018) 2272.

T.L. Fernandes, D. de Oliveira, G. D. Donatelli, T. M. Zilio, G. G. Soares, L. A. de Andrade. Preliminary study of the effect of dimensional variation on flow measurement using a cone-type meter, 2018.

Tutak, W., Lukacs, K., Szwaja, S., és Bereczky, A. (2015). Alcohol–diesel fuel combustion in the compression ignition engine. Fuel, 154, 196-206.

Virt, M., & Arnold, U. (2022). Effects of Oxymethylene Ether in a Commercial Diesel Engine. Cognitive Sustainability, 1(3). https://doi.org/10.55343/cogsust.20

https://doi.org/10.55343/CogSust.74

Modified sectoral average greenhouse gas intensity for more accurate carbon risk analysis

Renato Ritter
Hungarian National Bank
Budapest, Hungary
ritterr@mnb.hu

Abstract

Thanks to the existing data gaps, Monitoring climate-related financial risks is no easy task. Greenhouse gases play an important role in climate change; therefore, examining economic activities' emissions can be a suitable way to analyse climate risks. However, relying only on activities' average greenhouse gas emissions can be misleading, as substantial emitters can greatly divert the average values upwards. Modifying these sectoral average emissions by subtracting substantial emitters' data results in a clearer picture of climate risk analysis. The modified sectoral average method, which primarily relies on data from the emission trading systems, treats companies under the EU Emissions Trading System (ETS) regulation separately. With this treatment, sectors with high emission values can experience a significant drop in sectoral average emission intensity compared to companies under ETS regulation. To use an already available but not applied data cluster is an affordable, easy-to-implement way to increase the accuracy of climate risk indicators. Results in Hungary show that implementing this information reduces sectoral average intensity for companies not under ETS regulation.

Keywords
transition risk, sectoral exposure, greenhouse gas intensity

1. Introduction

Failure to adequately mitigate and adapt to climate change could have serious socio-economic impacts, whether in terms of physical risks from extreme weather or negative impacts on economic activities. In recent years, the financial risks associated with climate change have become increasingly pronounced, alongside the natural impacts of climate change. Climate change and its negative economic impacts can significantly impact human society, including its economic performance. Economic effects may include market shocks, physical damage, regulatory responses, and other impacts due to climate change and changing human attitudes. To ensure financial stability, it is important to examine these possible negative effects in order to be able to prepare for their potential impacts. Taking timely preventive action can help avoid potential risks. We consider the negative impacts of measures taken to create a low-carbon and climate-resilient economy as transition risks. Risks can include rapid repricing of financial assets (both positive and negative) due to rapidly changing consumer and regulatory expectations, reduced profitability of the company's activities or obsolescence of its business model (carbon tax, production technologies becoming more expensive due to the banning of old technologies), which can lead to a drastic change in the default rate of certain activities. Late, simultaneous, large-scale regulatory measures to achieve a climate-neutral economy could cause even greater financial shocks, affecting investment and credit portfolios. In many areas, a green transition is required to reduce these shocks. For example, monetary policy can aid a smoother and faster transition to a low-carbon economy (Kolozsi et al., 2022). Although ESG ratings improve and sustainable disclosure requirements are becoming more stringent (Marczis et al., 2023), we still lack, in many cases, meaningful, easy-to-interpret environmental data, which is needed at the very least to aid a rapid transition. Emission trading systems appear to be effective tools for lowering regulated companies' emissions (David, 2022), and as later shown, they could be useful input for risk analysis as a source of much-needed environmental data.

A fully developed, common methodology for assessing climate risks does not yet exist. A reliable way to assess climate transition risks is to quantify the risks based on the greenhouse gas intensity (hereafter GHG intensity) of individual

https://doi.org/10.55343/CogSust.74

economic activities. GHG intensity reveals how much pollution is needed in each economic sector to produce one value added. Combining this sectoral average data with company-level information is usually the first step for financial institutions for transition risk calculations. Several methods have already been defined to quantify carbon risks based on GHG intensity. If individual emission data is unavailable, these methods all build on the sectoral average GHG intensity of economic activities. However, incorporating individual emission data without modifying sectoral average GHG intensity values could lead to false results. A new method was developed to incorporate individual emission data into GHG intensity-based studies, achieving more realistic results compared to calculations based only on sectoral average GHG intensity values.

2. Basis of GHG intensity analysis

GHG intensity shows the GHG emissions per unit of added value. Equation (1) shows how GHG intensity values are calculated for a sector (Eurostat, 2023d).

$$INT_i = \frac{GHG_i}{GVA_I} \quad \text{(Eq. 1)}$$

where

INT_i is the GHG intensity value for the sector I in grams GHG/euro,

GHG_i is the total GHG emissions for the sector I in grams,

GVA_i is the Gross Value Added (hereafter GVA) for sector *i* in Euro using current exchange rates.

GHG intensity data are compiled by Eurostat (Eurostat, 2023b) for countries in the European Union (hereafter EU). Data coverage is complete at a section level (NACE Rev2 level 1: A-T) and non-complete at a sector level (NACE Rev2 level 2: A01-T98). GHG intensity data are available for Hungary from 1995 onwards, with the latest data for 2020. Eurostat regularly updates the database with a lag of about two years. This data is the basis of the MNB's Bank Carbon Risk Index (BCRI), used since 2021. The index captures the potential risks arising from the GHG intensity of the activities financed by banks based on two types of functions: (i) the linear function assumes that the higher the GHG intensity of the activity, the higher the associated risk, which is mainly due to regulatory activities (e.g. carbon taxes); while (ii) Gompertz's function assumes that up to a certain level, the exposure of activity to climate regulations is negligible, but beyond a critical point the activity is certainly affected (Bokor, 2021; MNB, 2021, 2022b) (Figure 1). The Gompertz function can be considered the more stringent one of the two functions. Therefore, it estimates higher risk levels. Based on the index, at the end of 2022, between 9 and 15 per cent of the Hungarian corporate exposures can be considered risky credit exposures (calculated based on the linear Gompertz function). The EU Emissions Trading System (hereafter ETS) supports the logic of the latter function, as only activities considered large polluters have been covered so far, while activities below a certain emission level are completely unaffected by the regulation.

https://doi.org/10.55343/CogSust.74

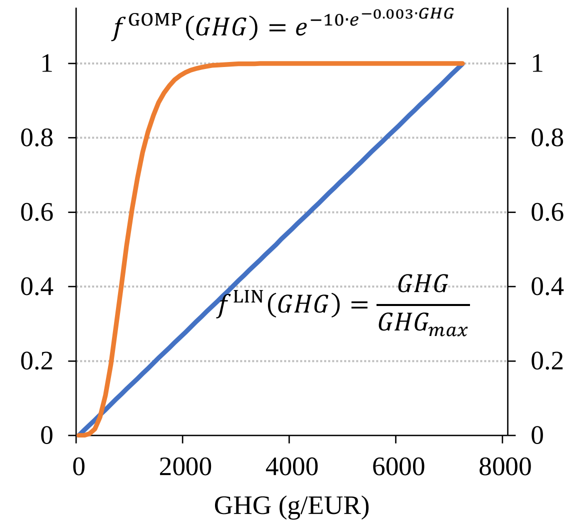


Figure 1: Transition risk weights as a function of GHG intensity (Bokor, 2021: 9)

The Task Force on Climate-Related Financial Disclosures (hereafter TCFD) also recommends the use of GHG intensity data (TCFD, 2017b), defining indicators to assess the carbon exposure of the portfolios of banks, insurers, fund managers, and investment service providers (TCFD, 2017a). Their most important indicator is the weighted average carbon intensity indicator, which was used, among others, by the MNB in its TCFD 2022 report to quantify the carbon intensity of the foreign exchange reserve, the government bond purchase programme and the Bond Funding for Growth scheme (287 tonnes CO_2e/million EUR GDP, 495 tonnes CO_2e/million EUR GDP, 551 tonnes CO_2e/million EUR GDP were the intensity of the named programmes, respectively) (MNB, 2022a). In their surveys, the European Banking Authority (hereafter EBA) and Ritter (EBA, 2021; Ritter, 2022) created 6 GHG groups (from very low to very high) based on the GHG intensity of activities by ranking the values and breaking them up at certain points.[1] Based on the 2021 results, 53 per cent of domestic corporate exposures financed exposures above the median GHG intensity, while in the EU, only 35 per cent of corporate exposures financed such activities in the EBA survey (EBA, 2021; Ritter, 2022). An important difference was that EBA used individual emissions data in creating the groups in addition to the available sectoral average intensity data without modifying it, while the domestic analysis used only sectoral average intensity data.

The advantage of using GHG intensity data is data availability at the sectoral level. The problem lies in the reliance on the debtor's main activity, the NACE code only. Companies are usually active in more than one economic sector. Therefore, using only one NACE code per debtor can mislead the analysis. The other problem arises from the usage of average GHG intensity data. As shown in the EBA analysis, including individual intensity data could significantly affect the results, thus providing a more nuanced picture of our credit portfolio's riskiness. Although the calculation and publication of GHG intensity values at the company level is not widespread nationally, this data type is available for a certain number of companies.

3. Modified sectoral average GHG intensity

Using sectoral average GHG intensity data can be a good starting point for analysis, yet altering it by treating large polluting companies separately can greatly change these values. Due to the functioning of the EU ETS, individual emissions data for large polluting companies have been available since the start of the trading scheme in 2005 (EC, 2023a). The EU database publishes site-level CO_2 emissions data for each ETS period. After aggregating the site-level data to the company level, we obtain company-level annual CO_2 emissions data. Since sectoral emissions data are available from Eurostat and the economic sector of the companies operating under the ETS regulation can be easily found, subtracting the two from each

[1] In the analysis, the 10th percentile, 1st quartile (25th percentile), median (50th percentile), 3rd quartile (75th percentile), 90th percentile were the cut-off points for the categories, with exposures above the median being considered risky in particular.

https://doi.org/10.55343/CogSust.74

other gives the annual GHG emissions of a given economic sector without the values of the companies whose emissions data are available. It is important to note that the Eurostat data include a range of GHG emissions, whereas the EU ETS database only considers CO_2 emissions. As Eurostat GHG data are published in CO_2 equivalent, it is possible to subtract the two. In Hungary, between 2005 and 2021, between 120 and 150 companies per year were covered by the EU ETS regulation, with the manufacturing and energy supply sectors (C and D) being the most affected (Table 1).

The expected future tightening of the EU ETS regulation will decrease the amount of allowances available for free allocation and purchase, hence raising the allowance's price and extending the range of activities covered by the regulation. EU policymakers would phase in CO_2 emissions from all large ships in maritime transport entering EU ports, which is responsible for a significant amount of emissions globally, accounting for 3 per cent of total emissions in 2018 (EC, 2023b). In addition, a completely separate scheme will be established to regulate pollution from land transport and buildings from 2025 onwards (Liese, 2022), which is expected to affect the H-transport and F-building sectors directly.

In order to modify the sectoral average intensity data presented in Equation (1) with the individual emission data of companies under ETS regulation, it was also necessary to calculate the GVA data in the denominator at the company level (Equation 2). The GVA values were calculated based on the methodological guidelines of the Hungarian Central Statistical Office on gross domestic product (GDP) (KSH, 2023). For this purpose, the annual reporting data available on the Ministry of Justice website (MJ, 2023) were used. Most companies use the two most common profit and loss account procedures, so the GVA values calculated on their basis are presented in more detail below.

$$GVA = output - intermediate\ consumption \qquad \text{(Eq. 2)}$$

$$GVA_1 = A \pm B - C \qquad \text{(Eq. 3)}$$

where GVA_1 is the GVA of the companies using the total cost method,
A is the net sales revenue,
B is the capitalised value of one's performance,
C is the value of material expenditure.

$$GVA_2 = A - B - C \qquad \text{(Eq. 4)}$$

where GVA_2 is the GVA of the companies using the turnover cost method,
A is the net sales revenue,
B is the direct costs linked to sales,
C is the value of the material expenditure.

Not all companies examined had income statement data available in all cases from which GVA could be derived. Companies with data gaps were excluded from the analysis in years without income statement information, so their emissions data were not considered in subsequent calculations in those periods. Depending on the year, between 95 and 99 per cent of the annual reports (based on the number of companies) were available for companies under ETS regulation.

In the remainder of the analysis, companies under ETS regulation with individual emissions data are referred to as ETS companies, and the remaining companies not under ETS regulation are non-ETS companies. For perspective, presenting the CO_2 emissions and value-added data of ETS companies concerning the national data might be worthwhile. It can be said that ETS companies contribute largely to CO_2 emissions in Hungary, while they are responsible to a much lesser extent for the generated gross value added in the country. The CO_2 pollution of these companies was responsible for 37–44% of the domestic pollution between 2005 and 2021, while their gross value added accounted for only 4.8–6.6% of the national value added (Figure 2) on the same time horizon. Based on these, we can assume that the GHG intensity of non-ETS companies in sectors affected by ETS regulation may be much lower than their ETS-affected counterparts.

Table 1: Breakdown of all Hungarian companies under ETS regulation by sector of their main activity (top), a ratio of Hungarian companies under ETS regulation with annual reports to all companies participating in the ETS in each sector (bottom)
Source: EC 2023a, MJ 2023, own analysis

https://doi.org/10.55343/CogSust.74

Year	Sector (number of companies)				Total
	C	D	H	Other	
2005	87	50	3	12	152
2006	87	51	3	13	154
2007	85	52	3	9	149
2008	80	52	3	9	144
2009	76	51	4	10	141
2010	76	52	4	10	142
2011	75	52	4	10	141
2012	72	51	8	10	141
2013	67	49	4	10	130
2014	64	46	6	8	124
2015	61	42	6	8	117
2016	61	41	6	8	116
2017	65	42	6	7	120
2018	64	44	7	7	122
2019	66	46	6	5	123
2020	67	43	5	5	120
2021	71	41	5	4	121

Year	Sector (proportion of annual reports available by number of companies)				Total
	C	D	H	Other	
2005	93%	100%	100%	92%	95%
2006	94%	98%	100%	92%	95%
2007	94%	98%	100%	100%	96%
2008	95%	98%	100%	100%	97%
2009	96%	100%	100%	100%	98%
2010	92%	100%	100%	100%	96%
2011	93%	100%	100%	100%	96%
2012	96%	100%	88%	100%	97%
2013	97%	100%	100%	100%	98%
2014	95%	100%	100%	100%	98%
2015	95%	100%	100%	100%	97%
2016	97%	100%	100%	100%	98%
2017	98%	100%	100%	100%	99%
2018	98%	100%	100%	100%	99%
2019	98%	100%	100%	100%	99%
2020	99%	100%	100%	100%	99%
2021	97%	95%	100%	100%	97%

* Note: C – Manufacturing; D – Energy supply; H – Transportation; Other – other sectors aggregated

https://doi.org/10.55343/CogSust.74

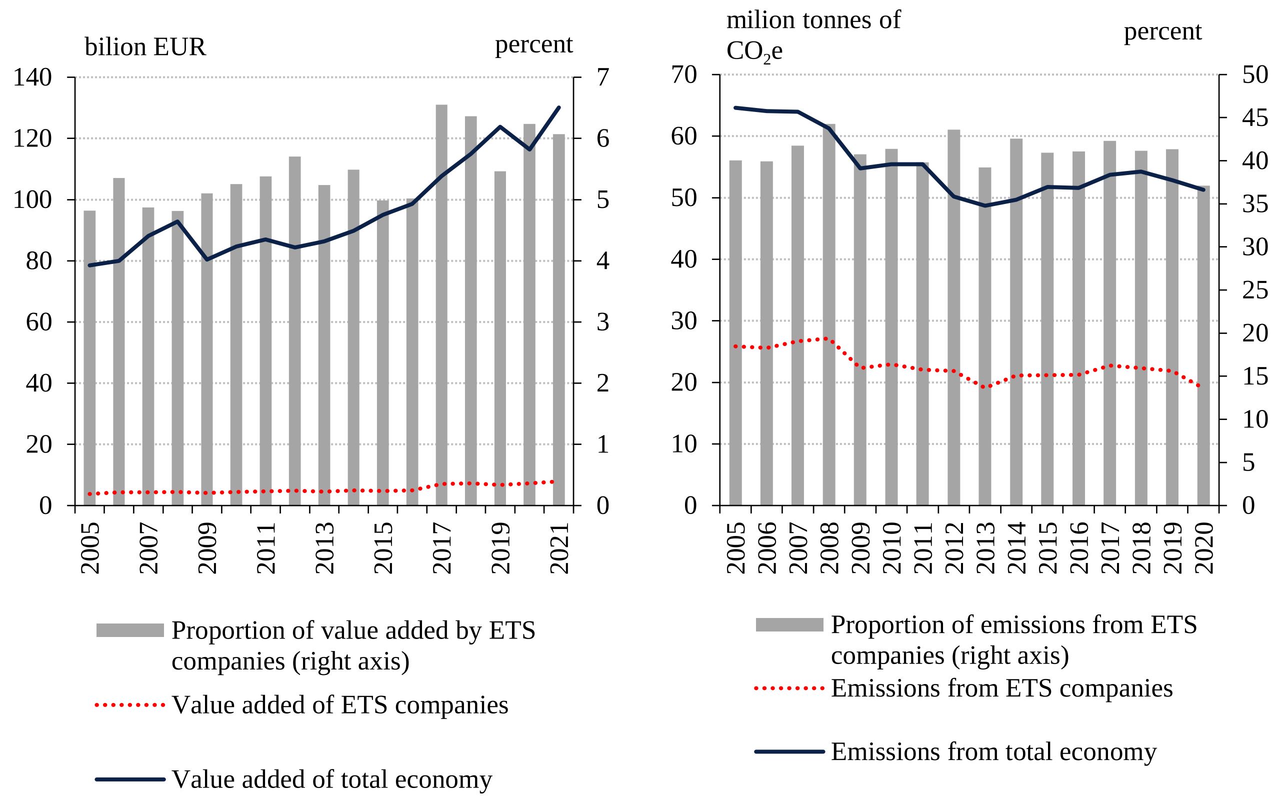


Figure 2: The share of companies under ETS regulation in value-added generated by the Hungarian economy (left) and national GHG emissions (right)

Source: Eurostat 2023a, Eurostat 2023c, EC 2023a, MJ 2023, own analysis

The modified sectoral average GHG intensity differs from the one presented in Equation (1) because, from both the denominator and the numerator, the emissions and value-added data of ETS companies active in the sector in the given year were removed if both data points were available. The resulting modified sectoral average GHG intensity data can be taken as non-ETS companies' new average intensity data.

$$INT_{NONETS_i} = \frac{GHG_i - GHG_{ETSi}}{GVA_i - GVA_{ETSi}} \quad \text{(Eq. 5)}$$

where $INT_{NONETSi}$ is the GHG intensity of non-ETS companies in sector i,
GHG_i is the total GHG emissions in sector i,
GHG_{ETSi} is the GHG emissions of ETS companies in sector i,
GVA_i is the total GVA of companies in sector i,
GVA_{ETSi} is the GVA of ETS companies in sector i.

Both GHG_i and GVA_i correspond to the original sectoral values published by Eurostat (Eurostat, 2023a, 2023c), while GHG_{ETSi} values are from the EU ETS GHG emissions inventory of ETS companies (EC, 2023a) and GVA_{ETSi} values are calculated using the two GVA calculation methods described above using annual reporting data (MJ, 2023).

3. Results and discussion

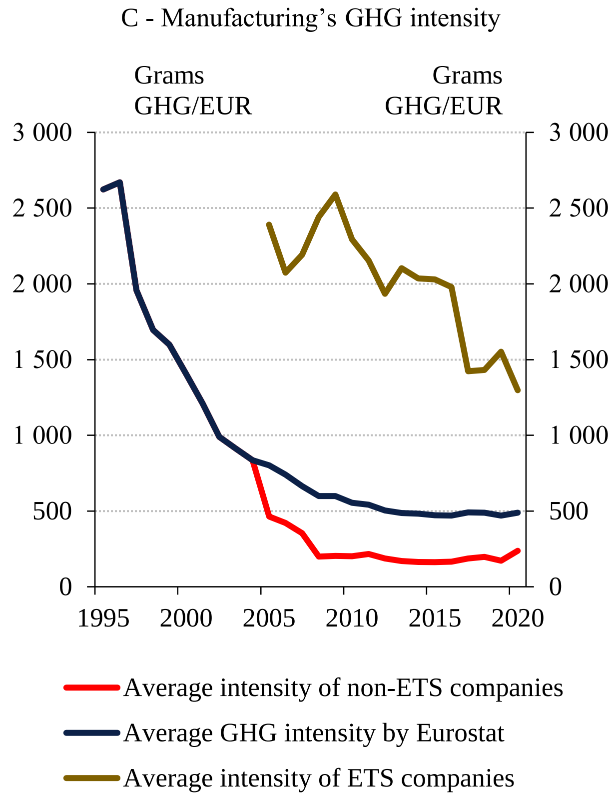
The modified sectoral average GHG intensity values calculated for non-ETS companies are no longer influenced by ETS companies' values, which are the largest CO_2 emitters. Using the modified values gives a more accurate picture of the GHG intensity of non-ETS companies in each ETS-affected sector, and hence, their GHG intensity-based climate change risk values will likely decrease. We can refine our existing risk measurement tools based on sectoral GHG intensities using modified sectoral average GHG intensity data for non-ETS companies. Unsurprisingly, the two sectors most heavily involved in the ETS are the ones where the modification results are most noticeable (Figure 3).

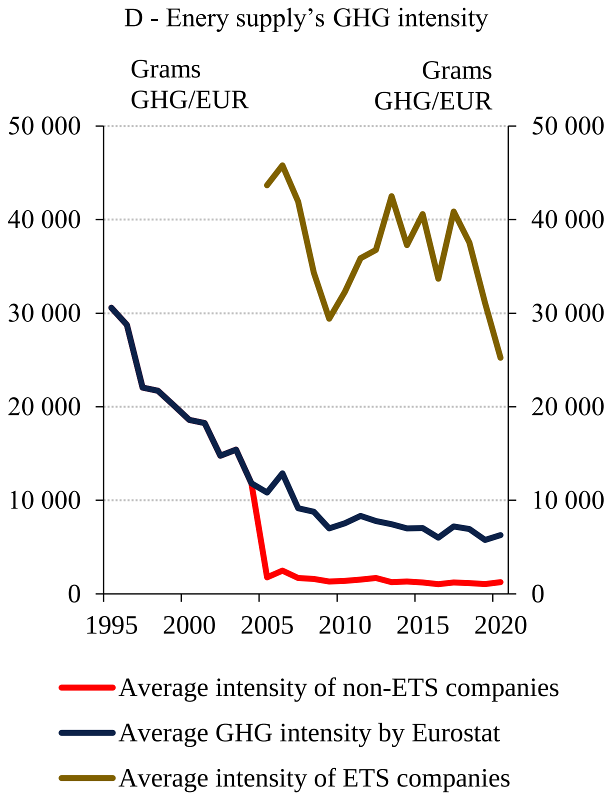
https://doi.org/10.55343/CogSust.74

Sectors C–Manufacturing and D–Energy supply show a drastic GHG intensity decrease compared to the average intensity values produced by Eurostat (blue lines), with the former showing a 42–67% decrease and the latter a 78–84% decrease in average GHG intensity values for non-ETS companies (red lines) depending on the year. We see drastic differences if we compare the sectoral average GHG intensity of non-ETS and ETS companies in the two sectors. On average, ETS companies (brown lines) can have up to 7–33 times higher average GHG intensity than their non-ETS counterparts in the D–Energy sector, while the same value was 3–12 times higher for ETS-companies in the C–Manufacturing sector.

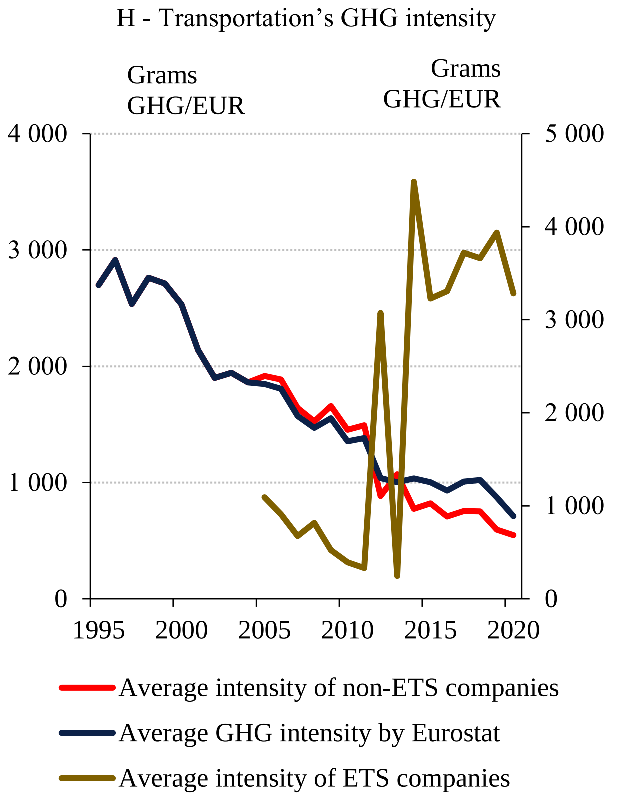
Surprisingly, results for the initial period can be observed in the H-Transportation sector. In the first years, several ETS companies produced much higher added value than the rest of the sector. As a result, although fewer GHG emissions remained in the numerator, the large reduction in value-added emissions resulted in higher GHG emissions per value-added produced than before the modification. So, the modified average GHG intensity values for non-ETS companies were initially higher than those calculated by Eurostat. This trend was reversed in 2014, and since then, as in the other two sectors above, the modified average intensity values of non-ETS companies have been lower than the original ones.

Looking at the average GHG intensity of the Hungarian economy, there has also been a significant reduction. Although there has been a steady decline in the economy's GHG intensity since 1995 with the decline of heavy industry and the growth of the service sector, modifying it with ETS companies' data has resulted in a further 30–41% reduction compared to the average intensity values calculated by Eurostat, depending on the year. When comparing the non-ETS and ETS companies, the differences are as striking as for the first two mentioned sectors (manufacturing and energy supply). The average GHG intensity of ETS companies in the whole economy was 7–15 times higher than the average GHG intensity of non-ETS companies.





https://doi.org/10.55343/CogSust.74



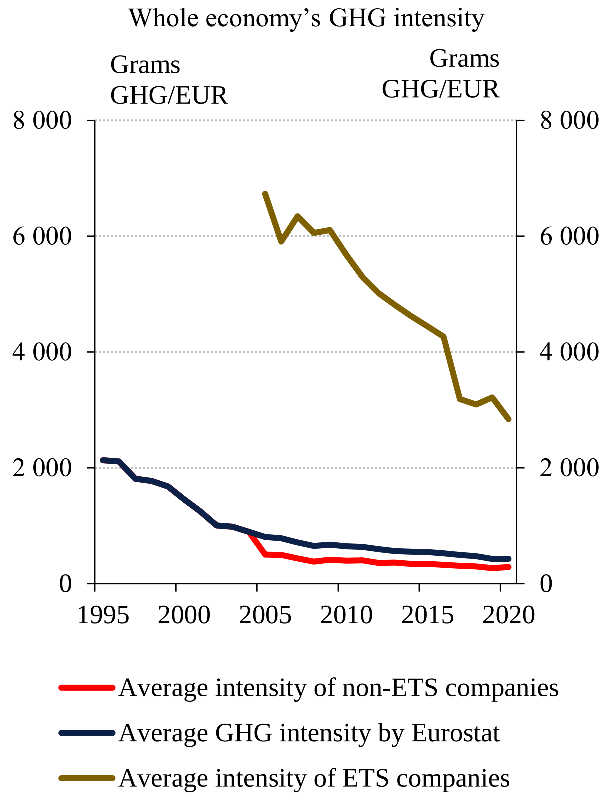


Figure 3: Change in sectoral average GHG intensity of most affected sectors and economy as a whole
Source: Eurostat 2023a-c, EC 2023a, MJ 2023, own analysis

https://doi.org/10.55343/CogSust.74

4. Conclusion

Introducing new individual emission data to already in-use sectoral average GHG intensity data for climate risk analysis could significantly improve our existing analytical tools. However, applying only the individual emission data to the company, where possible, and not modifying the sectoral average data can cause misleading results. To overcome this, a modification method was introduced to the sectoral average GHG intensity values. With the method applied, where most of the companies with individual emission data were active, a drastic change could be observed in the sectoral average GHG intensity value. An already available data cluster, namely the emissions data from the EU ETS, was used for the modification. Many companies will likely disclose their CO_2 emissions values shortly. Using these values the way introduced above will benefit the disclosing companies and their sectoral average, thus avoiding double counting their emissions.

However, one problem arises when relying on GHG intensity data: the effect of the value added by high GHG emitter companies. The high enough value-added generation of companies and sectors can mask high pollution. Thus, GHG intensity follows the weak sustainability approach, which postulates the full substitutability of natural capital, as GHG intensity does not recognise some companies as high polluters as long as they do it with high enough value-added generation. On the other hand, strong sustainability assumes that human and natural capital are complementary but not interchangeable. Therefore, only GHG emission values should be used instead of GHG intensity. This approach could result in higher climate risks, as high polluters with high value-added would be considered risky exposures compared to the current situation. Of course, some now high GHG intensity activities could become low emission activities as well, as high GHG intensity can occur when low pollution is coupled with even lower value added. Further research and comparison between the two approaches is necessary to reveal the true impact of this methodological transition.

References

Bokor, L. (2021). *Bank Carbon Risk Index – A simple indicator of climate-related transition risks of lending activity*. Central Bank of Hungary, MNB Occasional Paper (141). URL: https://www. mnb.hu/letoltes/mnb-op-141-final.pdf

David, A. (2022). *Carbon emission trading as a climate change mitigation tool*. Cognitive Sustainability, 1(3). DOI: https://doi.org/10.55343/cogsust.33

EBA. (2021). *Mapping climate risk: Main findings from the EU-wide pilot exercise*. EBA/ Rep/2021/11. URL: https://www.eba.europa.eu/sites/default/documents/files/document_library/Publications/Reports/2021/1001589/Mapping%20Climate%20Risk%20-%20Main%20findings%20from%20the%20EU-wide%20pilot%20exercise%20on%20climate%20risk.pdf

EC. (2023a). *European Union Transaction Log – Emission data of ETS companies (public database)*. URL: https://ec.europa.eu/clima/ets/allocationComplianceMgt.do?languageCode=en

EC. (2023b). *Reducing emissions from the shipping sector*. URL: https://climate.ec.europa.eu/eu-action/transport-emissions/reducing-emissions-shipping-sector_en

Eurostat. (2023a). *Eurostat/Air emissions accounts by NACE Rev. 2 activity [ENV_AC_AINAH_R2] (public database)*. URL: https://ec.europa.eu/eurostat/databrowser/view/env_ac_ainah_r2/default/table?lang=en

Eurostat. (2023b). *Eurostat/Air emissions intensities by NACE Rev. 2 activity [ENV_AC_AEINT_R2] (public database)*. URL: https://ec.europa.eu/eurostat/databrowser/view/env_ac_aeint_r2/default/table?lang=en

Eurostat. (2023c). *Eurostat/National accounts aggregates by industry (up to NACE A*64) [NAMA_10_A64] (public database)*. URL: https://ec.europa.eu/eurostat/databrowser/view/nama_10_a64/default/table?lang=en

Eurostat. (2023d). *Methodological guide: Air emissions accounts by NACE Rev. 2 activity (env_ac_ainah_r2)*. URL: https://ec.europa.eu/eurostat/cache/metadata/en/env_ac_ainah_r2_sims.htm

Kolozsi, P. P., Horváth, B. I., Csutiné-Baranyai, J., Tengely, V. (2022). Monetary Policy and Green Transition. *Financial and Economic Review*. 21(4), 7–28. DOI: https://doi.org/10.33893/FER.21.4.7

KSH (2023): *Methodology of Gross domestic product*. URL: https://www.ksh.hu/docs/eng/modszgyors/egdpmodsz15.html

Liese, P. (2022). *Revision of the EU emission trading system (ETS)*. URL: https://www.europarl.europa.eu/legislative-train/package-fit-for-55/file-revision-of-the-eu-emission-trading-system-(ets)

Marczis, D., Mihálovits, Zs., Sebestyén, G. (2023). Sustainability and Climate Risk Data: A New Era for Investment Decision–Making in the Age of Climate Change. *Cognitive Sustainability*. 2(2). https://doi.org/10.55343/cogsust.64

MJ. (2023). *Ministry of Justice: Business Information and Electronic Company Registration Service*. Individual annual reporting data for ETS companies was collected from the database. URL: https://e-beszamolo.im.gov.hu/oldal/kezdolap

MNB. (2021). *Green Finance Report: Status report on the sustainability of the Hungarian financial system*. Central Bank of Hungary. URL: https://www.mnb.hu/letoltes/20210303-zold-penzugyi-jelentes-angol.pdf

MNB. (2022a). *The Magyar Nemzeti Bank's Climate-related Financial Disclosure*. *Central Bank of Hungary*. Central Bank of Hungary. URL: https://www.mnb.hu/letoltes/tcfd-jelente-s-2022-en.pdf

MNB. (2022b). *Macroprudential Report*. Central Bank of Hungary. URL: https://www.mnb.hu/letoltes/makroprudencialis-jelentes-2022-eng.pdf

Ritter, R. (2022). Banking Sector Exposures to Climate Risks – Overview of Transition Risks in the Hungarian Corporate Loan Portfolio. *Financial and Economic Review*. 21(1), 32–55. DOI: https://doi.org/10.33893/FER.21.1.32

TCFD. (2017a). Implementing the Recommendations of the Task Force on Climate-related Financial Disclosures. *Task Force on Climate-related Financial Disclosures*. URL: https://assets.bbhub.io/company/sites/60/2020/10/FINAL-TCFD-Annex-Amended-121517.pdf

https://doi.org/10.55343/CogSust.74

TCFD. (2017b). Recommendations of the Task Force on Climate-related Financial Disclosures. *Task Force on Climate-related Financial Disclosures*. https://assets.bbhub.io/company/sites/60/2020/10/FINAL-2017-TCFD-Report-11052018.pdf

https://doi.org/10.55343/CogSust.84

Integrating Sustainability: Unveiling the Quadruple Helix – A Study on Pre-development of Smart City Strategy

Natalia Stalmasekova,
0000-0003-2314-1984,
University of Zilina
Univerzitná 8215/1, 010 26 Zilina, Slovakia
natalia.stalmasekova@uniza.sk

Juraj Fabus
0000-0002-5969-1155
University of Zilina
Univerzitná 8215/1, 010 26 Zilina, Slovakia

Abstract

Urban centers, irrespective of their size, grapple with critical issues concerning urban transport, parking, and environmental preservation. Finding solutions to these challenges requires proactive measures, such as research initiatives yielding practical applications. Sustainable innovation flourishes through collaborative efforts across diverse sectors, encouraging mutual learning and optimal solution generation. Embracing the principles of sustainability, one promising approach lies in the Smart City concept adoption. This article delves into the significance of the Smart City paradigm, emphasizing the integration of the Quadruple Helix model, with a specific focus on the involvement of the fourth helix – society. The article highlights the societal engagement in shaping Smart City strategies on a pre-development phase questionnaire survey, emphasizing the pivotal role of societal involvement in formulating sustainable urban strategies.

Keywords

Smart city, quadruple helix model, sustainability

1. Introduction

The "Smart City" is becoming a common term in the media and among the general public. In practice, the individual cities, which call themselves intelligent, differ significantly. This is mainly due to the fact that the concept itself is still evolving. As each city is unique and the implementation of the Smart City concept depends on the city's policies, goals and financial capabilities, the outcomes show a versatile picture. The strategy dedicated to the Smart City concept is gaining popularity mainly due to the growing population of cities and the related effort to ensure sustainable urban development. There are large economic differences between the countries of the European Union, including unemployment and inflation rates and the amount of public debt. The growing population, relatively low birth rate and aging population are considerable problems not only in Europe, but all over the world.

In order to highlight the main trends regarding population changes, some urbanization facts are presented here from Ritchie and Roser's article (2019):

- More than 4 billion people live in urban areas globally.
- The UN estimates 2007 was the year when, for the first time, more people in the world lived in urban than in rural areas.
- Estimates on urban populations vary – mainly as a result of disagreements on the exact definition of an "urban area' and what this includes.
- For most of human history, populations lived in very low-density rural settings. Urbanization is a trend unique to the past few centuries.
- By 2050 it's projected that more than two-thirds of the world population will live in urban areas.
- It is projected that close to 7 billion people will live in urban areas in 2050.
- People tend to migrate from rural to urban areas as they become richer.
- Living standards tend to be higher in urban areas.

https://doi.org/10.55343/CogSust.84

The world's cities are growing in both size and number. At the turn of the century in 2000, there were 371 cities with 1 million inhabitants or more worldwide. By 2018, the number of cities with at least 1 million inhabitants grew to 548, and by 2030, 706 cities are expected to have at least 1 million residents. Cities with more than 10 million inhabitants are often termed "megacities". Globally, the number of megacities is expected to rise from 33 in 2018 to 43 in 2030. In 2018, 48 cities had populations between 5 and 10 million. By 2030, 10 of these are believed to become megacities. Projections also indicate that 28 additional cities will cross the 5 million mark between 2018 and 2030, of which 13 are located in Asia and 10 in Africa. In 2030, 66 cities are expected to have between 5 and 10 million inhabitants (UN, 2018).

In the light of these data, it is clear that the city administration will have to face significant problems in the future related to waste management, natural resources, environmental pollution, traffic jams, deterioration of transport infrastructure and human health. The Smart City concept has the potential to address many of the challenges.

This article deals with the adoption of the concept of Smart city in the Slovak city, in its pre-development phase and it highlights the society engagement, because the successful implementation of the Smart city concept requires collaboration among various stakeholders, including government, private enterprises, and the local community (Peters, 2017).

1.1. Smart city definition

The concept of Smart City, as we see it, connects citizens, the city administration and the private sector through technology into one unit that works together to improve urban life experience, create solutions in challenging areas as ecology, transport, sustainability and wellbeing in general.

An overview of the definitions of the term Smart City was provided by the authors Gil-Garcia et al. (2015). They stated that definitions of a Smart City and many related phrases and terms refer to the same or very similar phenomena. However, each highlight different aspects or components, seemingly according to the focus of the proposing author's research, disciplinary backgrounds, and/or what different authors see as more important or prevalent. We have chose the ones that we think best capture the meaning of Smart City in line with our view.

- Kourtit et al. (2012: 234): "Advanced business and socio-cultural attractiveness [...], presence of a broad (public and private) labor force and public facilities [...], and presence and use of sophisticated e-services."
- Komninos (2009: 338): "Territories with high capacity for learning and innovation, which is built-in the creativity of their population, their institutions of knowledge creation, and their digital infrastructure for communication and knowledge management."
- Kourtit and Nijkamp (2012:93): "A promising mix of human capital (e.g. skilled labor force), infrastructural capital (e.g. high-tech communication facilities), social capital (e.g. intense and open network linkages) and entrepreneurial capital (e.g. creative and risk-taking business activities)."
- Woods and Citron (2014:1) "The integration of technology into a strategic approach to sustainability, citizen well-being, and economic development."
- Rios (2008:4): "A city that gives inspiration, shares culture, knowledge, and life, a city that motivates its inhabitants to create and flourish in their own lives."
- Giffinger et al. (2007:11): "A city well performing in a forward-looking way in [economy, people, governance, mobility, environment, and living] built on the smart combination of endowments and activities of self-decisive, independent and aware citizens."

1.2. Benefits of the Smart cities

The introduction of the concept of a Smart City brings many benefits. The positive reasons for the spread of this concept include, in particular, the rapid progress in technological development and the reduction of costs for data collection, sharing and analysis. With the right planning and investment, cities can improve their functionality, long-term sustainability and raise the living standards of their citizens (Smartcity.gov):

- Increased level of functionality of the city: means a functional economy of the city, the possibility of employment, access to basic aspects of prosperity – to infrastructure services such as interconnection and connectivity; reliable, sustainable and low-cost energy sources; adequate training opportunities; affordable forms of housing and efficient transport.
- Sustainability: means giving citizens access to the resources they need with regard to security for future generations. Sustainability is a method by which resources are not depleted or permanently destroyed. Sustainability is not only about the environment, but also about the economy. Smart cities make efficient use of natural resources, economic resources and human capital in order to create their urban infrastructure that delivers the highest possible outputs and costs as few inputs as possible.

https://doi.org/10.55343/CogSust.84

- Raising citizens' living standards: In a Smart City, citizens have access to a comfortable, healthy, clean, safe and active lifestyle, which includes several aspects such as cheap energy, convenient public transport, quality education, faster public services, clean water and air, low crime rates and access to various entertainment and cultural opportunities.

These activities help to increase the standard of living of the population, to improve the overall management of the city and to regulate the negative impacts on the environment (Jaculjaková et al., 2019).

2. Data and methods

Huong Thu Nguyen from Universitat Autònoma de Barcelona is currently conducting research on the project "From triple to quadruple helix, The role of society/citizens/customers in innovation." The triple helix model was the core concept of innovation developed since the mid 1990s where the collaborations among three helices of university, industry and government were encouraged. It aimed to enhance knowledge transfer, product and service development; and therefore, it aimed to lead to positive impacts on innovation and regional development. Having such top-down approach, the effectiveness of the triple helix model has recently been questioned . It is argued by some authors that the model does not ensure a long-term sustainable growth due to the lack of society's involvement (Cai and Lattu, 2022), (Pique et al., 2018).

Based on the concern above, the purpose of developing the Quadruple helix model is adding to the Triple helix model a fourth helix that represents societal needs. Thus the new model explicitly includes society in the process of knowledge creation. The advantage of this model is the combination of both top-down and bottom-up approaches. Bottom-up initiatives strengthened by top-down programs are believed to lead to the most successful results (Nguyen, 2018).

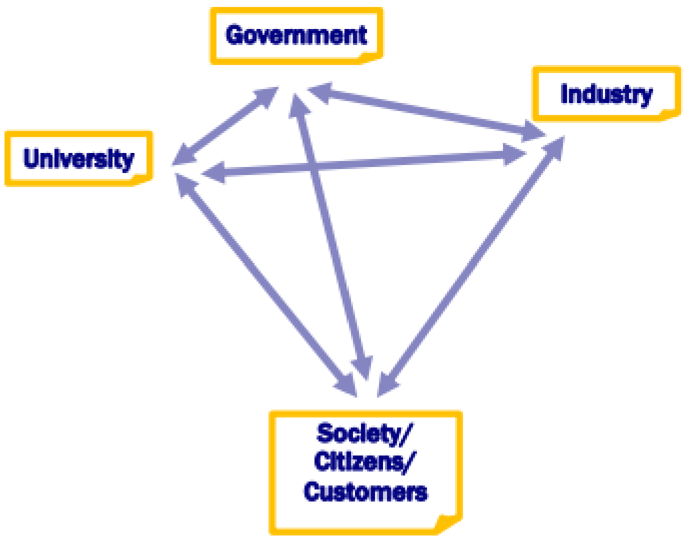


Figure 1 Quadruple helix model

Today, involving the public in research, development, and innovation is the dominant paradigm both in international STI-policy and in innovation research. The concept forms the backbone of several national innovation policies, strengthening regional innovation systems and enabling better evaluation of research organizations and research proposals (Cavallini et al., 2016). Collaborating with societal actors not only meets an established standard – cooperation and collaboration is the duty of every actor in an innovation system (Schütz et al., 2019).

3. Case study

For the purposes of this paper, we chose the city of Žilina, Slovakia. The city of Žilina is the capital of the Žilina Region, which covers an area of 6809 km and has a population of 691 613, including 82 656 inhabitants living in the city itself, while the district of Žilina has 161 377 inhabitants (SODB, 2021).

The reasons why Žilina is an eligible city for Smart city concept are as follows.

- population over 100 000 citizens (in the whole district),
- university,
- headquarters of several multinational companies.

https://doi.org/10.55343/CogSust.84

The Smart City concept in Žilina should correspond with Strategic Implementation Plan created by European Innovation Partnership on Smart Cities and Communities (Strategic Implementation Plan, 2013):

- Sustainable Urban Mobility – Alternative energies, public transport, efficient logistics, planning.
- Sustainable Districts and Built Environment – improving the energy efficiency of buildings and districts, increasing the share of renewable energy sources used and the livability of our communities.
- Integrated Infrastructures and processes across Energy, ICT and Transport – connecting infrastructure assets to improve the efficiency and sustainability of cities.

3.1. Creation of the Smart City strategy

Mora and Bolici (2016) published a case study about developing a Smart City strategy in Barcelona, which started in 2010. The good practice of creating the Smart City strategy in the cities around the world should became the source of knowledge and inspirations for cities of any size. The development process for creation of Smart City strategy consist of these steps:

1. Starting – Based on the motivation to make a decision to create a strategy and find the right people to create it.
2. Planning – analysis, vision formulation, action plan, team building.
3. Development of the projects – start of the implementation process, support of financing.
4. Monitoring and evaluation – supervision and possible adjustments.
5. Communication – getting the Smart city concept into the consciousness of the population, arousing enthusiasm.

To implement such a fundamental and complex strategy, the positive engagement of the citizens is needed. According to case study by Mora et al. (2018), the collaboration of the public and the private sector is the core engine behind the Smart City development strategies in chosen European cities under investigation and the program of activities that they have implemented. These programs mainly result from a Triple helix model of collaboration based on the interaction between Research, Industry and Government. Civil society organizations are the less represented organization type, along with citizens. But the data in the study also show that the four chosen cities have made an effort to connect urban innovation and civil society by stimulating citizens' active involvement in the development of Smart City projects and initiatives, thus acting according to Quadruple helix model. Involvement of the citizens should be important in the developing the Smart City strategy in Žilina and in subsequent expansion of the concept. As Žilina with its 82 656 citizens does not belong to the large European cities, the sense of community is important to build on.

For the involvement of citizens in the Smart City strategy in the pre-developing phase, a questionnaire survey could be the first step. Citizens can feel positive that they are thus directly involved in improving their city, and the planned change will also be based on their views.

In further phases of development, locals can be involved in the project in the following ways.

1. Involvement in workshops, discussions, forums with other stakeholders.
2. Making citizens beta testers for purposed smart solutions.
3. Encourage citizens to participate in the developing of the new digital ideas – hackathon and other creative events.

3.2. Proposal for a questionnaire survey for Smart City strategy in Žilina

Research characteristics:

- Object of the investigation: Residents of Žilina city.
- Type of the research: Primary research, quantitative research.
- Data collection method: Electronic inquiry, personal inquiry.
- Collection technique: Inquiry via electronic questionnaire Google Docs, inquiry via paper questionnaire.
- Time period: It will be specified based on current situation, but the collection must take place for at least a month.
- Method of the research testing: Pre-test on a small sample of respondents.

https://doi.org/10.55343/CogSust.84

Through a survey, the city of Žilina could bring the concept of Smart city to the attention of residents and start promoting its benefits. The city could create a campaign through which this survey would also be carried out. Campaign name suggestions: “Be a smart citizen,” “Creating a Smart City together,” “Smart change together.”

Questionnaires should be collected electronically, through the campaign website, which would be advertised on the largest electronic media in Žilina: Žilinak.sk, myzilina.sme.sk. Questionnaires in paper form could be available on the premises of the client center, where it could be filled in by the center’s clients while waiting for the services. The questionnaires would be anonymous. The design of the Smart City concept strategy for the Žilina should also take into account the needs of the citizens who live in the city on the basis of the results of the questionnaire survey. The questionnaire should focus on satisfaction in the areas that are the most important for the citizen. Topics of the questionnaire should be based on citizens’ suggestions from the “Odkaz pre starostu” (odkazprestarostu.sk) platform, including:

- city management,
- cleanliness,
- urban transport
- communication with citizens.

The questionnaire should not be time consuming for respondents to fill in. The main goal of the survey is to identify the most important areas that citizens consider fundamental to include in the Smart City strategy of Žilina. according to our proposal, the survey form should contain 16 closed questions with a choice of answers. A draft questionnaire is given in Appendix. We based the creation of the questionnaire on the strategy of the Czech city Písek and its “Blue and Yellow Book (Modrožlutá kniha Písek, 2015),” which contains the strategy of the Smart City concept in this city.

Every city with the motivation for implementing the Smart City concept should follow several steps. The main step should be to set up a working group composed of experts to deal with this agenda. As a matter of priority, the city should anchor the Smart City concept in a strategic document, which would define long-term goals and projects that the city should address. An appropriate way to set the priorities is to take into account the needs of the citizens who live in the city. The city should also set up a special fund to support the Smart City concept, to which it will contribute a certain amount from its budget and use all available resources from the government and European Union The city should also be open to cooperation with the academia and commerce, and develop cooperation across cities. Citizens should be informed about the projects and the concepts, and they should ideally be involved in the development.

The introduction of the Smart City concept, with the right timing and implementation, can be a significant step forward for all city stakeholders. It is a concept without which cities are unlikely to function in the future.

4. Conclusion

The Quadruple Helix Model integrates industry with three other major actors: science, policy, and society, fostering innovation. Utilizing information and communication technologies, diverse tools, and knowledge management strategies in organizations can lead to innovations applicable to the Smart City concept. Following the Quadruple Helix model, governments increasingly prioritize greater public involvement in innovation processes, aligning with sustainability goals.

As urban centers grapple with multifaceted challenges related to transportation, environmental preservation, and overall quality of life, the Smart City paradigm emerges as a promising avenue for holistic solutions. The article advocates for a proactive approach that involves the active participation of the community in the formulation of Smart City strategies. By incorporating the insights and preferences of the residents through pre-development phase questionnaire surveys, the article underscores the importance of aligning technological advancements with the values and needs of the society it serves.

In essence, the success of the Smart City concept lies not only in the sophistication of its technological infrastructure, but also in the depth of its integration with the values and aspirations of the society it serves.

Acknowledgement

The paper is published with the support of project VEGA 1/0460/22 and project CIS 07711134 by Ministry of Educations SR.

https://doi.org/10.55343/CogSust.84

References

Cai, Y., Lattu, A. (2022), Triple Helix or Quadruple Helix: Which Model of Innovation to Choose for Empirical Studies?. Minerva 60, 257–280. https://doi.org/10.1007/s11024-021-09453-6

Cavallini, S., Soldi, R., Friedl, J., Volpe, M. (2016). Using the Quadruple Helix Approach to Accelerate the Transfer of Research and Innovation Results to Regional Growth. *Brussels: European Union*. DOI: https://doi.org/10.2863/408040

European Innovation Partnership on Smart Cities and Communities. (2013). *Strategic Implementation Plan*. URL: https://smartcities.at/wp-content/uploads/sites/3/sip-final-en.pdf

Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanovic, N., Meijers, E. (2007). *Smart Cities: Ranking of European Medium-Sized Cities*. Centre of Regional Science (SRF), Vienna University of Technology, Vienna. URL: http://www.smart-cities.eu/download/smart_cities_final_report.pdf.

Gil-Garcia, J. R., Pardo, T. A., Nam, T. (2015). What makes a city smart? Identifying core components and proposing an integrative and comprehensive conceptualization. *Information Polity*., 20(1), 61–87. DOI: https://doi.org/gqrq5b

Jaculjaková, S., Štofková, K. R., Laitkep, D. (2019). SMART City Concept within selected countries. *CBU International Conference Proceedings*. Praha: CBU Research Institute, 151–158. DOI: https://doi.org/k3wg

Komninos, N. (2009). Intelligent cities: Towards interactive and global innovation environments. *International Journal of Innovation and Regional Development*. 1(4), 337–355. DOI: https://doi.org/chbr39

Kourtit, K., Nijkamp, P. (2012). Smart cities in the innovation age. Innovation: The European Journal of Social Science Research, 25(2), 93–95. DOI: https://doi.org/k3wh

Kourtit, K., Nijkamp, P., Arribas, D. (2012). Smart cities in perspective – A comparative European study by means of self-organizing maps. *Innovation: The European Journal of Social Science Research*. 25(2), 229–246. DOI: https://doi.org/gj253b

Modrožlutá kniha Písek (2015). Smartplan obce Písek. URL: &https://www.mesto-pisek.cz/assets/File.ashx?id_org=12075&id_dokumenty=5399

Mora, L., Bolici, R. (2016). The development process of Smart city strategies: the case of Barcelona. In Rajaniemi, J. (ed.): *Re-City: Future City – Combining Disciplines*. School of Architecture Tampere University of Technology, Tampere. 155–181. URL: https://trepo.tuni.fi/bitstream/handle/10024/116760/DATUTOP_34.pdf?sequence=1&isAllowed=y

Mora, L., Deakin, M., Reid, A. (2018). Strategic Principles for Smart City Development: A Multiple Case Study Analysis of European Best Practices. *Technological Forecasting and Social Change*, 142(May), 70–97. DOI: https://doi.org/gf3w9p

Nguyen, T. H. (2022). From the Triple to the Quadruple Helix of Innovation: Theory and Practice. PhD Thesis. Universitat Autònoma de Barcelona, Barcelona. URL: https://tesisenred.net/bitstream/handle/10803/687801/thn1de1.pdf?sequence=1&isAllowed=y

Peters, B. (2017). Smart City Strategy Success Factors. URL: https://www.ibigroup.com/wp-content/uploads/2017/10/Smart-Cities-Bruno-White-Paper.pdf

Pique, J.M., Berbegal-Mirabent, J. & Etzkowitz, H. (2018). Triple Helix and the evolution of ecosystems of innovation: the case of Silicon Valley. Triple Helix 5, 11. DOI: https://doi.org/10.1186/s40604-018-0060-x

Rios, P. (2008). Creating "the Smart city". URL: https://archive.udmercy.edu/handle/10429/393

odkazprestarostu.sk. (n.d.). Odkaz pre starostu Žilina. URL: https://www.odkazprestarostu.sk/zilina

Ritchie, H., Roser, M. (2019). Urbanization. *Our World in Data.* URL: https://ourworldindata.org/urbanization

Smartcity.gov.sk. (n.d.). Benefity Smart Cities [Benefits of Smart Cities]. URL: https://www.smartcity.gov.sk/benefity-smart-cities/index.html

Schütz, F., Heidingsfelder, M. L., Schraudner, M. (2019). Co-shaping the Future in Quadruple Helix Innovation Systems: Uncovering Public Preferences toward Participatory Research and Innovation. *She Ji: The Journal of Design, Economics, and Innovation*. 5(2), 128–146. DOI: https://doi.org/ggk5zn

Štatistický úrad Slovenskej republiky. (2021). Sčítanie obyvateľov a domov a bytov. URL: https://www.scitanie.sk/obyvatelia/zakladne-vysledky/pocet-obyvatelov/OB/SK031B517402/OB

UN. (2018). Data. URL: https://www.un.org/en/development/desa/population/publications/pdf/urbanization/the_worlds_cities_in_2018_data_booklet.pdf

Woods, E., Citron, R. (2016). Smart Technologies and Infrastructure for Energy, Water, Mobility, Buildings, and Government: Global Market Analysis and Forecasts. URL: https://kapost-files-prod.s3.amazonaws.com/uploads/asset/file/5ae1f293b0782600b300007e/2016%20Navigant-Executive-Summary.pdf

https://doi.org/10.55343/CogSust.84

Appendix

Dear Citizen, we would like to kindly ask you to express your opinion about life in our city, also on the basis of which we will begin to create a strategy for the concept of Smart City in the city of Žilina. With the Smart City concept, we want to create an environment that will bring the citizens of the city the opportunity to benefit from the introduction of sophisticated modern technologies that systematically complement each other and go beyond the perspective of one field (transport, energy, environment, etc.).

1. What is your age category?
   a) >20 b) 20–30 c) 31–40 d) 41–50 e) 51–65f) 65 >
2. Do you use public transport for transport around the city?
   a) Yes, regularly (Skip Q nr.4) b) yes, sporadically c) No (go to Q nr.4)
3. Are you satisfied with the frequency of the lines you normally use?
   a) Yes b) No
4. Why don't you regularly use public transport services for transport around the city?
   a) insufficient frequency of the line b) uncomfortable type of transport c) unsatisfactory routes of the lines
   d) use of other types of transport (car, train, bicycle…) e) other
5. Are you satisfied with the flow of car traffic in your city?
   a) Yes b) Mostly yes c) Mostly no d) No e) I don't know
6. Do you think the parking situation has been getting worse over last years in the city?
   a) Yes b) No c) I don't know
7. Would you use car sharing/rental services for city transport?
   a) Yes b) Probably yes c) Probably no d) No e) I don't know
8. On a scale of 1–10, rate the cleanliness of your city center (1 – the city center is constantly dirty, 10 – the city center is constantly clean)
   1 2 3 4 5 6 7 8 9 10
9. Are you satisfied with the frequency of waste collection at your place of residence?
   a) Yes b) Mostly yes c) Mostly no d) No e) I don't know
10. Are you interested in the economy of the city?
    a) Yes b) Mostly yes c) Mostly no d) No e) I don't know
11. Can you easily get information about the city's economy?
    a) Yes b) Mostly yes c) Mostly no d) No e) I don't know
12. Are you sufficiently informed about what is happening in the city?
    a) Yes b) Mostly yes c) Mostly no d) No e) I don't know
13. Which information channels between the city and the citizens do you use?
    a) webpage of the city b) local news (paper, or electronic) c) social media
    d) public debates e) information tables f) other g) none
14. Would you use a mobile application that would bring together the whole city agenda?
    a) Yes b) Probably yes c) Probably no d) No e) I don't know
15. On a scale of 1–10, rate the importance of free WIFI and mobile chargers & charging stations in your city (1 – absolutely irrelevant, 10 – very important)
    1 2 3 4 5 6 7 8 9 10
16. Is there enough sports and cultural activities in your city?
    a) Yes b) Mostly yes c) Mostly no d) No e) I don't know

If you want to, please express your opinion on what other areas the city should focus on?

Thank you for your time.

https://doi.org/10.55343/CogSust.81

THE ROLE OF ARTIFICIAL INTELLIGENCE IN THE DEVELOPMENT OF RAIL TRANSPORT

Peter FICZERE
0000-0003-3207-5501
Budapest University of Technology and Economics, Department of Railway Vehicles and Vehicle System Analysis
Budapest, Hungary
ficzere.peter@kjk.bme.hu

Abstract

Artificial intelligence plays a revolutionary role in modern transport systems. The article discusses the role of artificial intelligence in railway transport and its potential impact on the sector. This article presents different types of artificial intelligence technologies used in this sector, explores the advantages of artificial intelligence in this field, and discusses the challenges associated with using artificial intelligence in rail transport. Artificial Intelligence is revolutionary in rail transport systems by enhancing efficiency, safety, and overall performance. There are several ways in which AI influences rail transport and its impact on the cognitive load of human resources. These factors are examined in this article.

Keywords

Artificial Intelligence, Rail Transport, Rail Industry, Digital Automatic Coupling, Cognitive Load

1. Introduction

Rail transport is a mode of transportation that involves the movement of passengers and goods on wheels of vehicles running on railway tracks. It is a widely used and efficient means of mass transit and freight transport. The fundamental components of rail transport include:

Railway Tracks: Fixed tracks made of steel or other materials that guide the train and provide a stable path for its movement (*Shi et al., 2023*).

Train: A connected series of rail vehicles (such as locomotives and wagons) that move along the tracks as a single unit (*Waters, 2007*).

Locomotive: A powered railway vehicle that provides the motive power for a train. Locomotives are typically fuelled by electricity, diesel, or steam (*Babcock, Bunch, 2007*).

Cars or Wagons: Non-powered rail vehicles coupled and pulled by a locomotive. They can be designed for the transportation of passengers or cargo, depending on the type of service.

Human Resource: Nowadays, human labour is essential in the maintenance and operation of train services (*Crawford, Kift, 2018*).

Rail transport is known for its efficiency, particularly in moving large quantities of goods over long distances. It is also a popular mode of commuter transportation in many urban areas. The advantages of rail transport include lower fuel consumption per ton mile, reduced traffic congestion, and lower environmental impact compared to other modes of transportation. However, it requires significant infrastructure, such as well-maintained tracks, stations, and terminals. Rail transport is one of the world's oldest and most important modes of transport. Since its establishment, it has played a vital role in passenger and freight transport and is an important part of the transport system. However, the rail sector faces many challenges, including increasing competition from other modes of transport, ageing infrastructure and safety concerns. To address these challenges, the rail industry is turning to artificial intelligence technologies to improve its operations and provide better customer service.

https://doi.org/10.55343/CogSust.81

2. Applications of AI in rail transport

Artificial intelligence (AI) can revolutionise rail transport, and this paper explores the potential role of artificial intelligence in developing this field through the outstanding publications:

Artificial intelligence has played an increasingly important role in the development of rail transport, offering numerous benefits that improve efficiency, safety, and overall operations. Here are some key areas where artificial intelligence has influenced the development of rail transport.

Predictive maintenance: AI is used to analyse vast amounts of data from sensors on trains and tracks to predict potential equipment failures before they occur. This proactive approach helps schedule maintenance activities, reduce unexpected breakdowns, and minimise disruptions. As a result, maintenance crews can focus their efforts more efficiently, reducing the cognitive load associated with reactive problem-solving (*Yao et al., 2023*).

Autonomous trains: AI algorithms are employed in automatic train control systems. These systems can optimise train schedules, manage traffic, and ensure safe distances between trains. Human operators can concentrate on more strategic decision-making tasks by automating certain control functions, reducing the cognitive burden associated with routine operations (*Niu, You, 2022*).

Enhanced safety and security: Artificial intelligence-powered video analytics and sensor technologies can monitor and detect potential security threats or safety hazards in railway infrastructure. This helps prevent accidents and ensure the safety of passengers and rail personnel. AI contributes to digital signalling systems that enhance traffic management efficiency on rail networks. These systems can dynamically adjust train schedules based on real-time conditions, minimising delays and improving traffic flow. This reduces the cognitive load on human operators who would otherwise need to manage and adapt to changing circumstances manually. AI-powered safety systems, such as Positive Train Control (PTC) and collision avoidance systems, enhance safety by monitoring and controlling train movements. These systems can automatically apply brakes or take other corrective actions in emergencies, reducing human operators' cognitive load to react quickly to potential hazards (*Hartong, Goel, Wijesekera, 2011*).

Optimised operations and scheduling: AI algorithms optimise crew schedules, ensuring that human resources are allocated efficiently. This helps minimise fatigue and stress among railway personnel and ensures that the right personnel are available when needed.

Passenger experience: AI is also employed to improve passenger services, including ticketing, customer service chatbots, and personalised travel recommendations. This can reduce the cognitive load on staff dealing with routine inquiries and transactions.

Energy efficiency: Artificial intelligence can reduce energy consumption and optimise energy use in rail systems. Through data analysis and predictive modelling, artificial intelligence can help develop energy-efficient strategies and reduce the environmental impact of rail transport.

Infrastructure planning and management: Artificial intelligence can assist in the planning and management of railway infrastructure, including the design of new rail lines, maintenance of existing infrastructure, and the identification of areas for improvement.

Integrating artificial intelligence in the development of rail transport can revolutionise the industry, improving operational efficiency and safety and creating better customer experiences. While AI brings numerous benefits to rail transport, it is crucial to consider the human-machine interface and provide adequate training for personnel to understand and trust AI systems. Additionally, human oversight is often essential to handle unexpected situations and ensure the ethical and safe operation of the railway system. The successful integration of AI in rail transport requires a collaborative approach that considers both technological and human factors.

Several types of artificial intelligence technologies are used in rail transport. There are several types of ITS (Intelligent Transportation Systems) technologies. Machine learning is an artificial intelligence that allows machines to learn from data without being explicitly programmed (*Borsodi and Takács, 2022*). This technology is used in rail transport to develop predictive maintenance models that can help prevent breakdowns and reduce downtime. Another important type of artificial

https://doi.org/10.55343/CogSust.81

intelligence technology is computer vision. Computer vision is artificial intelligence that allows machines to recognise and interpret visual data, such as images and video (*Fedorko, 2021*). Computer vision is being used in rail transport to develop intelligent video surveillance systems that can help improve safety and security (*Ulewicz et al., 2019*).

3. Benefits

3.1. The Benefits of Using Artificial Intelligence in Rail Transport

There are many advantages to using artificial intelligence in rail transport. One of the most important ones is increased safety. Artificial intelligence technologies, such as predictive maintenance models and intelligent video surveillance systems, can help identify potential safety hazards and prevent accidents before they happen. Another benefit is increased efficiency. Artificial intelligence technologies can help optimise rail operations, reducing delays and improving customer service. In addition, artificial intelligence can help reduce costs by reducing the need for manual labour and optimising the allocation of resources. In the rail sector, artificial intelligence can support research and development in several ways, including (*Tang et al., 2022*):

- Predictive modelling and simulation: artificial intelligence can help researchers develop predictive models and simulations that simulate different scenarios and outcomes. For example, machine learning algorithms can be used to analyse data from railway systems – such as train schedules, traffic patterns, and weather conditions – to predict potential problems and improve the efficiency of railway operations (*Liu et al., 2023*).
- Data analysis and decision-making: artificial intelligence can help researchers and engineers analyse large amounts of data to make informed decisions. Machine learning algorithms can identify patterns and correlations in the data, allowing researchers to make predictions and recommendations based on the data (*Lieophairot, Rojniruttikul, 2023*).
- Autonomous systems: artificial intelligence can also be used to develop and optimise autonomous systems for rail operations. For example, artificial intelligence can be used to develop algorithms that allow trains to operate autonomously, reducing the need for human intervention.
- Optimising resources: Artificial intelligence can be used to optimise the use of resources in railway operations. Machine learning algorithms can be used to analyse train schedules, maintenance schedules and other operational data to identify areas where resources can be optimised, for example, to reduce downtime or increase efficiency.
- Risk management: Artificial intelligence can help researchers and engineers identify potential risks to rail operations and develop strategies to mitigate them. For example, machine learning algorithms can analyse train accident data to identify patterns and potential risks.

In summary, artificial intelligence can be a useful tool to support research and development in the rail sector, helping to improve efficiency, safety and reliability.

3.2. Digital Automatic Coupling

Digitalisation is one of the most important steps in properly taking advantage of new opportunities. An essential element of this is the introduction of digital automatic coupling (DAC). Digital Automatic Coupling (DAC) is a concept and technology aimed at improving the efficiency and automation of coupling and uncoupling railway vehicles (such as wagons). It involves the use of digital communication and control systems to automate the process of connecting and disconnecting railcars. The traditional coupling method involves manual labour, where railway workers physically connect and secure the couplings between train cars. Digital automatic coupling seeks to replace or augment this manual process with automated, digitally controlled mechanisms. Digital automatic coupling technology can improve the efficiency of rail transport in several ways:

- Faster and more efficient coupling: the DAC technology enables faster and more efficient coupling and uncoupling of rail wagons. This can reduce the time needed to load and unload cargo and improve the overall efficiency of rail operations.
- Increased safety: DAC technology can improve safety by providing more accurate and up-to-date information on the position and movement of train wagons. DAC can help reduce the risk of collisions and other accidents (Takacs, 2023; Mekonnen et al., 2023).
- Real-time tracking: DAC technology can enable real-time tracking of the position and movement of trainsets, allowing control centres to make more informed decisions on train scheduling, routing, and other operational issues. This can help reduce delays and improve the overall efficiency of rail operations.
- Increased capacity: DAC technology can increase the capacity of rail wagons by enabling more accurate coupling and uncoupling. DAC can help reduce the number of wagons needed to transport a given amount of freight, improving efficiency and reducing costs.

https://doi.org/10.55343/CogSust.81

- Better planned maintenance: DAC technology can improve maintenance operations by providing real-time information on the components and performance of rail wagons. This can help maintenance teams identify and repair potential problems faster, reducing downtime and improving efficiency.

4. Challenges and potential threats

4.1 Challenges for the application of artificial intelligence in rail transport

Thus, DAC technology can improve rail transport efficiency, safety and reliability through faster and more efficient interconnection, real-time monitoring and more accurate information on the position and movement of train sets.

While introducing DAC has clear benefits, it may not be widespread shortly. This is because rail transport is fragmented. Separate departments are responsible for traction, passenger transport, freight and infrastructure management. As these outsourced companies are economically independent, there is no common interest. Passenger and freight wagons must be equipped with DAC equipment, so the costs would logically be linked to these departments. However, the profits are expected to be realised by the infrastructure provider. Understandably, the companies operating and servicing the wagons do not want to spend around €20,000 per wagon, as they will never profit. At the same time, the company that maintains and manages the infrastructure does not want to spend money upgrading wagons owned by another company. Another difficulty is that the profitability of old freight wagons is very uncertain.

To solve this problem, the companies in the different areas would have to work together, but none is interested in doing so separately. However, the social benefits are unquestionable so top-down legislation could be the solution. It is important to note that this investment does not make sense in small steps, and the benefits can only be realised if everyone simultaneously moves to the new system. This is, of course, the most important thing for international traffic. It is also important that standardised switching equipment is introduced.

As with the economic benefits (*Evans, 2013*), this is the only way to maximise the potential of artificial intelligence. The resulting increase in capacity could also significantly reduce passenger and freight traffic on roads, reducing emissions (*Torok and Sipos, 2022*).

4.2. Potential threats to the application of artificial intelligence in rail transport

While digitalisation can bring many benefits to rail transport, potential threats must be addressed. Some of the main risks of digitalisation for rail transport are outlined below (*Török, 2023*):

- Cybersecurity risks: As rail systems become increasingly interconnected and rely more and more on digital technology, they may become more vulnerable to cyberattacks. This could include attacks against rail IT systems, control systems or trains. Cybersecurity threats can disrupt rail operations, compromise passenger safety, and cause significant financial losses.
- Reliance on technology: While digitalisation can improve the efficiency and reliability of rail services, it also means that rail companies rely more on technology. This means that significant disruption to rail operations can occur in the event of a technological failure (*Hussain, Zefreh, Torok, 2018*).
- Job losses: using digital technology in the rail sector could lead to job losses as automation and artificial intelligence take over some of the tasks previously done by humans. This could have a significant impact on workers and their communities.
- Privacy concerns: As digitisation generates large amounts of data about rail passengers and operations, there are concerns about how these data are collected, stored and used. There is a risk that this data could be misused or hacked, leading to privacy breaches and other negative consequences.
- Infrastructure Challenges Digitisation requires significant investments in IT infrastructure and communication networks, which can be costly and time-consuming. Railway operators may need to upgrade their infrastructure to support digitalisation, which may disrupt the implementation phase.

Recognising the potential threats of digitalisation in rail transport is essential, and measures must be taken to mitigate them. This includes investing in cyber security, ensuring that workers are properly trained and protected, addressing data protection concerns, and carefully managing the introduction of new digital technologies.

Despite the many benefits of artificial intelligence in rail transport, its implementation faces several challenges. One of the biggest challenges is the cost of implementing artificial intelligence technologies. Implementing artificial intelligence requires significant investments in hardware, software and personnel. Another challenge is the lack of standardisation in the industry. Standardised data formats and communication protocols are needed to integrate different artificial intelligence systems. Finally, privacy and security concerns may also arise. For example, intelligent video surveillance systems raise concerns about using personal data and the potential for misuse.

https://doi.org/10.55343/CogSust.81

5. Summary

In conclusion, artificial intelligence plays an increasingly important role in the development of rail transport. The benefits of artificial intelligence, including improved safety, greater efficiency, and cost reduction, make it a promising technology. Despite the challenges, the future of artificial intelligence in rail transport is promising. Artificial intelligence technologies are expected to continue to play an important role in the development of rail transport, and new technologies and applications will emerge over time. One area of particular interest is using artificial intelligence in autonomous trains. Autonomous trains could revolutionise rail transport, improving safety and efficiency while reducing costs. Another interesting area is artificial intelligence in customer service, where chatbots and virtual assistants are being developed to improve the customer experience.

Rail transport has been a vital mode of transport for both people and goods for many years, but it faces many challenges, such as competition from other modes, ageing infrastructure and safety risks. Artificial intelligence-based technologies can revolutionise rail transport by improving efficiency, reducing costs, and increasing safety. Machine learning and computer vision are among the technologies based on artificial intelligence used in rail transport. Artificial intelligence can support research and development in the rail sector in several ways, including predictive modelling and simulation, data analysis, autonomous systems development, resource optimisation, and risk management. Digital automatic coupling (DAC) technology can further improve the efficiency of rail transport through faster and more efficient coupling, improved safety, real-time monitoring, and increased capacity. However, the fragmentation of the rail sector may limit the deployment of DAC technology shortly.

References

Babcock, M. W., & Bunch, J. L. (2007). Energy Use and Pollutant Emissions Impacts of Shortline Railroad Abandonment. *Research in Transportation Economics*, 20, 225-257. DOI: https://doi.org/c5cww8

Borsodi, E., Takács, Á. (2022). Generative Design: An Overview and Its Relationship to Artificial Intelligence. *Design of Machines and Structures*. 12(2), 54–60, DOI: https://doi.org/k6gc

Crawford, E. G., & Kift, R. L. (2018). Keeping track of railway safety and the mechanisms for risk. *Safety Science*, 110, 195-205. DOI: https://doi.org/k6fw

Evans, A. W. (2013). The economics of railway safety. *Research in Transportation Economics*. 43(1), 137–147. DOI: https://doi.org/k6gd

Fedorko, G. (2021). Application possibilities of virtual reality in failure analysis of conveyor belts. *Engineering Failure Analysis*. 128, 105615. DOI: https://doi.org/grb69m

Hartong, M., Goel, R., & Wijesekera, D. (2011). Positive train control (PTC) failure modes. *Journal of King Saud University-Science*, 23(3), 311-321. DOI: https://doi.org/fbrdps

Hussain, B., Zefreh, M. M., & Torok, A. (2018). Designing the appropriate data collection method for public transport passenger satisfaction analysis. *International Journal for Traffic & Transport Engineering*, 8(2). DOI: https://doi.org/k6gk

Lieophairot, C., & Rojniruttikul, N. (2023). Factors affecting State Railway of Thailand (SRT) passenger train service use decision: A structural equation model. *Heliyon*, 9(5). DOI: https://doi.org/k6gb

Liu, X., Dabiri, A., Xun, J., & De Schutter, B. (2023). Bi-level model predictive control for metro networks: Integration of timetables, passenger flows, and train speed profiles. *Transportation Research Part E: Logistics and Transportation Review*, 180, 103339. DOI: https://doi.org/k6f2

Mekonnen, A. A., Sipos, T., Szabó, Zs. (2023). Generalised Linear Modeling of Crashes on Urban Road Links. *Periodica Polytechnica Transportation Engineering*. 51(2), 140–146. DOI: https://doi.org/k6gf

Niu, R., & You, S. (2022). Research on run-time risk evaluation method based on operating scenario data for autonomous train. *Accident Analysis & Prevention*, 178, 106855. DOI: https://doi.org/k6fz

Shi, C., Zhou, Y., Xu, L., Zhang, X., & Guo, Y. (2023). A critical review on the vertical stiffness irregularity of railway ballasted track. *Construction and Building Materials*, 400, 132715. DOI: https://doi.org/gsj9bq

Takacs, A. (2023). Safe In and Out of the Car. In: Jármai, K., Cservenák, Á. (eds) *Vehicle and Automotive Engineering* 4. VAE 2022. Lecture Notes in Mechanical Engineering. Springer, Cham. DOI: https://doi.org/k6gg

Tang, R., De Donato, L., Besinović, N., Flammini, F., Goverde, R. M., Lin, Z., ... & Wang, Z. (2022). A literature review of Artificial Intelligence applications in railway systems. *Transportation Research Part C: Emerging Technologies*, 140, 103679. DOI: https://doi.org/gp58cf

Török, Á. (2023). Do Automated Vehicles Reduce the Risk of Crashes – Dream or Reality? *IEEE Transactions on Intelligent Transportation Systems*. 24(1), 718–727. DOI: https://doi.org/k6gj

Torok, A., Sipos, T. (2022). Can the Marginal Cost Be Extended to Life Cycle Cost? A Theoretical Case Study for Transport. *International Journal for Traffic and Transport Engineering*. 12(2). DOI: https://doi.org/k6gh

Ulewicz, R. Nový, F., Novák, P., Palček, P. (2019). The investigation of the fatigue failure of passenger carriage draw-hook. *Engineering Failure Analysis*, 104, 609–616. DOI: https://doi.org/gmj8vd

Waters II, W. G. (2007). Evolution of railroad economics. *Research in transportation economics*, 20, 11-67. DOI: https://doi.org/dsw492

Yao, Y., Wu, L., Xie, B., Lei, L., Wang, Z., & Li, Y. (2023). A two-stage data quality improvement strategy for deep neural networks in fault severity estimation. *Mechanical Systems and Signal Processing*, 200, 110588. DOI: https://doi.org/k6fx

https://doi.org/10.55343/CogSust.80

Exploring Cognitive Sustainability Concerns in Public Responses to Extreme Weather Events: An NLP Analysis of Twitter Data

Rihem Berbère
National Engineering School of Tunis, University of Tunis El Manar,
Tunis, Tunisia
rihem.berbere@etudiant-enit.utm.tn

Safa Elkefi
0000-0002-4293-0404
School of Nursing
Columbia University
New York, USA
se2649@cumc.columbia.edu

Safa Bhar Layeb
0000-0003-2536-7872
LR-OASIS
National Engineering School of Tunis, University of Tunis El Manar,
Tunis, Tunisia
safa.layeb@enit.utm.tn

Achraf Tounsi
0000-0001-8166-616X
Department of Civil, Environmental & Ocean Engineering
Stevens Institute of Technology
Hoboken, USA
atounsi@stevens.edu

Abstract

The United States has a long history of experiencing extreme weather events. Hurricanes are among the most devastating natural disasters that have significant economic and physical impacts on the country. By applying Natural Language Processing (NLP) to Twitter data for sentiment analysis, emotion detection, and topic modelling, this study provides a more thorough understanding of public response and concerns during five study cases of hurricanes that hit the United States: Harvey, Irma, Maria, Ida, and Ian. The findings on sentiment analysis revealed that 64.75% of the tweets were classified as Negative and 35.25% as Positive. For emotion detection, the predominant emotion was anger, with 39.91%. These results were centred around the main public concerns shown by the topic modelling: hurricane management, donation and support, and disaster impacts. Our future work will focus on understanding people's responses to extreme weather events through the evolving concept of Cognitive Sustainability.

Keywords

Hurricane, people response, Sentiment Analysis, Emotion detection, Topic Modeling, Natural Language Processing.

1. Introduction

Extreme weather events (EWE), characterized by their increasing intensity and frequency, clearly indicate the impact of climate change (*Fischer, Knutti, 2015*). These disasters are defined as events in which some meteorological variables, such as strong wind speeds (e.g., hurricanes), reach extreme values that can lead to tragedies, resulting in serious physical and financial damages, with significant impacts on healthcare systems (*Stephenson, Diaz, Murnane, 2008; Ebi, Bowen, 2016*). The United States (U.S.) has a longstanding history of experiencing EWE, as evidenced by 355 weather-related and climatic disasters since 1980. These events have resulted in substantial losses and costs exceeding $2.540 trillion, demonstrating their significant economic impact on the country. Among the most devastating EWEs are hurricanes, which claim a high number of lives while destroying the infrastructure and damaging the healthcare system on numerous levels (*Waddell et al., 2021*). For example, Hurricane Harvey in 2017, the second-costliest natural disaster in U.S. history after Hurricane Katrina, accounted for $152.5B and more than 75 total deaths (*Oluyomi et al., 2021, NOAA National Centers for Environmental Information 2023*).
Similarly, Hurricane Ian, which impacted Florida in 2022, being the third costliest hurricane in U.S. history, accounted for $114.0B in economic losses and 148 reported deaths in mid-January. The severe consequences of these hurricanes extend beyond economic losses. The most tragic outcome is the loss of lives, often resulting from drowning, electrocution, or bodily

https://doi.org/10.55343/CogSust.80

injury, which is the most severe acute result of a hurricane. Mental health issues, on the other hand, are a critical aspect of public concerns and can be severely amplified or result in hurricane disasters (*Ebi et al., 2021*). Major effects of Hurricane Katrina included increases in unhealthy behaviours, post-traumatic stress disorder (PTSD), anxiety, and elevated relapse rates for pre-existing health conditions. More than 50% of evacuees with persistent and severe psychological effects are less likely to receive treatment contact from healthcare providers (*Mills, Edmondson, Park, 2007*).

Moreover, hurricanes can cause significant damage to infrastructure systems, affecting mostly vulnerable populations and working-class communities that struggle in the long term to recover (*Dash et al., 2007*). Electric power outages, in particular, can be very disruptive and affect other critical infrastructures on a high level, such as transportation and healthcare services (*Mitsova et al., 2019*). For instance, the major impact of Hurricane Maria on Puerto Rico was the electric power outage that initially affected the island and lasted more than ten months. The island's orography, logistical constraints, and the severity of the destruction all had a significant role in the exceedingly delayed restoration process that resulted in the island's power grid having very low resilience (*Kwasinski et al., 2019*). During these disasters, the public responds in various ways to cope with the potential damages. Emergency preparedness, evacuation, and migration become a priority, leading people to relocate to safer areas and create emergency plans by seeking timely and accurate information from local authorities and disaster management agencies and relying on community networks to share resources and support (*Sutton, Tierney, 2006*). This collective response reflects the importance of effective communication during emergencies in sharing resources, providing support, and disseminating critical updates (*Lu, Yang, 2011*).

Sustainability refers to the ability to meet the needs of the present without compromising the ability of future generations to meet their own needs. It involves making choices and implementing practices considering the long-term health of the environment, society, and the economy. Sustainability aims to balance social, economic, and environmental factors to ensure that resources are used in a way that preserves and protects them for the benefit of current and future generations. In this article, several levels of sustainability are considered. Cognitive sustainability is not a widely recognized term within the mainstream discussion on sustainability. However, it can be inferred to relate to maintaining cognitive well-being and resilience, particularly in the context of the challenges posed by the modern information age. Social sustainability could be defined as the ability of a society or community to develop in a way that meets the needs of the present without compromising the ability of future generations to meet their own needs. It focuses on fostering social well-being, equity, and justice. Physical sustainability could refer to individuals' or communities' long-term health and well-being. The physical aspects of sustainable development focus on the built environment and infrastructure. This might include the design and construction of buildings, transportation systems, and urban spaces that minimize environmental impact, use resources efficiently, and contribute to the overall resilience of communities. Emotional sustainability could also be the sustainability of interpersonal relationships and social connections. This might include fostering healthy relationships, effective communication, and community support. Building and maintaining positive social connections is essential.

Understanding this intricate web of responses through the lens of cognitive sustainability reveals the deeper cognitive processes at play. The prioritization of safety and the collaborative efforts to gather and disseminate information showcase the cognitive adaptability of individuals and communities in the face of environmental crisis, aligning with the main characteristics of cognitive sustainability. The ability to collectively navigate challenges and optimize human value creation during extreme weather events reflects a sustainable cognitive approach to social, emotional, and the implication of sustainability levels, harmonizing artificial and biological cognitive systems with the broader goal of enhancing resilience and well-being (*Zoldy et al., 2022*)**.** In this context, and with the increasing number of natural disasters, traditional methods for analyzing public concerns, such as surveys, have become constrained, expensive, and time-consuming (*Karami et al., 2020*). However, the emergence of social media has revolutionized how we track and understand people's concerns, offering unique and efficient ways of capturing real-time public sentiment and reactions. By gathering tweets from those who were impacted by Hurricanes Maria, Harvey, Irma, Ida, and Ian and using Natural Language Processing (NLP) techniques to analyze the textual data, we can examine people's responses, with the Twitter platform being a valuable source of real-time data due to its wide user base and potential for capturing diverse perspectives.

In 2023, Twitter will have 353.9 million users (please note that in April 2023, Twitter merged with X.), offering real-time feedback and using a timestamp to update users on conversations. As a result, Twitter has the potential to be a trustworthy and pertinent data source and offers a special chance to comprehend users' issues and contribute to enhancing situational awareness with a low waiting time access to data. Twitter has diverse applications, from finance to politics and health. However, because of problems with deception, attention-grabbing content, performance measurement, dependability, and the inability to independently check both the poster and the content they publish, using microblogging feeds as information sources during a significant event is still challenging (*Neppalli et al., 2017*). Still, research has been conducted to encourage outcomes. For example, a framework and Web-based system called Twitcident was developed by *Abel et al. (2012)* to monitor and filter information from social media streams, notably Twitter. Also, the extraction, processing, and classification of Twitter feeds have advanced significantly thanks to machine learning and natural language processing. A study on using this platform in public health and disasters shows that people are likely to communicate their feelings and thoughts about disasters on social

https://doi.org/10.55343/CogSust.80

media sites (*Vieweg et al., 2010*). Consequently, crisis response managers can benefit from social media by better knowing how the public reacts to disasters and analyzing their sentiments and expressions (*Yuan, Li, Liu, 2020*).

The EWE we worked on occurred when communities had access to vast amounts of data, which presented an opportunity to advance public health systems using cutting-edge tools like data-driven artificial intelligence (AI). Natural language processing (NLP) is a subset of AI technologies with significant promise in pinpointing trends and sentiments related to particular themes (*Al-Garadi, Yang, Sarker, 2022*). When tweets are gathered using Twitter's Academic API, the major goal is to determine the sentiment of each one by defining positive and negative polarity, detecting emotions, and extracting different related topics with topic modelling. Compared to surveys or questionnaires, this can be a simpler method of retrieving user's opinions and thoughts. Numerous research studies have used NLP to address public response and community vulnerabilities during disasters using sentiment and topic modelling approaches (*Albahli et al., 2021*). For instance, during Hurricane Laura in 2020, research proposed a supervised approach from NLP models with a Latent Dirichlet Allocation (LDA) workflow to investigate temporal latent topics from tweets to retrieve real-time situational awareness (SA) (*Zhou et al., 2023*).

This work fills this gap by thoroughly examining the public response during extreme weather events, taking into account the study case hurricanes that occurred between 2017 and 2022 in the USA. With this prolonged time frame, evaluating how people's concerns have evolved, noting shifting dynamics, and spotting patterns or trends in susceptibility is possible. This work, framed within the exploration of cognitive sustainability in public responses to extreme weather events, offers a more comprehensive knowledge of the effects of extreme weather events on individuals by looking at a sequence of events rather than isolated incidents, especially with the adoption of a novel strategy with Twitter being a rich source of real-time information. We aim to provide a more thorough and nuanced understanding of these concerns in such contexts by applying NLP analysis to Twitter data and looking at the voices and emotions of people directly or indirectly affected by these extreme weather events.

2. Data and methods

2.1 Extreme weather events

In this study, we aim to understand the public response during EWE through the lens of cognitive sustainability. We selected five hurricanes that took place between 2017 and 2022. These events include Hurricane Maria, Hurricane Harvey, Hurricane Irma, Hurricane Ida, and Hurricane Ian. Collecting tweets, we delve into people's cognitive, social, physical, and emotional sustainability concerns and responses to these disasters, examining their sentiments and expressions. Our selection of the hurricanes was mainly based on their severity as determined by their attributes. We first focused on hurricanes from the past ten years. By narrowing our scope to recent events between 2017 and 2022, the selection of certain hurricanes was determined by their effects on finances and mortality, with a concentration on hurricanes that caused large losses in property and lives. We focused on hurricanes that were at least Saffir-Simpson Category 4 in severity and escalated to Category 5 in some cases.

2.2 Twitter data

In this study, 113,475 tweets were collected from 2017 to 2022 and were stored in Twitter's Search API, which was used for collecting tweets in this research. Twitter's Standard Search API returns tweets from the previous seven days based on user-specified search criteria. In this study, we conducted keyword- and hashtag-based research for data collection using the "search/tweets" endpoint of the Twitter API. A set of carefully selected queries was used to achieve this, which could be combined using logical operators such as 'OR' and 'NOT'. This approach allowed for precise filtering of tweets, ensuring that the collected data pertained to the selected hurricanes and encompassed keywords related to hurricane impacts. We specified different queries for each timeline event and removed the replies and retweets. Then, we fetched the corresponding tweets located in the USA from specific time ranges. The collection was done over a 7 to 10-day period for each hurricane, which allowed the capture of reactions from the beginning to the end of the hurricane event during the preparation and impact phases. This is particularly useful as it allows us to collect many tweets containing relevant keywords in a single request. The endpoint supports pagination, which means we can retrieve up to 500 tweets per request. Another reason why this method was chosen is its flexibility and precision. Table I provides a summary of the Twitter data used in our study.

https://doi.org/10.55343/CogSust.80

Table 1. Twitter data summary

Hurricane Name	**Collection / start_time**	**Collection/ End_time**	**Number of tweets collected**
Ian	2022-09-22	2022-10-02	12 860
Ida	2021-08-26	2021-09-05	9 680
Harvey	2017-08-23	2017-09-02	16 204
Irma	2017-08-03	2017-08-12	20 638
Maria	2017-09-30	2017-10-10	54 093

Source: own compilation

2.3 Data cleaning

For the preprocessing stage of our study, we cleaned the collected Twitter dataset (only English data were considered) to meet the specific requirements of BERT and BERTopic *Fig. 1* illustrates the steps followed (see chapter 2.4). The cleaning process involved removing noise by removing special characters, numbers, and emojis from the tweets to concentrate on the textual content. It also removed HTML, different mentions (i.e., words starting with '@' and '#'), symbols, flags, Unicode characters, and punctuation. In addition, particular user references and linkages were removed from remarks and URLs. The text was changed to lowercase for consistency, and any leading or trailing whitespace was removed. To guarantee a unique dataset for analysis, duplicate tweets were also found and eliminated. The total number of tweets dropped from 113,475 to 68,373 after cleaning, creating a more accurate dataset that can be used for BERT model training and evaluation. We refrained from employing conventional text preprocessing techniques like lemmatization and stemming. This decision was underpinned by our use of BERT, a cutting-edge transformer-based model, which inherently captures contextual meaning and ensures the preservation of the text's richness and nuance.

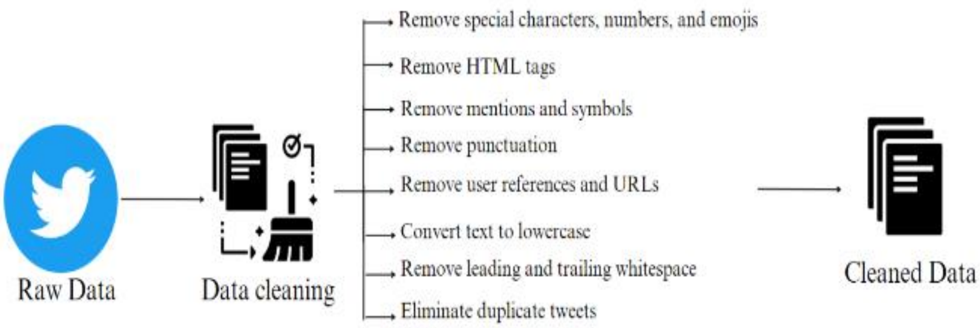


Fig. 1. Data cleaning process

2.4 Bert Model

This study used BERT for sentiment analysis, emotion detection, and topic modelling. BERT is a multi-layered structure of Bidirectional Transformer encoder layers based on the architecture of the transformer (*Nguyen et al., 2020*). The main benefit of BERT is the application of two-dimensional transformer training techniques. BERT combines how to look at a text string from 2 dimensions. This technique can significantly increase the retention of word representations in sentences. BERT models are an effective tool for sentiment analysis for binary and multiclass classification (*Catelli, Pelosi, Esposito, 2022*).
Additionally, BERT is also suitable for emotion detection; it can successfully categorize text into particular emotional categories by being trained on datasets that include various emotions (*Acheampong, Nunoo-Mensah, Chen 2021*). Our study used two pre-trained Bert-based models: The roBERTa-based model from Hugging Face for sentiment analysis and the BERT-base uncased emotion model for emotion detection. These models were chosen based on their established performance and suitability for our study. The RoBERTa-based model from Hugging Face is a state-of-the-art model that performs three tasks:
(i) sentiment analysis,
(ii) hate speech detection, and
(iii) offensive language identification.

https://doi.org/10.55343/CogSust.80

The model is trained on approximately 58 million tweets using the TweetEval benchmark. The sentiment analysis task aims to recognize if a tweet is positive, negative, or neutral. The training set comprised 45,389 tweets, the validation set included 2,000 tweets, and the testing set consisted of 11,906 tweets from the Semeval2017 dataset. The macro recall score for the testing performance was 68.5 (*Barbieri et al., 2020*).

Similarly, the BERT-base uncased emotion model is a popular and widely used model designed for emotion analysis tasks. It was trained on a diverse range of emotions and fine-tuned using PyTorch Lightning, providing a validation accuracy (val_acc) of 0.931. It has demonstrated excellent performance in various benchmark datasets. It has proven effective in accurately predicting emotions and their suitability with the specific emotions we aimed to analyze in our study.

For Topic modelling, we used BERTopic. This model has emerged as a promising alternative for topic modelling techniques like Probabilistic Latent Semantic Analysis and Latent Dirichlet Allocation. BERTopic offers several benefits over these traditional methods, frequently requiring complex changes and morpheme analysis for the best outcomes. Other models frequently ignore the semantic connections between words, but this model does not. BERTopic creates accurate and coherent topic expressions using the clustering technique and TF-IDF's class-based transformation. It incorporates BERT-based embedding and c-TF-IDF word weights during the text embedding stage to improve the quality of topic representations. Using text clustering, BERTopic also excels at locating pertinent subjects throughout various domains (*Sun et al., 2023*).

Applying topic modelling techniques helps to uncover hidden themes and "topics" within our dataset. Analyzing the words in the dataset, we identified patterns and connections that revealed certain topics discussed. We used BERTopic, simplifying the topic modelling process by employing various embedding techniques and c-TF-IDF to create dense clusters. Our approach involved extracting the 40 most common topics from the dataset. We then manually selected the most relevant topics and grouped them together based on the similarity between their keywords. For instance, we observed that 'topic 2' was characterized by keywords such as "donation," "fundraiser," "proceeds," and "donating," while 'topic 14' featured keywords such as "donation," "donate," "collecting," and "fundraising." Recognizing the similarity between these keywords, we grouped these topics under the same category: "hurricane and healthcare management." We plotted an inter-topic distance map to extract even more relevant topics, allowing us to observe clusters of closely related topics.

3. Results

3.1 General analysis

The dataset contains 113 475 tweets. We analyzed the average number of tweets during the day and night periods, provided in Table 2. We observed that the number of tweets during the day (from 5 am to 8 pm) exceeded those during the night (from 8 pm to 5 am). However, the number of night tweets remains high compared to the total number of Tweets in the dataset, emphasizing the ongoing discussion of disaster-related topics even during late hours. This highlights individuals' significant engagement and interest in addressing and sharing information about hurricanes throughout the day, including nighttime.

Table 2. Hurricane Twitter Activity Analysis

Hurricane	Total tweets	Day tweets	Night tweets	The average number of tweets per whole day
Ian	12860	8551	4309	459.28
Ida	9680	5938	3742	312.25
Harvey	16204	10022	6182	540.13
Irma	20638	11748	8890	665.74
Maria	54093	30660	23433	1865.27

Source: Own compilation

3.2 Word Cloud & Frequency Analysis

In order to address public concerns, we conducted a thorough analysis of the dataset, focusing on word frequency and Word cloud visualization techniques. The Word cloud analysis aimed to filter out irrelevant words in the first step, including political references and random speech-related terms. The second step consisted of excluding hurricane-related terms, such as the names of specific hurricanes. We aimed to highlight the key terms and concepts about people's responses and concerns. This approach allowed us to gain valuable insights into the predominant discussions during extreme weather events. Table 3 provides the most frequently used words shown by the Word cloud.

https://doi.org/10.55343/CogSust.80

Table 3: Most common Word cloud terms

Extracted terms	***Frequency***
emergency	19,269
evacuation	8,531
help	5,218
relief	3,821
disaster response	2,962
hospital	2,405

Source: own compilation

The word cloud in Fig. 2 provided valuable insights into the prevalent topics among the tweets collected. Several recurring emergency terms were frequently mentioned. Words such as “affected” highlight the widespread impact of these hurricanes on different areas of the USA. Other terms such as ‘family’, ‘home’, and ‘shelter’ emerged as key concerns. Medical assistance and healthcare facilities were represented by “emergency” and “hospital” words. Other terms worth noting are “victim” and “death”, which suggest the occurrence of an unfortunate loss of life and the potential emotional impact on individuals and families. On a more positive note, the mention of “donate” and “relief” indicates the existence of a spirit of generosity and a willingness among people to contribute and support relief efforts. It may indicate the need for support and donations during the emergency crisis, especially with the presence of phrases such as “mandatory evacuation”, “help”, and “supply resources” that signify the importance and urgency of evacuation measures for the affected areas and the challenges encountered to rebuild and provide essential aid and resources. This comprehensive analysis offers a glimpse into the cognitive concerns, needs, and actions of individuals, local government authorities, and healthcare providers, underscoring the importance of effective disaster response strategies and the strength of community support systems.

Fig. 2. Word cloud

3.3 Temporal analysis

For a temporal analysis, the study examined the temporal distribution of tweets for each of the five hurricanes. Specifically, we present Fig. 3 for Ida tweets distribution, which was the most representative as it captured the trend from the landfall to the dissipation of the hurricane over time. The collection of Ida-related tweets ranged from August 26, 2021, to September 5, 2021, covering the initial stages and the dissipation of the hurricane. On September 1st and 2nd, 2021, Hurricane Ida significantly impacted the Northeastern United States, becoming an extratropical cyclone. This resulted in catastrophic river and flash flooding, particularly affecting regions like Louisiana (*Shultz et al., 2022*). The histogram presented in Fig. 3 highlights a sharp increase in tweets one day after the disaster, reflecting the immediate aftermath and the public response to the event. The slight delay in the burst of tweets can be attributed to the progressive nature of the hurricane. As time progressed and the hurricane dissipated, the number of tweets per day gradually decreased, as depicted in the figure. This observation suggests public attention and discussion around the hurricane waned over time.

https://doi.org/10.55343/CogSust.80

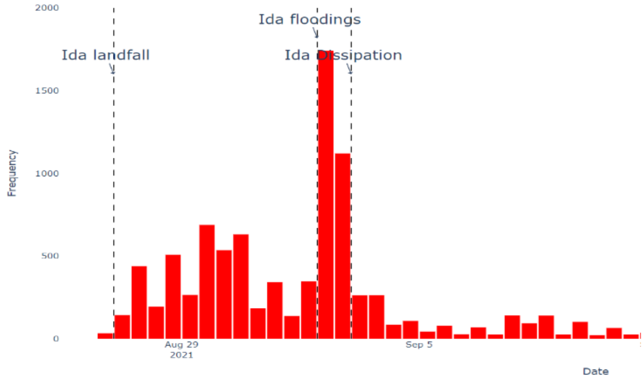


Fig. 3. Hurricane Ida Tweet's distribution

3.4 Sentiment analysis

In our study, we conducted the sentiment analysis for tweets using a pre-trained Roberta-base model from Hugging Face. We ran the model for the sentiment task. The sentiment analysis aims to recognize if a tweet is positive, negative or neutral. Noteworthy outcomes were found for the five hurricane cases. The model predicted a distribution of sentiments, with 64.75% classified as Negative, 35.25% as Positive, and 0.0% as Neutral. The sentiment analysis provided a binary classification for negative and positive sentiments, so our analysis will focus on this binary framework. We discovered distinct temporal patterns in the sentiments expressed, closely aligned with the real-life situations experienced during the hurricanes. The percentage of negative tweets indicating hurricane-related sentiments (64.75%) exhibited similar patterns across all five hurricanes. In each case, the number of tweets sharply increased as the hurricane approached and peaked, corresponding to the hurricane's intensity and impact on affected areas, such as floods and heavy rains. The hurricanes profoundly impacted people and the environment, as reflected in the sustained social media discussions.

We selected Hurricane Harvey and Irma as representative examples to explore the temporal patterns of sentiment predictions and their relationship with hurricane paths. Fig. 4 displays the sentiment distribution during the occurrence of Irma and Harvey. As Hurricane Harvey moved closer and made landfall in Texas, sentiments became more negative, indicated by the prevalence of negative terms, as shown in Fig. 4. This shift in sentiment likely resulted from the heavy rains and damage caused by the storm across Texas and Louisiana.

Following the dissipation of Harvey, Hurricane Irma, reaching Category 5 on September 4, 2017, emerged as the subsequent hurricane of interest. Irma inflicted widespread and catastrophic damage, particularly in the northeastern Caribbean and the Florida Keys. The predicted sentiments align with the temporal trends of hurricane landfall and intensity, evidenced by the substantial gap between positive and negative sentiments. As the relief efforts were initiated, evacuation procedures were implemented, disaster response management came into play, and the gap between sentiments gradually narrowed. These findings highlight the dynamic interplay between hurricane occurrences and the sentiments expressed on social media. They also underscore the crucial aspect of non-physical sustainability at the social and emotional levels. This dimension of sustainability is particularly evident in the evolving sentiments and nuanced temporal patterns observed during the occurrence and aftermath of hurricanes like Harvey and Irma. The emotional sustainability aspect becomes pronounced as the community navigates through the challenges posed by extreme weather events. It accentuates the vital role of relief efforts and emergency response measures in fostering emotional resilience and social cohesion. Thus, our analysis contributes to understanding sentiment dynamics and sheds light on the inherent sustainability challenges and coping mechanisms at the non-physical levels.

https://doi.org/10.55343/CogSust.80

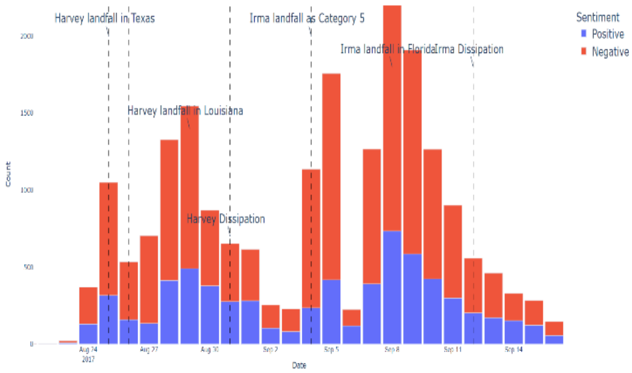


Fig. 4. Sentiment distribution during hurricanes: Irma and Harvey

3.5 Emotional analysis

We used the pre-trained BERT-based " bert-base-uncased-emotion " model from Hugging Face for the emotion detection analysis. This model was further fine-tuned on an emotion dataset using PyTorch Lightning. The selected model can predict six emotions: Anger, Fear, Sadness, Joy, Surprise, and Love. These emotions were deemed suitable for our study case, allowing us to capture a broad range of emotional responses. The choice of the model was based on the number of users and the evaluation parameters, and it has exhibited accurate results. Fig. 5 provides the emotion percentages in the dataset.

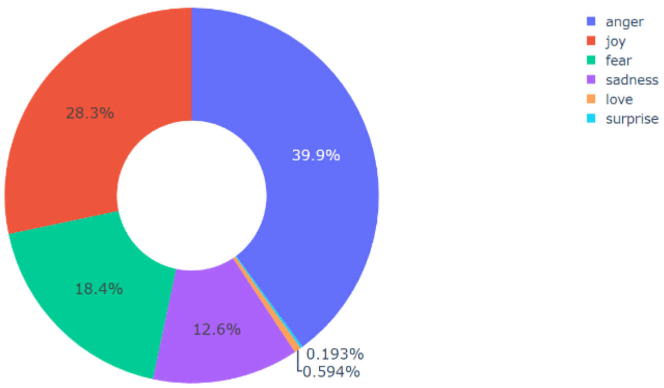


Fig. 5. Emotions distribution during all hurricanes

The results for emotion analysis revealed a notable distribution, with anger comprising the highest percentage at 39.91%, followed by joy at 28.30%. Fear and sadness accounted for 18.38% and 12.63%, respectively, while Love and Surprise had minimal representation, with percentages of 0.59% and 0.19%, respectively. For further analysis, we will concentrate on the four primary emotions: anger, joy, fear, and sadness. These emotions encompass a wide range of reactions and discussions surrounding hurricanes. To gain deeper insights into the main topics associated with each emotion, we have generated a word cloud, depicted in Fig. 6, which visually represents the most frequently occurring words about the four emotions considered in our study. To get a coherent analysis of emotions, we grouped them into Negative and Positive emotions. The negative emotions, which include anger, fear, and sadness, accounted for a total percentage of 70.92%, while the positive emotions were represented by joy (28.30%), which is consistent with our expectations.

3.5.1 Negative emotions

The dominant emotion observed in our dataset was anger, with a percentage of 39.91%. This finding reflects the discontent and frustration expressed by individuals during the occurrence of the hurricanes. The word cloud associated with anger reveals the presence of mandatory evacuation, emergency, power outage, and infrastructure protection, indicating dissatisfaction with healthcare services, disaster management preparedness, and infrastructure. These observations indicate that a significant number of users expressed their anger towards these aspects of the hurricanes. The fear detected in our dataset exhibited similar concerns to those associated with anger, as evidenced by the shared terms in the word cloud. These terms revolve around disaster management and emergency operations, reflecting the unease experienced by individuals during the hurricanes. Sadness reveals the level of intensity of the disasters, emphasizing the unfortunate loss of lives and the extent of devastation caused. Words

https://doi.org/10.55343/CogSust.80

such as death, devastation, and damage are prevalent in the Sadness word cloud, highlighting the profound impact of the hurricanes on affected communities.

3.5.2 Positive Emotions

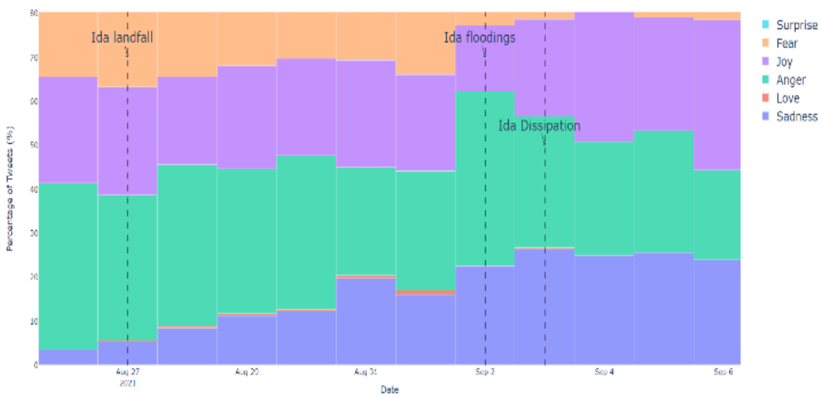
The Joy emotion accounted for 28.30% of the emotions detected. It can be attributed to relief efforts, assistance, and individual resilience. Words associated with joy in the word cloud include relief, help, life, and family, indicating a glimmer of positivity towards the challenges faced during the hurricanes.

Fig. 6. Word cloud for Anger, Joy, Sadness, and Fear (dominant) emotions

We also employed a temporal analysis approach to investigate the normalized emotion distribution over time presented in Fig. 7 to examine the evolving emotions detected during hurricane events. Among the hurricanes studied, Hurricane Ida emerged as the most representative case for our analysis. During the progression of Hurricane Ida, we observed distinct patterns in the emotions detected. Prior to the landfall of the hurricane, there was a noticeable increase in emotions such as fear and anger. These emotions peaked one day before and during the hurricane's landfall and the day of flooding caused by the disaster. This aligns with our expectations, as the heightened fear and anger can be attributed to the anticipation and direct impact of the hurricane, as well as the uncertainties and potential threats associated with the flooding. This finding suggests a prevailing sense of dissatisfaction and insecurity.

Furthermore, the emotion of sadness showed a continuous upward trend throughout the hurricane event. This observation corresponds to the increasing intensity of the disaster, reflecting the accumulation of devastating impacts and the loss of lives. The rising trend in sadness emotions can potentially indicate individuals directly affected by the hurricane and experiencing profound emotional distress. It seems the patterns of anger and fear are significantly different. Fear practically ceases when the worst happens, but anger is maintained even after that, though diminishingly.

On a contrasting note, the emotion of joy increased, particularly after the dissipation of the hurricane. This highlights the presence of individuals who have been relieved from the immediate impacts of the hurricane or have been successfully rescued and saved. The growing joy can be attributed to relief and gratitude among those spared from the hurricane’s direct consequences.



https://doi.org/10.55343/CogSust.80

Fig. 7 Emotion distribution during Hurricane Ida

3.6 Topic Modeling

Through this analysis, we identified three main topics from the dataset provided in Table 4. that highlighted the most people's concerns during hurricane occurrence. Analyzing the generated topics shows that the public's response to hurricanes primarily revolved around hurricanes and healthcare management. The first topic highlighted the public's concerns about evacuation efforts, disaster response, and infrastructure preparedness to cope with such disasters effectively. Individuals seek viable strategies for ensuring their safety and minimizing potential damage. Another prominently discussed topic revolved around support and donations. Within this theme, individuals proposed and sought assistance through various means, including collecting 'help', 'fundraising' initiatives, and 'donations'.

Furthermore, hurricane intensity and forecasting emerged as a significant area of interest for the public. People actively sought information regarding hurricane evolution, intensity levels, and ways to disseminate this information to others. Keywords such as 'shortage' and 'death' emphasized the seriousness of the situation, indicating the public's awareness of the human and resource implications caused by hurricanes.

Table 4. Most discussed topics

Topic groups	***Related keywords***
Hurricane/healthcare management	'kit','Irma','operation','monitoring', 'management','center','prepared', 'evacuation','mandatory','underprepared', 'preparedness','response','hospital', 'patient', 'infrastructure
Support and donation	'donation','fundraiser','proceeds','donating', 'donate','collecting', 'fundraising', 'victim', 'help', 'affected', 'helping'
Hurricane intensity and forecasting	'toll', 'death', 'rise', 'serious', 'shortage', 'category','strengthens','declares', 'intensifies''become','rapidly','expected', 'intensify'

Source: Own compilation

4. Discussion

This study aims to examine and understand the public's response during extreme weather events by exploring cognitive sustainability concerns. The study approach was based on collecting Twitter data using healthcare/hurricane-related keywords for each of the five study cases and applying NLP models for sentiment, emotion, and topic analysis. A first noteworthy observation is the prevailing culture of social media engagement, even during disasters and at all hours of the day, which was emphasized by the significant number of tweets collected during the night. The word cloud and frequency analysis findings unveiled a notable presence of words related to disaster response, infrastructure, and help-seeking, where people were deeply concerned about their safety and the need to relocate to safer areas when faced with the imminent threat of extreme weather. This response reflects the urgency and importance of evacuation measures in protecting lives and mitigating potential damages. Emotional aspects, when talking about home and family, highlight the impact of EWE on the personal and mental health of individuals and communities. People's concerns centred on protecting their lives, homes, and loved ones, underlining the sense of vulnerability and the need for effective disaster preparedness plans.

Another aspect worth noting is the discussions related to providing and seeking help, which is compelling evidence of a robust information-spreading culture. It showcases the public's heightened awareness and proactive approach towards disaster preparedness, reflecting how individuals use communication channels to share critical information, warn others, and offer support during extreme weather events. These findings were supported by the outcomes of the topic modelling analysis that provided a concise and informative summary of what we classified as the main three concerns expressed by the public during EWE. The identified concerns and active engagement observed in the public response to hurricanes are deeply connected to the cognitive sustainability concept. The first prominent concern centred around emphasizing hurricane management, reflecting the need for extended sustainability where decision-makers actively manage various aspects and dimensions of the disaster. This involves responding to the immediate impact and strategically planning for infrastructure preparedness and evacuation, ensuring a comprehensive and sustainable disaster response. The second topic, revolving around donation and support, resonates with the principle of inter-cognitive sustainability. The community actively engages in collective efforts, such as fundraising and donations, showcasing a collaborative response that transcends individual cognition to support the broader affected community. The third topic, focusing on hurricane intensity and forecasting, emphasizes a shared understanding and concern for the evolving situation, with information dissemination playing a crucial role. The temporal analysis, which represented a pattern from the initial stages of the hurricane until its dissipation, further highlights this cognitive sustainability aspect, showcasing the

https://doi.org/10.55343/CogSust.80

community's active participation in sharing real-time updates and urgent messages and experiences with social media being a way to stay informed and connected, promoting a collective and informed response.

As the hurricane intensifies, the frequency of tweets increases as people express their sentiments, emotions, and concerns about the unfolding situation. In this context, sentiment analysis results indicated a higher prevalence of negative sentiment than positive sentiment. Neutral sentiments were not detected, which may be caused by imbalanced training data and potential limitations in the training process of the pre-trained model we applied. With the dominance of the negative sentiment, it becomes evident that the overall public opinion, as reflected in the word cloud, is negative. Further analysis revealed that anger was the predominant negative sentiment extracted from emotion detection, indicating a significant dissatisfaction among citizens with disaster response and emergency efforts. This finding was corroborated by the word cloud associated with anger, highlighting the presence of these topics.

Additionally, the continuous increase in the emotions of fear and sadness, as shown in the normalized emotion distribution for Hurricane Ida, suggests a failure in effectively managing the hurricane disaster. This is particularly evident in the persistence of sadness, even after the dissipation of Hurricane Ida, reflecting the long-lasting impacts generally associated with the physical, mental, and financial difficulties faced by those directly or indirectly affected by the disaster. These findings align with our expectations for a topic focused on disasters and emergencies, with the expression of love being understandably diminished during such events. Additionally, the absence of surprise can be attributed to the availability of advanced forecasting systems that inform individuals of the hurricane's occurrence in advance. By examining the temporal aspects of our sentiment and emotion detection findings, we can investigate how fast people's responses evolve during extreme weather events. While negative sentiment was consistently prevalent, the gap between negative and positive sentiments, as shown in the daily sentiment distribution for hurricanes Irma and Harvey, rapidly decreased during hurricane dissipation and in the following days. This finding aligns with the emotion distribution of Hurricane Ida, where the emotion of joy increased the day after the hurricane's dissipation, likely attributed to relief efforts, as indicated by the word cloud associated with the joy emotion. This suggests a positively fast change in people's emotions that effective disaster response and emergency interventions can explain. However, the parallel increase in the emotion of sadness in the same graph raises questions about which populations received more assistance during the hurricanes. This leads us to one of the study's limitations: our dataset's absence of geographical indications.

This study was the first attempt to examine the public response during disasters. However, it is important to acknowledge several limitations that should be considered for future work. First, the dataset was imbalanced for each hurricane. This discrepancy in the number of tweets collected makes it challenging to compare the study case events directly. Additionally, the absence of geolocation information associated with the tweets limited our ability to capture the areas most affected by hurricanes. Moreover, good evaluation parameters for pre-trained models do not necessarily mean better performance or accurate results. This was evident in the case of the RoBERTa model, where the absence of neutral tweets in our dataset affected the model's ability to classify news-related content accurately.

We propose training an NLP model on a well-labelled and balanced dataset for future work. Additionally, we suggest the use of geo-located tweets in order to capture disparities. Also, inspired by the topic modelling results that revealed evacuation and disaster response as predominant concerns, we can work on assessing emergency management and healthcare system vulnerability during EWE, which will facilitate the development of targeted interventions to strengthen disaster response capabilities for cognitive sustainability and propose alternatives with a focus on exploring the role that E-health and telemedicine as innovative solutions.

5. Conclusion

In conclusion, this study explored the public's response during EWE by analyzing Twitter data and highlighting the concept of cognitive sustainability. The analysis revealed three key concerns: hurricane management, donation and support, and disaster impacts. We provided insight into how people react to these topics by applying NLP techniques for sentiment and emotional analysis. However, it is essential to acknowledge the limitations of this study. The imbalanced data collection and lack of geographical indications in the dataset may have influenced the representation of certain areas and topics, potentially impacting the overall findings.

For future work, it is recommended to address the limitations by ensuring a more balanced data collection and incorporating geographical parameters. Moreover, further research could focus on refining sentiment and emotion analysis by training NLP models rather than relying on pre-trained ones to capture the nuances of public responses more accurately. Exploring other social media platforms and languages could provide a more comprehensive understanding of the public's response during EWE. Finally, this study lays the groundwork for better understanding the public's concerns and responses during extreme weather events, offering valuable insights for disaster management and emergency preparedness efforts.

https://doi.org/10.55343/CogSust.80

References

Abel, F., Hauff, C., Houben, G. J., Stronkman, R., & Tao, K. (2012, June). Semantics+ filtering+ search= twitcident. exploring information in social web streams. In *Proceedings of the 23rd ACM conference on Hypertext and social media* (pp. 285-294). DOI: https://doi.org/gftcp3

Acheampong, F. A., Nunoo-Mensah, H., & Chen, W. (2021). Transformer models for text-based emotion detection: a review of BERT-based approaches. *Artificial Intelligence Review*, 1-41. DOI: https://doi.org/gmxnr9

Albahli, S., Algsham, A., Aeraj, S., Alsaeed, M., Alrashed, M., Rauf, H. T., ... & Mohammed, M. A. (2021). COVID-19 Public Sentiment Insights: A Text Mining Approach to the Gulf Countries. *Computers, Materials & Continua*, 67(2). DOI: https://doi.org/k9bj

Al-Garadi, M. A., Yang, Y. C., & Sarker, A. (2022, November). The Role of Natural Language Processing during the COVID-19 Pandemic: Health Applications, Opportunities, and Challenges. In *Healthcare* (Vol. 10, No. 11, p. 2270). MDPI. DOI: https://doi.org/k9bh

Barbieri, F., Camacho-Collados, J., Neves, L., & Espinosa-Anke, L. (2020). Tweeteval: Unified benchmark and comparative evaluation for tweet classification. *arXiv preprint arXiv*:2010.12421. DOI: https://doi.org/k9bp

Catelli, R., Pelosi, S., & Esposito, M. (2022). Lexicon-based vs. Bert-based sentiment analysis: A comparative study in Italian. *Electronics*, 11(3), 374. DOI: https://doi.org/k9bn

Dash, N., Morrow, B. H., Mainster, J., & Cunningham, L. (2007). Lasting effects of Hurricane Andrew on a working-class community. *Natural Hazards Review*, 8(1), 13-21. DOI: https://doi.org/fqpbsn

Ebi, K. L., & Bowen, K. (2016). Extreme events as sources of health vulnerability: Drought as an example. *Weather and climate extremes*, 11, 95-102. DOI: https://doi.org/gg3r27

Ebi, K. L., Vanos, J., Baldwin, J. W., Bell, J. E., Hondula, D. M., Errett, N. A., ... & Berry, P. (2021). Extreme weather and climate change: population health and health system implications. *Annual review of public health*, 42(1), 293-315. DOI: https://doi.org/gnhm97

Fischer, E. M., & Knutti, R. (2015). Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. *Nature climate change*, 5(6), 560-564. DOI: https://doi.org/35h

Karami, A., Shah, V., Vaezi, R., & Bansal, A. (2020). Twitter speaks: A case of national disaster situational awareness. *Journal of Information Science*, 46(3), 313-324. DOI: https://doi.org/ggwkc4

Kwasinski, A., Andrade, F., Castro-Sitiriche, M. J., & O'Neill-Carrillo, E. (2019). Hurricane Maria effects on Puerto Rico electric power infrastructure. *IEEE Power and Energy Technology Systems Journal*, 6(1), 85-94. DOI: https://doi.org/gp497r

Lu, Y., & Yang, D. (2011). Information exchange in virtual communities under extreme disaster conditions. *Decision Support Systems*, 50(2), 529-538. DOI: https://doi.org/fnnh27

Mills, M. A., Edmondson, D., & Park, C. L. (2007). Trauma and stress response among Hurricane Katrina evacuees. *American journal of public health*, 97(Supplement_1), S116-S123. DOI: https://doi.org/brsvmf

Mitsova, D., Escaleras, M., Sapat, A., Esnard, A. M., & Lamadrid, A. J. (2019). The effects of infrastructure service disruptions and socio-economic vulnerability on hurricane recovery. *Sustainability*, 11(2), 516. DOI: https://doi.org/10.3390/su11020516

Neppalli, V. K., Caragea, C., Squicciarini, A., Tapia, A., & Stehle, S. (2017). Sentiment analysis during Hurricane Sandy in emergency response. *International journal of disaster risk reduction*, 21, 213-222. DOI: https://doi.org/gbj94q

Nguyen, Q. T., Nguyen, T. L., Luong, N. H., & Ngo, Q. H. (2020, November). Fine-tuning bert for sentiment analysis of vietnamese reviews. In 2020 *7th NAFOSTED conference on information and computer science* (NICS) (pp. 302-307). IEEE. DOI: https://doi.org/k9bm

NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters (2023). https://www.ncei.noaa.gov/access/billions/, DOI: https://doi.org/gmhcnv

Oluyomi, A. O., Panthagani, K., Sotelo, J., Gu, X., Armstrong, G., Luo, D. N., ... & Bondy, M. (2021). Houston hurricane Harvey health (Houston-3H) study: assessment of allergic symptoms and stress after hurricane Harvey flooding. *Environmental Health*, 20(1), 1-15. DOI: https://doi.org/gjxrb8

Shultz, J. M., Trapido, E. J., Kossin, J. P., Fugate, C., Nogueira, L., Apro, A., ... & Galea, S. (2022). Hurricane Ida's impact on Louisiana and Mississippi during the COVID-19 delta surge: Complex and compounding threats to Population Health. *The Lancet Regional Health–Americas*, 12. DOI: https://doi.org/k9br

Stephenson, D. B., Diaz, H. F., & Murnane, R. J. (2008). Definition, diagnosis, and origin of extreme weather and climate events. *Climate extremes and society*, 340, 11-23. URL: https://empslocal.ex.ac.uk/people/staff/dbs202/publications/2008/extremes.pdf

Sun, S., Kim, J. H., Jung, H. S., Kim, M., Zhao, X., & Kamphuis, P. (2023). Exploring Hype in Metaverse: Topic Modeling Analysis of Korean Twitter User Data. *Systems*, 11(3), 164. DOI: https://doi.org/k9bq

Sutton, J., & Tierney, K. (2006). Disaster preparedness: Concepts, guidance, and research. Colorado: *University of Colorado*, 3(1). URL: https://www.bencana-kesehatan.net/arsip/images/referensi/april/Disaster%20Preparedness%20Concepts_Jurnal.pdf

Vieweg, S., Hughes, A. L., Starbird, K., & Palen, L. (2010, April). Microblogging during two natural hazards events: what twitter may contribute to situational awareness. *In Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1079-1088). DOI: https://doi.org/fjck8p

Waddell, S. L., Jayaweera, D. T., Mirsaeidi, M., Beier, J. C., & Kumar, N. (2021). Perspectives on the health effects of hurricanes: a review and challenges. *International journal of environmental research and public health*, 18(5), 2756. DOI: https://doi.org/gr934k

https://doi.org/10.55343/CogSust.80

Yuan, F., Li, M., & Liu, R. (2020). Understanding the evolutions of public responses using social media: Hurricane Matthew case study. *International journal of disaster risk reduction*, 51, 101798. DOI: https://doi.org/k9bg

Zhou, S., Kan, P., Huang, Q., & Silbernagel, J. (2023). A guided latent Dirichlet allocation approach to investigate real-time latent topics of Twitter data during Hurricane Laura. *Journal of Information Science*, 49(2), 465-479. DOI: https://doi.org/k9bk

Zoldy, M., Csete, M. S., Kolozsi, P. P., Bordas, P., & Torok, A. (2022). Cognitive sustainability. *Cognitive Sustainability*, 1(1). DOI: https://doi.org/htfq