JOURNAL OF HYDRO - METEOROLOGY

ISSN 2525 - 2208

VIETNAM METEOROLOGICAL AND HYDROLOGICAL ADMINISTRATION

No 17
12-2023

Publishing licence
No: 166/GP-BTTTT - Ministry of Information and Communication dated 17/04/2018

Editorial office
No 8 Phao Dai Lang, Dong Da, Ha Noi
Tel: 024.39364963
Email: tapchikttv@gmail.com

Engraving and printing
Vietnam Agriculture Investment Company Limited
Tel: 0243.5624399

JOURNAL OF HYDRO-METEOROLOGY
Volume 17 - 12/2023

TABLE OF CONTENT

JOURNAL OF
HYDRO-METEOROLOGY

Rerearch Article

Efficient oil removal from water using carbonized rambutan peel: Isotherm and kinetic studies

Trinh Trong Nguyen[1,2], Nguyen Dinh Loc[1], Le Huy Ba[2], Thai Van Nam[1]*

[1] HUTECH Institute of Applied Sciences, HUTECH University, 475A Dien Bien Phu Street, Ward 25, Binh Thanh District, Ho Chi Minh City 700000, Vietnam; tt.nguyen@hutech.edu.vn; 6009220001@hufi.edu.vn; lochenni@gmail.com; tv.nam@hutech.edu.vn

[2] Ho Chi Minh City University of Industry and Trade, 140 Le Trong Tan Street, Tay Thanh Ward, Tan Phu District, Ho Chi Minh City 700000, Vietnam; tt.nguyen@hutech.edu.vn; 6009220001@hufi.edu.vn; lhuyba@gmail.com

*Corresponding author: tv.nam@hutech.edu.vn; Tel: +84–945007990

Received: 08 September 2023; Accepted: 02 October 2023; Published; 25 December 2023

Abstract: The study aims to assess the diesel oil adsorption capacity of activated carbon derived from rambutan peel, a prevalent agricultural by-product in Vietnam. The adsorption process will be investigated using various adsorption isotherms (Langmuir, Freundlich, Temkin, Dubin-Radushkevich), adsorption kinetics (pseudo-first order and pseudo-second order). The characterization analysis of RPAC revealed the following values: BET surface area of 786.0143 m^2/g, BJH adsorption cumulative volume of pores at 0.05392 cm^3/g, and BJH adsorption average pore diameter of 55.2432 nm. Individual assessments of factors influencing oil removal efficiency and adsorption capacity identified the optimal initial oil concentration at 1% volume/volume, an optimum contact time of 60 minutes, and an equilibrium adsorption time of 80 minutes. The optimal adsorbent dosage was 250 mg, and the ideal pH was 6. The adsorption process adhered to the Langmuir isotherm with an R^2 value of 0.9993, yielding a maximum adsorption capacity of 5,712.0 mg/g. The pseudo-second-order model provided a better fit for RPAC's oil adsorption process with an R^2 value of 0.9969.

Keywords: Activated carbon; Diesel oil; Materials science; Modeling; Rambutan peel.

1. Introduction

The total volume of oil released into the environment from oil tanker spills in 2021 was approximately 10,000 tons [1]. Oil-polluted wastewater affects water quality and underwater ecosystems worldwide [2], posing severe health consequences [3].

Methods for oil removal in water can be categorized into in-situ burning, chemical methods (solidification, dispersion), biological methods, and physical methods (skimming, adsorbents) [4]. Among these, oil adsorbents remain the preferred technique for oil clean-up due to their speed, simplicity, environmental friendliness [5], and cost-effectiveness. The choice of adsorbent material depends on factors such as availability, cost, and safety considerations [6]. Currently, about 200 different types of adsorbent materials are being produced and used [6], classified into three main groups: inorganic minerals, synthetic organic products, and natural organic products [7]. Among these, adsorbent materials of natural organic origin offer significant advantages compared to others, especially in terms of environmental friendliness in marine environments and their lightweight nature, making them easily recoverable and reusable [8].

Activated carbon, derived from carbonization, is used for adsorbing various compounds [9–10]. Activated carbon from plant biomass is particularly favored due to its abundant and readily available source, resulting in significantly lower production costs compared to commercial activated carbon [11]. Consequently, many different activated carbon materials have been synthesized from various plant-based sources, such as coconut shells [11], coconut coir [12], safou seeds [13], corn cobs [10, 14], rice husks [15, 16], and barley straw [17].

According to the General Statistics Office data, Vietnam's total agricultural by-products in 2020 amounted to over 156.8 million tons, with 88.9 million tons of post-harvest by-products from crops and agricultural processing (accounting for 56.7%) [18]. The current challenge is effectively utilizing and recycling agricultural residues to support farmers' livelihoods while minimizing environmental impact. Among these residues, those with high lignocellulose content present an important area of research for selecting suitable agricultural by-products to fabricate efficient and cost-effective oil-adsorbent materials. Rambutan belongs to the 10 main fruit-bearing trees group in Vietnam [19]. Rambutan peel (RP) is an agricultural waste with a high cellulose content, constituting approximately 24.28% [20]. Cellulose is a crucial structural component of adsorbent materials [21], capable of enhancing the oil recovery efficiency of these materials in aqueous environments [22]. Consequently, RP is considered a cost-effective and suitable adsorbent material for oil removal in water. Numerous studies have demonstrated the potential of using activated carbon derived from rambutan peel for the removal of Fe(II) from aqueous solutions [23] and the removal of Remazol Brilliant Blue R [24].

This study aimed to investigate the individual effects of factors influencing RP's oil removal efficiency and adsorption capacity, followed by conducting isotherm and kinetic studies on the adsorption process.

2. Material and methods

The research diagram is presented in Figure 1.

2.1. Materials

Rambutan peel was sourced from the main market in HCMC, Vietnam, and the rambutan variety used is Dona rambutan. The diesel oil (DO) 0.05S was obtained from Petrolimex Aviation in HCM, Vietnam, with a specific gravity at 15°C of 820-860 kg/m^3. Potassium hydroxide, n-Hexane, and sodium sulfate anhydrous were sourced from Xilong Scientific Co., Ltd. in Shenzhen, China. Sulfuric acid (c(H_2SO_4) = 0.05 mol/l or 0.1 N) for a 1000 ml solution and hydrochloric acid (c(HCl) = 0.1 mol/l or 0.1 N) were provided by Merck in Germany.

2.2. Preparation of adsorbent and simulated oil spill

Rambutan peel underwent rinsing with tap water, subsequent drying at 105°C for 24 hours to eliminate moisture, grinding, and sieving to achieve a particle size of 1-2 mm. The resulting dried RP was introduced into a stainless steel vertical tubular reactor positioned within a tube furnace. Carbonization was executed at 550°C for a duration of 2 hours in an environment of purified nitrogen (99.99%). The resulting product, rambutan peel char (RPC), was subsequently combined with KOH pellets at a 1:2 impregnation ratio [24]. Deionized water was added to dissolve the entire KOH pellets, resulting in a KOH concentration of 1M, and the mixture was soaked for 24 hours. The mixture was dried in an oven at 105°C for 24 hours to remove moisture and then subjected to activation in a muffle furnace at 550°C for 1 hour. The resulting rambutan peel activated carbon (RPAC) was subsequently cooled, treated with a 0.1 M HCl solution to eliminate ash content, and rinsed with distilled water until the pH of the washing solution stabilized at 6-7. The prepared RPAC was dried in an oven at 105°C for 2 hours, crushed, and sieved into various particle sizes before being stored in a

desiccator for use in the next experiment. The simulated oil-water mixture was prepared by mixing 2.5 mL of DO with 250 mL water in a beacher of 250 ml of brine water (1.0% volume/volume).

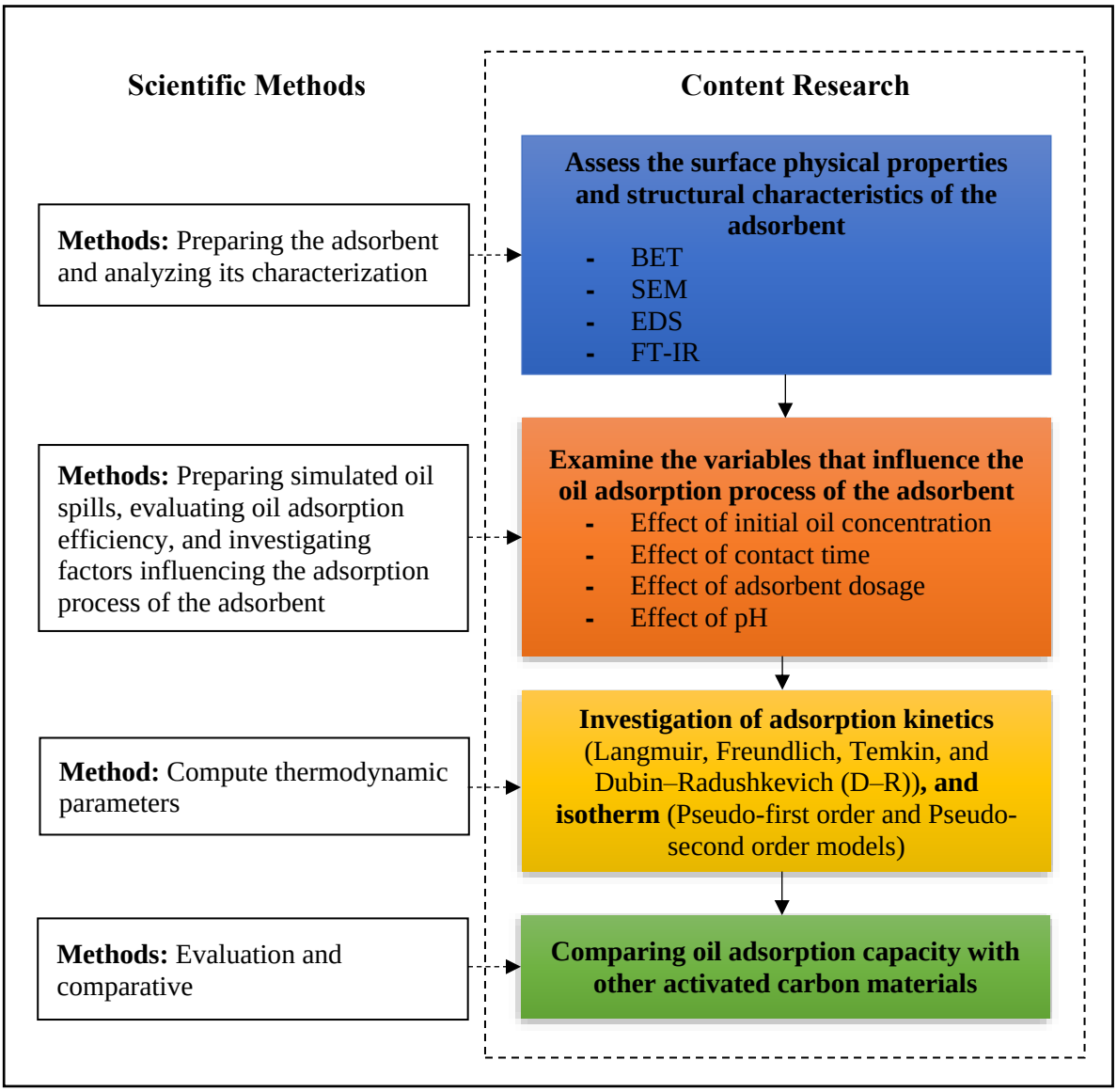


Figure 1. Research diagram.

2.3. Characterization of adsorbent

The surface area, pore volume, and average pore diameter of the RPAC prepared under optimal conditions were assessed using the Micromeritics® TriStar II Plus Version 3.03. Scanning Electron Microscope (SEM) and Energy Dispersive X-ray Spectroscopy (EDS) analyses were performed with the JSM-IT500 InTouchScope™ Scanning Electron Microscope. Fourier-transform infrared spectroscopy (FT-IR) measurements were conducted using the Agilent Cary 630 FTIR spectrometer.

2.4. Method to evaluate the efficiency of oil adsorption

For each experiment, the amount of adsorbed oil per gram of adsorbent at equilibrium is denoted as q_e (mg/g), at time t is q_t (mg/g), and the adsorption efficiency is determined using the following formula [24]:

$$q_e = \frac{(C_o - C_e) * V}{M} \quad (1)$$

$$q_t = \frac{(C_o - C_t) * V}{M} \quad (2)$$

$$\%\, Effective\ removal = \frac{(C_o - C_e) * 100}{C_o} \quad (3)$$

where C_o is the initial oil concentration (mg/L), C_e is the equilibrium concentration (mg/L), C_t is the oil concentration at time t (mg/L), V is the volume of the solution (ml), and M is the mass of the adsorbent used (mg).

2.5. Investigating the factors affecting the oil adsorption process of the adsorbent

The factors affecting the oil adsorption process of the adsorbent will be sequentially investigated from Experiments 1 to 4. The optimized values obtained from each experiment will be utilized for the subsequent experiment. The experimental layout is presented in Table 1.

Table 1. Experiment to investigate the factors affecting the oil adsorption process.

Experiment	Initial oil concentration (% v/v)	Contact time (min)	Adsorbent dosage (mg)	pH	Adsorbent particle size (mm)	Temperature (°C)
Experiment 1: Investigating the effect of initial oil concentration	Range 0.25 - 2 (jump interval 0.25)	60	250	7 ± 0.2	0.6 ÷ 1.0	25 ± 1
Experiment 2: Investigating the impact of contact time	Optimization of Experiment 1	Range 10 - 90 (jump interval 10)	250	7 ± 0.2	0.6 ÷ 1.0	25 ± 1
Experiment 3: Investigating the influence of adsorbent dosage	Optimization of Experiment 1	Optimization of Experiment 2	Range 250 - 1,500 (jump interval 250)	7 ± 0.2	0.6 ÷ 1.0	25 ± 1
Experiment 4: Investigating the effect of pH	Optimization of Experiment 1	Optimization of Experiment 2	Optimization of Experiment 3	Range 3 - 11 (jump interval 1)	0.6 ÷ 1.0	25 ± 1

2.6. Sorption Isotherms and Kinetics

In this study, four commonly used sorption isotherm models, namely Langmuir (4), Freundlich (5), Temkin (6), and Dubin–Radushkevich (D–R) (7) equations, were used to fit the experimental data [25]:

$$\frac{C_e}{q_e} = \frac{C_e}{q_{max}} + \frac{1}{q_{max} K_L} \quad (4)$$

$$\ln q_e = \ln K_f + \frac{1}{n} \ln C_e \quad (5)$$

$$q_e = B_T \ln K_T + B_T \ln C_e \quad (6)$$

$$\ln q_e = \ln q_{D-R} - K_{ad} \varepsilon^2 \quad (7)$$

where C_e is the equilibrium concentration of adsorbate (mg/L), q_e is the amount of oil adsorbed per gram of the adsorbent at equilibrium (mg/g), q_{max} is the maximum monolayer adsorption capacity (mg/g), K_L is Langmuir isotherm constant (L/mg), K_f is Freundlich isotherm constant (mg/g), n is adsorption intensity, B_T is constant related to heat of sorption(J/mol), K_T is Temkin isotherm equilibrium binding constant (L/g), q_{D-R} is theoretical

isotherm saturation capacity (mg/g), K_{ad} is Dubinin-Radushkevich isotherm constant (mol^2/kJ^2), and ε is Polanyi potential.

First-order and pseudo-first-order kinetic model [26] and Second-order and pseudo-second-order kinetic model [27] are presented in equations (8) and (9).

$$log(q_e - q_t) = log(q_e) - \frac{k_1}{2.303}t \quad (8)$$

$$\frac{t}{q_t} = \frac{1}{k_2 q_e^2} + \frac{1}{q_e}t \quad (9)$$

where q_t and q_e are the amount of oil sorbed per mass of sorbent (mg/g) at time t (min) and equilibrium, respectively. k_1 (1/min) and k_2 (g/mg.min) are the constant rate parameters of the pseudo-first-order and the pseudo-second-order sorption, and h (mg/g.min) is the initial sorption rate.

3. Results and discussion

3.1. Characteristics of the adsorbent

3.1.1. Surface physical characteristics of the adsorbent

Surface physical characteristics of RPC and RPAC were determined using Brunauer-Emmett-Teller (BET) analysis and are presented in Table 2.

Table 2. Surface physical characteristics of RPC and RPAC.

Surface physical characteristics	RPC	RPAC
BET surface area (m^2/g)	485.7680	786.0143
BJH adsorption cumulative surface area of pores (m^2/g)	0.9787	10.3904
BJH adsorption cumulative volume of pores (cm^3/g)	0.005293	0.05392
BJH adsorption average pore diameter (nm)	21.6336	55.2432

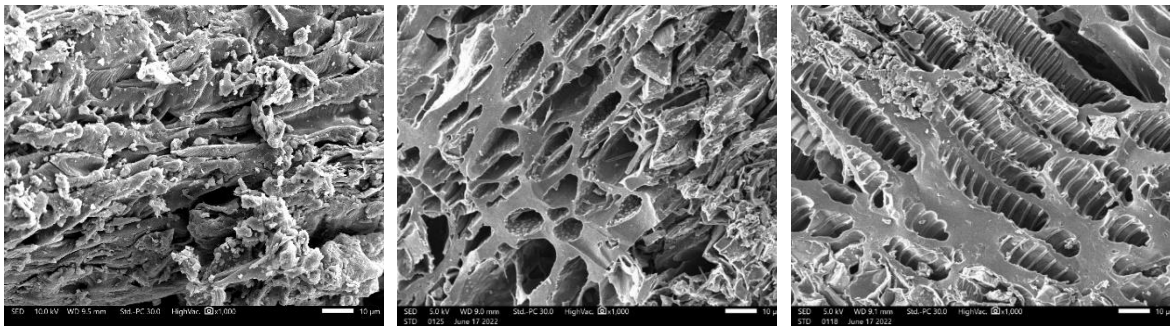
The adsorbents' pore properties and surface area were determined using BET isotherm analysis. The BET surface area of RPC and RPAC were 485.7680 (m^2/g) and 786.0143 (m^2/g), respectively. The increased BET surface area in RPAC indicated more active sites, enhancing oil uptake [28]. Compared to other naturally derived organic materials that have undergone carbonization, such as coconut coir (with a surface area of 691.800 m^2/g) [12] and plantain peels fiber (with a surface area of 234.51 m^2/g) [28], the surface area of RPAC is higher.

The pore volume increased from 0.005293 to 0.05392 (cm^3/g), while the pore diameter increased from 21.6336 to 55.2432 nm. The average pore diameter of RPAC (55.2432 nm) is higher than the average pore size of activated carbon from corncobs (45 nm) [10], rambutan peel-based activated carbon (2.63 nm) [24], and carbonized plantain peels fiber (4.77 nm) [28].

From these findings, it is evident that RPAC exhibits larger surface area, total pore volume, and pore diameter in comparison to RPC. This creates advantageous circumstances for the adsorption process to occur on the adsorbent's surface, consequently augmenting RPAC's oil retention and adsorption capacity. The enhancement in BET surface area can likewise be ascribed to the removal of volatile organics during thermal treatment [29]. The enhanced pore characteristics will facilitate the diffusion of a greater number of oil molecules onto the RPAC surface [28, 30].

3.1.2. Surface and structural characteristics of the adsorbent

The SEM images depicting the surface and structural characteristics of raw RP, RPC, and RPAC at ×1,000 magnification are presented in Figure 2.



a) b) c)

Figure 2. SEM images (×1,000) of raw RP (a), RPC (b), and RPAC (c).

Based on the SEM images at ×1,000 magnification, the surface structure of raw RP displays uniform fiber-like crystal bonds and small pores, indicating the presence of minerals (Figure 2a). The surface mainly comprises a lignocellulose network and fiber matrix containing lignin, cellulose, volatile organic compounds, and hemicellulose. After carbonization, RPC's surface exhibits uneven, rough, and rugged structures with numerous existing pores (Figure 2b). Alkaline treatment results in larger, deeper pores, increasing their number on the material's surface (Figure 2c). The KOH-C reaction enhances pore development during activation, thereby increasing surface area and adsorption capacity. A nearly heterogeneous pore structure pattern is also observed on the RPAC surface.

The immersion of activated carbon in KOH solution enhances the adsorption activity of the adsorbent by creating potassium carbonate crystals on the carbon surface. These crystals generate airflow holes on the surface of adsorbents, boosting adsorption capacity and increasing the existing surface area for adsorbing organic compounds, including oil from water, during adsorption. Furthermore, the immersion process also forms active functional groups on the carbon surface, playing a vital role in adsorbing organic substances.

The observed formation of pores on the surface of RPAC material aligns with findings by [24] regarding rambutan peel activation and with the study by [12] on activated carbon production from coconut coir.

3.1.3. FT-IR analysis of the adsorbent

The analysis of raw RP and RPAC samples using FT-IR spectra revealed the presence of various peaks, indicative of distinct functional groups, across the wavelength range of 650 - 4000 cm^{-1}, as depicted in Figures 3 and 4.

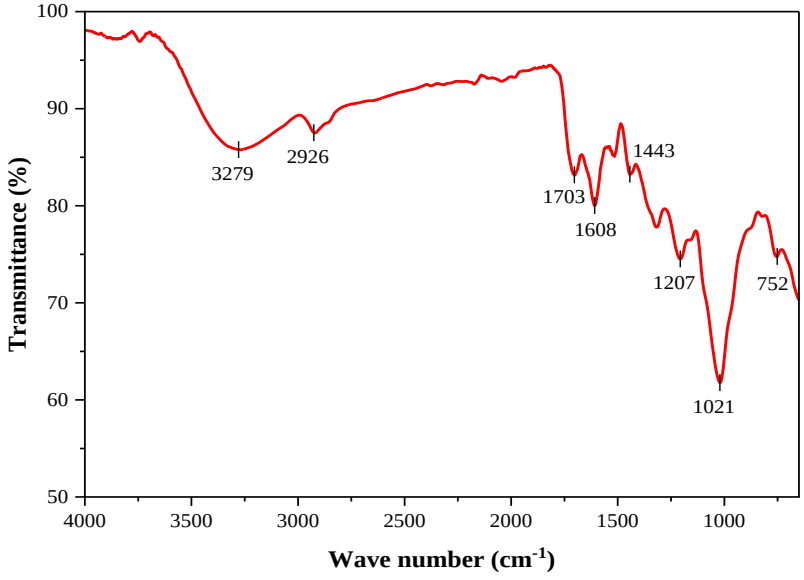


Figure 3. FT-IR of raw RP.

The FT-IR spectrum of rambutan peel reveals distinct functional groups. A peak at 3279 cm^{-1} signifies the presence of (O-H) groups, originating from alcohol and carboxylic acid moieties. Furthermore, an aliphatic C-H group is evident at 2926 cm^{-1}. Notably, a non-ionic carbonyl group (C=O) is identified at 1703 cm^{-1}. Other significant peaks include 1608 cm^{-1}, representing a -COO-carboxylate group, 1443 cm^{-1}, signifying a symmetrical carboxylate group, and 1207 cm^{-1}, indicating a C-O group derived from ether (resulting from lignin bonding with cellulose). These findings align with previous studies [31] regarding the functional group characteristics of rambutan peel. Additionally, the band at 1021 cm^{-1} is attributed to C-OH [32], while at 752 cm^{-1}, the presence of Methylene—$(CH_2)_n$—rocking ($n \geq 3$) is observed [33]. This comprehensive characterization enhances our understanding of the chemical composition of rambutan peel.

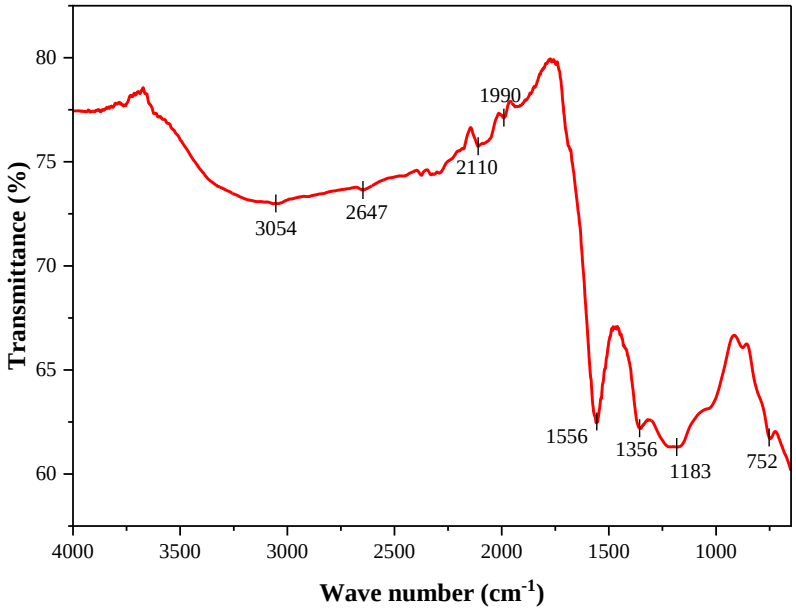


Figure 4. FT-IR of RPAC.

For RPAC, at 3054 cm^{-1}, we observe a distinct C-H stretching indicative of an alkene group. The presence of a carboxylic acid group is evident at 2647 cm^{-1}, characterized by its -OH stretching. Additionally, the spectrum features a sharp peak at 2110 cm^{-1}, corresponding to the C≡C stretching of an alkyne group [12]. At 1990 cm^{-1}, we discern a C-H bending pattern, which is characteristic of an aromatic compound. Furthermore, the spectrum highlights several important peaks: at 1556 cm^{-1}, there is a pronounced C=C stretching, characteristic of an aromatic benzene ring group [12]. At 1356 cm^{-1}, we observe an O-H bending pattern, indicative of phenol. Lastly, at 1183 cm^{-1}, there is a distinct C-O stretching signal associated with a tertiary alcohol group. In addition to these functional group vibrations, the presence of methylene groups is clearly detected at 752 cm^{-1}, where the Methylene —$(CH_2)_n$—rocking ($n \geq 3$) motion is observed [33].

Both raw RP and RPAC exhibit the presence of various functional groups, as depicted by distinct FT-IR spectral peaks at different wavenumbers. These functional groups include C-H, -OH, -COO, and Methylene—$(CH_2)_n$—rocking ($n \geq 3$), and they appear in both samples, albeit at different positions and intensities. However, significant differences also exist, such as the presence of a non-ionic carbonyl group (C=O) and a -COO-carboxylate group in raw RP but not in RPAC. Conversely, RPAC displays features like C≡C stretching of an alkyne group and C=C stretching of an aromatic benzene ring group that are absent in raw RP. Moreover, each sample's FT-IR spectrum exhibits distinctive spectral peaks, aiding in the identification of their specific chemical compositions and properties. The results of the

EDX analysis will further specify the exact concentrations of substances within the adsorbent material.

3.1.4. EDS analysis

The results of EDS analysis of the raw RP and RPAC are shown in Table 3 and Figures 5 and 6.

Table 3. Result of EDS analysis of raw RP and RPAC.

Elements	Raw RP	RPAC
Carbon (C)	64.1	79.8
Oxygen (O)	35.9	15.9
Nitrogen (N)	0.0	0.0
Potassium (K)	0.0	4.3
Toatl (wt%)	100.0	100.0

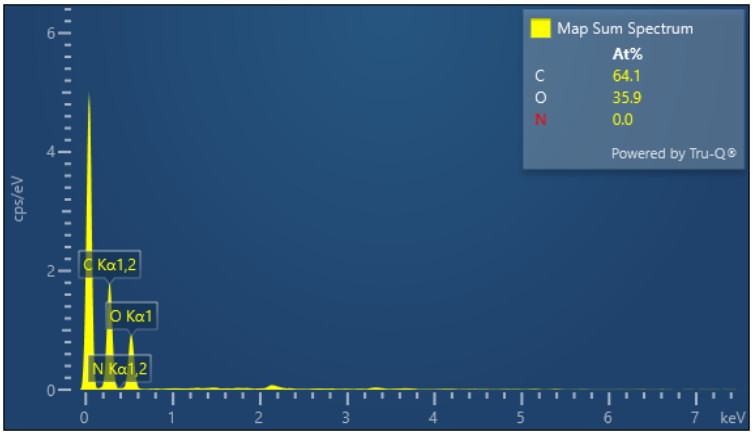


Figure 5. EDS spectrum of raw RP.

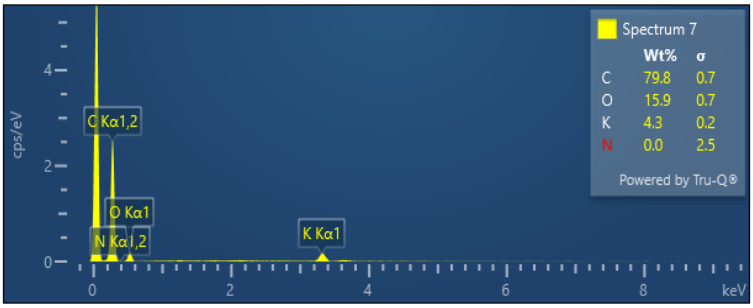


Figure 6. EDS spectrum of RPAC.

The EDS analysis revealed that carbon is the predominant component in both raw RP and RPAC. The carbon content in raw RP is 64.1 wt%, accompanied by an oxygen content of 35.9 wt%. In the case of RPAC, the carbon content increases to 79.8 wt%, while the oxygen content decreases to 15.9 wt%. Notably, there is also a presence of potassium (K) at 4.3 wt% in RPAC. The appearance of K in RPAC results from the treatment of raw RP with KOH during the activation process aimed at producing activated carbon. The levels of C, O and K in RPAC closely resemble those found in activated carbon derived from durian shells [34]. According to the Indonesian Standard SNI 06-3730-1995, the minimum required pure carbon component for activated carbon is 65% [35]. In this experiment, nearly 80% of the composition in RPAC consists of carbon, thus meeting the requirements of the Indonesian Standard (SNI). The high carbon content in rambutan peel can be attributed to its substantial

lignocellulose component. These findings highlight the potential of rambutan peel as a promising raw material for activated carbon production.

3.2. Factors influencing the batch adsorption equilibrium investigations

The batch adsorption equilibrium experiments concerning the adsorption of oil onto RPAC yielded the following results:

3.2.1. Effect of initial oil concentration and isotherm studies

Studying initial oil concentration is crucial due to its strong influence on the thermodynamics and kinetics of the adsorption process [36]. The results of investigating the impact of initial oil concentration on RPAC's oil adsorption capacity and removal efficiency are presented in Figure 7.

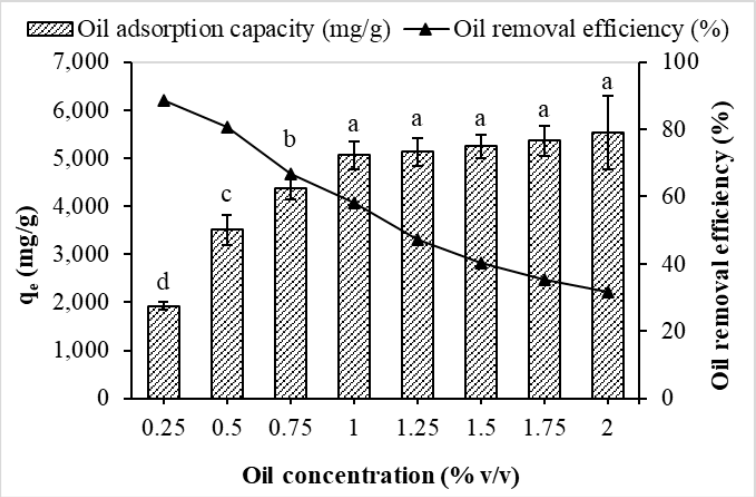


Figure 7. Effect of initial oil concentration on oil adsorption capacity and removal efficiency (Experimental conditions: adsorbent particle size = 0.6 - 1.0, adsorbent dosage = 250 mg per 250 ml of oil-contaminated water, temperature = 25 ± 1°C, contact time = 60 min, pH = 7 ± 0.2, and initial oil concentration = 0.25 - 2.0 % v/v). The error bars represent the mean (SD) for three replicates (n = 3).

At an initial oil concentration of 0.25% v/v, the removal efficiency reached 88.57%. However, this ratio decreased to 31.76% as the initial oil concentration increased to 2% v/v. This decrease in removal efficiency at higher initial concentrations (for the same adsorbent dosage) can be attributed to the reduction in the number of available active sites on the adsorbent's surface responsible for oil binding. This finding aligns with Abel et al. (2020) [12], who observed a similar trend when using coconut coir-activated carbon to remediate oil spills.

The adsorption capacity of the adsorbent increases with higher initial oil concentrations, particularly in the range of initial oil concentrations from 0.25% v/v (1,926.3 mg/g) to 1% v/v (5,064.3 mg/g). This is attributed to the higher ratio of surface active sites on the adsorbent to oil molecules at low initial concentrations (0.25% v/v), causing more oil molecules to interact with the adsorbent and occupy the active sites on its surface. As the initial oil concentration exceeds 1% v/v, oil molecules rapidly occupy the surface active sites of the adsorbent, leading to early saturation. The limited number of active sites available on the adsorbent's surface for oil molecules results in a gradual reduction in removal efficiency. Consequently, the adsorption capacity of the adsorbent only slightly increases when the initial oil concentration shifts from 1% v/v (5,064.3 mg/g) to 2% v/v (5,526.9 mg/g). This finding aligns with experimental data from Cheu et al. (2016), indicating that oil adsorption capacity increases with oil concentration until reaching a certain threshold and reaching

equilibrium. This result is attributed to the saturation of active sites on the adsorbent's surface with oil, hindering further oil adsorption onto these active sites [37].

The ANOVA results indicate significant differences in RPAC's oil adsorption capacity between the initial oil concentration of 1% v/v and the concentrations of 0.25, 0.5, and 0.75% v/v. However, no substantial distinctions were found among the higher initial oil concentrations and 1% v/v. Within the scope of this study, the authors will select the optimal initial oil concentration of 1% v/v to investigate the impact of contact time on RPAC's oil adsorption process.

Based on the initial oil concentration study, the calculated isotherms of Langmuir, Freundlich, Temkin, and Dubin-Radushkevich (D-R) are presented in Figure 8.

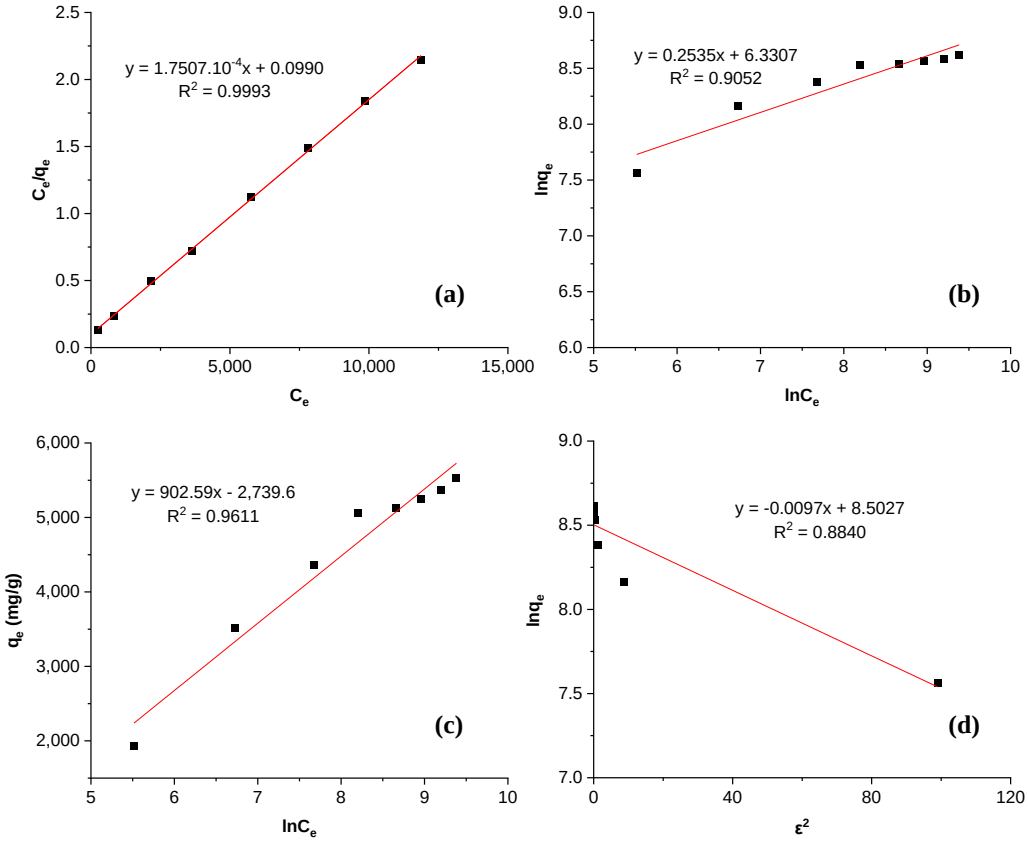


Figure 8. Equilibrium isotherms for oil adsorption onto prepared RPAC: (a) Langmuir, (b) Freundlich, (c) Temkin, and (d) Dubin-Radushkevich isotherms.

Table 4. Parameters and correlation coefficients for various isotherm models in sorption study.

Isotherm	Parameter	Calculated value
Langmuir	q_{max} (mg/g)	5,712.0
	K_L (L/mg)	0.0018
	R^2	0.9993
Freundlich	K_f (mg/g)	561.55
	1/n	0.2535
	R^2	0.9052
Temkin	K_T (L/g)	0.0481
	B (J/mol)	902.59
	b_T	2.7463
	R^2	0.9611
D-R	$Q_{D–R}$ (mg/g)	4,928.1
	K_{ad} (mol^2/KJ^2)	0.0097
	E (KJ/mol)	7.1796
	R^2	0.8840

The calculated isotherms results reveal that the Langmuir isotherm exhibits a higher coefficient R^2 (0.9993) compared to other sorption isotherm models. This implies that the Langmuir isotherm is the most suitable model for describing RPAC's oil adsorption process under the conditions of this study. These results demonstrate that the oil adsorption process on RPAC can be explained using a monolayer adsorption model with consistent conditions for all active sites on RPAC's surface. Accordingly, RPAC's adsorption capacity saturates when all surface active sites are filled with oil molecules. The Langmuir model indicates that as the oil concentration increases, RPAC's adsorption capacity rises and reaches a maximum value (q_{max}) at a certain point. Subsequently, as the initial oil concentration continues to increase, RPAC's adsorption capacity doesn't increase significantly and gradually reaches a stable value. This aligns with the experimental results presented in Figure 7 of this study.

Additionally, the Langmuir isotherm model also shows that RPAC's adsorption process is physical adsorption. This can be explained by the property of RPAC that the existence of pores on the surface. These pores can physically interact with oil molecules through the Van der Waals force.

The Langmuir isotherm model allows the computation of RPAC's maximum adsorption capacity. This can aid in assessing RPAC's effectiveness in oil removal and its practical applications. Accordingly, the maximum adsorption capacity (q_{max}) of RPAC is 5,712.0 mg/g, and the Langmuir constant (K_L) is 0.0018 L/mg.

3.2.2. Effect of contact time on oil sorption and kinetics studies

Adsorption is a process that varies with time, and it is crucial to understand the adsorption rate for designing and assessing the efficiency of adsorbents in oil removal [38]. The influence of contact time on the oil adsorption capacity and oil removal effectiveness of RPAC was investigated within the range of 10 to 90 minutes, and the results are shown in Figure 9.

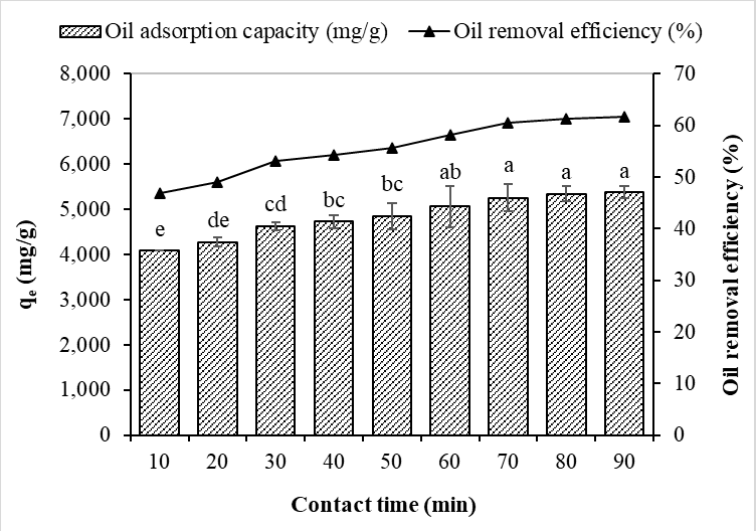


Figure 9. Effect of contact time on oil adsorption capacity and removal efficiency (Experimental conditions: adsorbent particle size = 0.6 - 1.0, adsorbent dosage = 250 mg per 250 ml of oil-contaminated water, temperature = 25 ± 1°C, pH = 7 ± 0.2, and initial oil concentration = 1.0 % v/v, contact time = 10 - 90 min). The error bars represent the mean (SD) for three replicates (n = 3).

The oil adsorption capacity of RPAC rapidly increases within the first 10 minutes of contact with oil, reaching 4,077.7 mg/g and achieving a removal efficiency of 46.87%. This indicates RPAC's strong affinity for diesel oil. After 10 minutes, the oil adsorption capacity continues to rise gradually, showing a tendency to reach equilibrium at 80 minutes (achieving 5,338.3 mg/g adsorption capacity and 61.36% removal efficiency). This can be explained by the fact that during the initial phase, a significant portion of RPAC's surface active sites

remains unoccupied by oil molecules. Consequently, their adsorption capacity and removal efficiency can increase rapidly. However, as the surface active sites become saturated with oil molecules, competition among these molecules gradually reduces the adsorption rate, ultimately reaching equilibrium. This trend aligns with the findings of Ukpong et al. (2020), where the adsorption of crude oil onto activated carbon from coconut husks was rapid in the initial 15 to 30 minutes of contact, followed by a slower pace before finally achieving saturation after 105 minutes [12].

The ANOVA analysis results indicate no significant difference in oil adsorption capacity at different time intervals beyond 60 minutes of contact. Therefore, a 60-minute contact time was chosen as the optimal duration for conducting experiments to investigate the effect of adsorbent dosage on RPAC's oil adsorption capacity. This finding is consistent with Hussein's study, which compared the adsorption capacity of activated carbon from bagasse with a commercial adsorbent and found the maximum adsorption time to be 60 minutes [39]. Compared to the study by [12], the adsorption time of RPAC in this study is faster than the 105 minutes for activated carbon from coconut husks. This difference is due to the type of oil used in the test. Crude oil has a higher viscosity than diesel oil, so the adsorption time to reach equilibrium is longer [39].

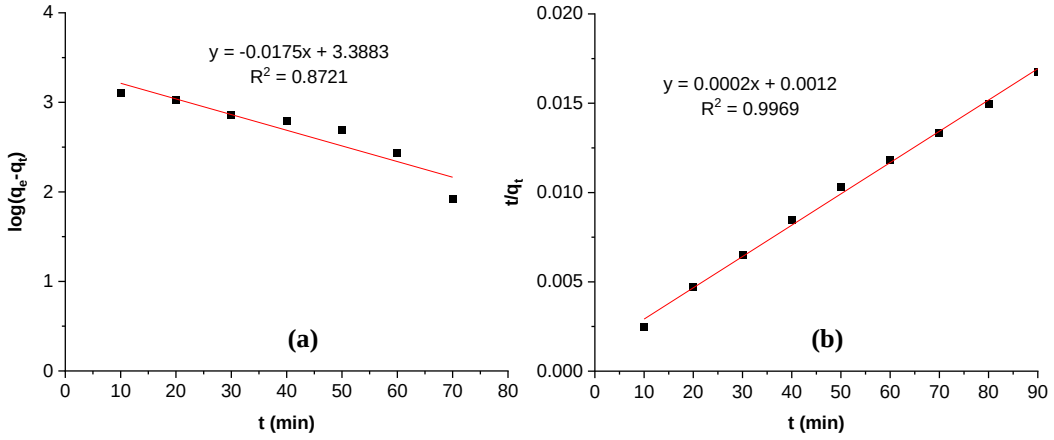


Figure 10. Comparing the (a) pseudo-first-order and (b) pseudo-second-order rate equation models for RPAC.

Table 5. Thermodynamics parameters of crude oil sorption by RPAC for first and second-order models.

$q_{e\ exp}$ (mg/g)	Pseudo-first order			Pseudo-second order			
	$q_{e\ calc}$ (mg/g)	k_1 (1/min)	R^2	$q_{e\ calc}$ (mg/g)	k_2 (g/mg.min)	h (mg/g.min)	R^2
5,338.3	2,445	0.0403	0.8721	5,707	$2.6465.10^{-5}$	862.07	0.9969

where: q_e: the amount of oil adsorbed at equilibrium (mg/g); q_t: the amount of oil adsorbed at time t (mg/g); $q_{e\ exp}$: the experimental adsorption capacity (mg/g); $q_{e\ calc}$: the theoretical adsorption capacity (mg/g).

The calculations based on the Pseudo-first order model yielded an adsorption capacity of 2,445 mg/g, significantly lower than the experimental adsorption capacity of 5,338.3 mg/g. On the other hand, according to the Pseudo-second order model, the theoretical adsorption capacity reached 5,707 mg/g, which closely resembles the experimental adsorption capacity. Additionally, the R-squared value (R^2) of the Pseudo-second order model (R^2 = 0.9969) is higher compared to the Pseudo-first order model (R^2 = 0.8721). Therefore, the Pseudo-second order model (with a rate constant k_2 of $2.6465.10^{-5}$ g/mg.min and a coefficient h of 862.07 mg/g.min) is better suited for describing the oil adsorption process of RPAC under the conditions of this study. The Pseudo-second order model indicates that RPAC's adsorption process exhibits strong adsorption characteristics, particularly at high oil concentrations, and can adsorb a significant amount of oil within a short time frame. This is clearly shown in the

experiment, where RPAC's oil adsorption capacity significantly increases to 4,077.7 mg/g within the first 10 minutes of oil contact. This finding aligns with the research by Anwana Abel et al., where the predominant kinetic model for the adsorption behavior of crude oil on coconut coir-activated carbon was the Pseudo-second order model [12].

3.2.3. Effect of adsorbent dosage

The results of the investigation into the influence of adsorbent dosage are presented in Figure 11.

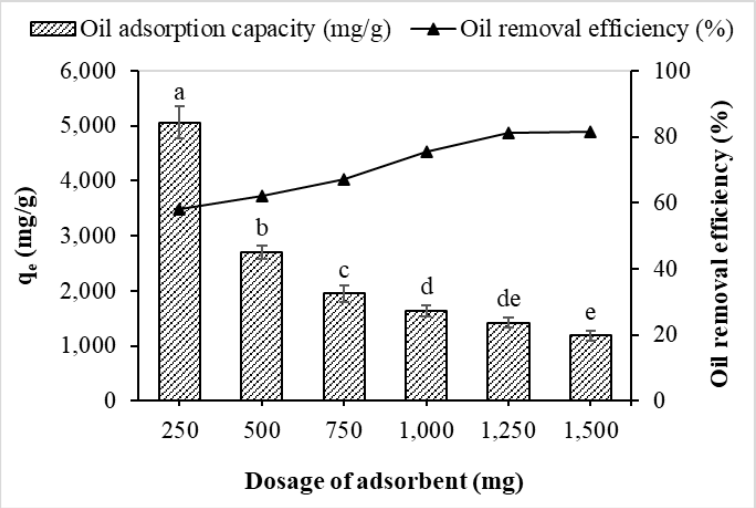


Figure 12. Effect of adsorbent dosage on oil adsorption capacity and removal efficiency (Experimental conditions: adsorbent particle size = 0.6 - 1.0, temperature = 25 ± 1ºC, pH = 7 ± 0.2, initial oil concentration = 1.0 % v/v, contact time = 60 min, and adsorbent dosage = 250 - 1,500 mg per 250 ml of oil-contaminated water). The error bars represent the mean (SD) for three replicates (n = 3).

As the dosage of the adsorbent increases from 250 g to 1,500 mg, the oil adsorption capacity rises from 1,266.1 mg to 1,774.6 mg, and the oil removal efficiency increases from 58.21% to 81.59%. This trend can be attributed to the higher availability of active sites on RPAC's surface as the dosage increases, which becomes occupied by oil molecules, resulting in enhanced oil adsorption. This result aligns with the findings [40], who demonstrated that the oil removal percentage of powdered activated lemon peels is directly proportional to the adsorbent dosage. Larger dosages lead to higher oil adsorption efficiency due to the increased surface area for adsorption to occur.

However, at higher dosages in this specific study, such as 1,250 mg (with an oil removal efficiency of 81.34%) and 1,500 mg (with an oil removal efficiency of 81.59%), the removal efficiency nearly plateaus. This could be attributed to an excess of RPAC in the environment, hindering the penetration of oil molecules to occupy active sites on the surface, thus limiting further adsorption capacity and preventing complete oil removal. This phenomenon has also been demonstrated by [41], where the adsorption capacity increased with increasing adsorbent dosage from 500 mg to 1,500 mg, followed by a decreasing trend. This reduction could be due to the compaction of fibers at higher dosages, potentially reducing the even penetration of oil into the fibers. Determining the optimal dosage of adsorbent is crucial as it significantly impacts the capacity and removal efficiency of the oil adsorption process in real-world conditions. This helps save resources, optimize processing efficiency, protect the environment, and reduce costs, while also creating favorable conditions for adsorbent reuse. The calculated oil adsorption capacity per gram of material (q_e) gradually decreases from 5,064.3 mg/g to 1,183.1 mg/g as the RPAC dosage increases from 250 mg to 1,500 mg.

The results of the ANOVA analysis indicate a significant difference in oil adsorption capacity between the dosage of 250 mg and other dosages. Therefore, the RPAC dosage of 250 mg was chosen as the optimal dosage for conducting experiments to investigate the effect of pH on the oil adsorption process of RPAC.

3.2.4. Effect of pH

The solution's pH significantly affects adsorption because it influences the adsorbent's surface properties and binding sites [42]. Moreover, changes in acidity or alkalinity in the oil solution can increase oil solubility, indicating a reduction in the oil's hydrophobicity. Consequently, this leads to a decrease in oil adsorption on the hydrophobic adsorbent surface, especially in highly acidic or alkaline solutions [37]. The investigation results regarding the impact of pH on the oil adsorption capacity and oil removal efficiency of RPAC are presented in Figure 13 within the pH range of 3 to 11.

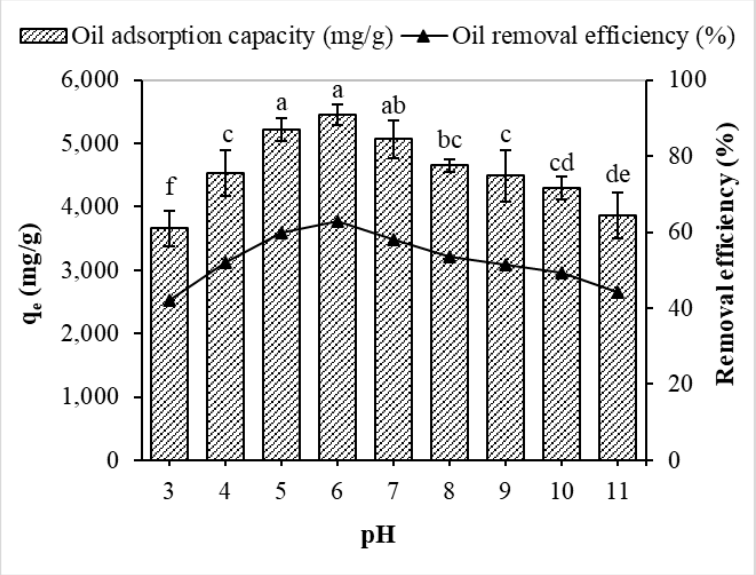


Figure 13. Effect of pH on oil adsorption capacity and removal efficiency (Experimental conditions: adsorbent particle size = 0.6 - 1.0, temperature = 25 ± 1°C, initial oil concentration = 1.0 % v/v, contact time = 60 min, adsorbent dosage = 250 mg per 250 ml of oil-contaminated water, and pH = 3 - 11). The error bars represent the mean (SD) for three replicates (n = 3).

Increasing the solution's pH from 3 to 6 enhances the oil adsorption capacity from 3,660.3 mg/g to 5,458.3 mg/g, increasing oil removal efficiency from 42.07% to 62.74%. At lower pH levels, the adsorbent's hydrophobic nature diminishes due to a considerable number of protons adsorbed on the adsorbent's surface [37]. Specifically, at lower pH (acidic conditions), there are many protons (H^+ ions) in the solution. These protons attach to the adsorbent's surface and neutralize its charged parts. This reduces the surface's ability to attract hydrophobic (water-repelling) molecules, making the biosorbent less hydrophobic and less effective at adsorbing such molecules. As the pH continues to rise, the oil adsorption capacity tends to decrease. Specifically, the oil adsorption capacity and removal efficiency drop to 3,859.0 mg/g and 44.36%, respectively, when the pH reaches 11. This phenomenon stems from pH's influence on the adsorbent's net charge and the oil surface. In an alkaline environment, protons are removed from charged functional groups on the surface of the adsorbent in the adsorption reaction mixture, thereby diminishing the overall positive charge and the affinity of the adsorption sites for negatively charged oil molecules [28]. In short, at acidic pH levels, the adsorbent surface carries a positive charge, which promotes the electrostatic attraction and subsequent adsorption of negatively charged contaminants (anions). In neutral pH conditions, the adsorbent surface charge approaches neutrality,

leading to adsorption through a combination of forces, including van der Waals interactions, hydrogen bonding, and electrostatic interactions. In alkaline environments, at high pH levels, adsorbent acquires a negative surface charge, allowing it to attract and adsorb positively charged contaminants (cations) through electrostatic interactions.

The ANOVA analysis results indicate no substantial difference in oil adsorption capacity among solution pH values of 5, 6, and 7. Nonetheless, a notable difference in oil adsorption capacity is observed between pH values of 5, 6, and 7 when compared to the pH values of the remaining solutions. Consequently, a solution pH of 6, demonstrating the most effective oil adsorption capacity, is selected as the optimal pH for the experimental optimization process. Other studies also show similar results, with optimal oil adsorption efficiency occurring within the neutral pH range. For instance, the research by [43], utilizing activated carbon derived from mango peels, determined an optimal pH of around 7.

3.3. Comparative analysis of oil adsorption capacity with other activated carbon materials

Several activated carbon materials sourced from different agricultural waste byproducts are compared with RPAC in this study by the author, and the results are presented in Table 6.

Table 6. Comparative analysis of oil adsorption capacity.

Materials	Method	Oil type	Maximum adsorption capacity	References
Barley Straw	Carbonized, 400 °C	Gas oil	8.5 - 9 g/g	[39]
Rice husks	Carbonized	Diesel oil	5.02 g/g	[44]
Rice husks	Carbonized	Diesel oil	2.78 g/g	[44]
Rice husks	Carbonized, 480 °C	Diesel oil	5.5 kg/kg	[45]
Coconut Coir	Carbonized, 600 °C	Crude oil	4,859.5 mg/g	[12]
RPAC	Carbonized, 550 °C	Diesel oil	5,712.0 mg/g	This study

Table 6 shows that the maximum adsorption capacity of RPAC in this study is higher than that of pyrolyzed rice husk reported in the research conducted by [44] and [45], as well as that of activated carbon from coconut coir [12]. However, compared to the findings [39], the maximum adsorption capacity of RPAC is lower than that of carbonized oat straw at 400°C for 0.5 to 3 hours. The activated carbon adsorbent derived from rambutan in this study excels in diesel adsorption compared to numerous activated carbon materials derived from agricultural by-products. This underscores the potential of RPAC as a highly effective oil absorbent for pollution remediation. Nevertheless, the treatment effectiveness of the adsorbent may vary depending on pyrolysis conditions, the type of tested oil, and the properties of the materials used in activated carbon production.

4. Conclusion

Activated carbon derived from rambutan peel demonstrated highly effective oil removal capabilities in water. The process involved carbonization and subsequent KOH activation, resulting in a more porous adsorbent with significantly increased surface area: 485.7680 (m^2/g) and 786.0143 (m^2/g), respectively. The pore volume also saw a substantial increase, going from 0.005293 to 0.05392 (cm^3/g), while the pore diameter expanded from 21.6 to 55.2 nm. Particularly noteworthy is the alkaline treatment, which generated larger and deeper pores, enhancing their abundance on the material's surface. Individual evaluations of factors influencing RPAC's oil removal efficiency and adsorption capacity identified optimal values. These include a 1% v/v oil concentration, a 60-minute contact time, a 250 mg adsorbent dosage, and a pH of 6. The adsorption process conformed well to the Langmuir isotherm model (R^2 = 0.9993), yielding a maximum adsorption capacity of 5,712.0 mg/g. Furthermore,

the pseudo-second-order model (R^2 = 0.9969) proved to be a more suitable fit for RPAC's oil adsorption process.

Author contribution statement: Defining and developing the research idea and researchframework: T.V.N., L.H.B.; Collecting data and literature, data analysis and synthesis: T.T.N., N.D.L.; Experimental research: T.T.N., N.D.L.; Drafting the manuscript: T.T.N.; Manuscript editing and revision: T.V.N., L.H.B.

Acknowledgement: We would like to express our gratitude for the support in terms of time, facilities, and financial resources provided by HUTECH University for conducting this study.

Competing interest statement: The authors declare that this article was the work of theauthors, has not been published elsewhere, has not been copied from previous research; therewas no conflict of interest within the author group.

Reference

1. ITOPF. *Tanker Spill Statistics 2021*; ITOPF Ltd, London, UK., 2022, 1–20.
2. Kolokoussis, P.; Karathanassi, V. Oil spill detection and mapping using Sentinel 2 imagery. *J. Mar. Sci. Eng.* **2018**, *6(1)*, 1–12.
3. Wardley-Smith, J. The control of oil pollution. Graham & Tortman Ltd, 1976.
4. Alaa El-Din, G.; Amer, A.A.; Malsh, G.; Hussein, M. Study on the use of banana peels for oil spill removal. *Alexandria Eng. J.* **2018**, *57(3)*, 2061–2068.
5. El-Nafaty, U.A.; Muhammad, I.M.; Abdulsalam, S. Biosorption and kinetic studies on oil removal from produced water using banana peel. *J. Civ. Eng. Res.* **2013**, *3(6)*, 125–136.
6. Idris, J.; Eyu, G.D.; Mansor, A.M.; Ahmad, Z.; Chukwuekezie, C.S. A preliminary study of biodegradable waste as sorbent material for oil-spill cleanup. *Sci. World J.* **2014**, *2014*, 1–5.
7. Lim, T.T.; Huang, X. Evaluation of Kapok (Ceiba Pentandra (L.) Gaertn.) as a natural hollow hydrophobic–oleophilic fibrous sorbent for oil spill cleanup. *Chemosphere* **2007**, *66(5)*, 955–963.
8. Banerjee, S.S.; Joshi, M.V.; Jayaram, R.V. Treatment of oil spill by sorption technique using fatty acid grafted sawdust. *Chemosphere* **2006**, *64(6)*, 1026–1031.
9. Kadirvelu, K.; Namasivayam, C. Activated carbon from coconut coirpith as metal adsorbent: Adsorption of Cd(II) from aqueous solution. *Adv. Environ. Res.* **2003**, *7(2)*, 471–478.
10. Maulion, R.V.; Abacan, S.A.; Allorde, G.G.; Umali, M.C.S. Oil spill adsorption capacity of activated carbon tablets from corncobs in simulated oil-water mixture. *Asia Pac. J. Multidiscip. Res.* **2015**, *3(5)*, 146–151.
11. Aljeboree, A.M.; Alshirifi, A.N.; Alkaim, A.F. Kinetics and equilibrium study for the adsorption of textile dyes on coconut shell activated carbon. *Arab. J. Chem.* **2017**, *10*, S3381–S3393.
12. Anwana Abel, U.; Rhoda Habor, G.; Innocent Oseribho, O. Adsorption studies of oil spill clean-up using coconut coir activated carbon (CCAC). *ASEAN J. Chem. Eng.* **2020**, *8(2)*, 36–47.
13. Atemkeng, C.D.; Anagho, G.S.; Tagne, R.F.T.; Amola, L.A.; Bopda, A.; Kamgaing, T. Optimization of 4-nonylphenol adsorption on activated carbons derived from safou seeds using response surface methodology. *Carbon Trends* **2021**, *4*, 100052.
14. Igwegbe, C.A.; Umembamalu, C.J.; Osuagwu, E.U.; Oba, S.N.; Emembolu, L.N. Studies on adsorption characteristics of corn cobs activated carbon for the removal of oil and grease from oil refinery desalter effluent in a downflow fixed bed adsorption equipment. *Eur. J. Sustain. Dev.* **2020**, *5(1)*, 1–14.
15. Sahu, J.N.; Agarwal, S.; Meikap, B.C.; Biswas, M.N. Performance of a modified multi-stage bubble column reactor for lead(ii) and biological oxygen demand

removal from wastewater using activated rice husk. *J. Hazard. Mater.* **2009**, *161(1)*, 317–324.
16. Angelova, D.; Uzunov, I.; Uzunova, S.; Gigova, A.; Minchev, L. Kinetics of oil and oil products adsorption by carbonized rice husks. *Chem. Eng. J.* **2011**, *172(1)*, 306–311.
17. Husseien, M.; Amer, A.; El- Maghraby, A.; Taha, N. Experimental investigation of thermal modification influence on sorption qualities of barley straw. *J. Appl. Sci. Res.* **2008**, *4(6)*, 652–657.
18. Huong, D. Agricultural by-products: Resources are being wasted. Government Electronic Newspaper, 2021. (in Vietnam)
19. Department of Horticulture. Handbook of farming techniques according to VietGAP for 10 major fruit trees. Ministry of Agriculture and Rural Development of Vietnam, 2022.
20. Oliveira, E.; Santos, J.; Goncalves, A.P.; Mattedi, S.; Jose, N. Characterization of the rambutan peel fiber (nephelium lappaceum) as a lignocellulosic material for technological applications. *Chem. Eng. Trans.* **2016**, *50*, 391–396.
21. Olga, V.R.; Darina, V.I.; Alexandr, A.I.; Alexandra, A.O. Cleanup of water surface from oil spills using natural sorbent materials. *Procedia Chemistry* **2014**, *10*, 145–150.
22. Al-Jammal, N.; Juzsakova, T. Review on the effectiveness of adsorbent materials in oil spills clean up. Proceedings of the 7th International Conference of ICEEE, Budapest, Hungary **2016**, 131–138.
23. Normah, N. et al. Hydrothermal carbonization of rambutan peel (*Nephelium lappaceum L.*) as a Green and low-cost adsorbent for Fe(II) removal from aqueous solutions. *Chem. Ecol.* **2022**, *38*, 1–17.
24. Ahmad, M.A.; Alrozi, R. Optimization of rambutan peel based activated carbon preparation conditions for remazol brilliant blue R removal. *Chem. Eng. J.* **2011**, *168(1)*, 280–285.
25. Omer, A.M.; Khalifa, R.E.; Tamer, T.M.; Ali, A.A.; Ammar, Y.A.; Mohy Eldin, M. S. Kinetic and thermodynamic studies for the sorptive removal of crude oil spills using a low-cost chitosan-poly (butyl acrylate) grafted copolymer. *Desalin. Water Treat.* **2020**, *192*, 213–225.
26. Kowanga, K.D.; Gatebe, E.; Mauti, G.O.; Mauti, E.M. Kinetic, sorption isotherms, pseudo-first-order model and pseudo-second-order model studies of cu(ii) and pb(ii) using defatted moringa oleifera seed powder. *J. Psychopharmacol.* **2016**, *5(2)*, 71–78.
27. Ho, Y.S.; McKay, G. Pseudo-second order model for sorption processes. *Process Biochem.* **1999**, *34(5)*, 451–465.
28. Nnamdi Ekwueme, B.; Anthony Ezema, C.; Asadu, C.O.; Elijah Onu, C.; Onah, T. O.; Sunday Ike, I.; Chinonyelum Orga, A. Isotherm modelling and optimization of oil layer removal from surface water by organic acid activated plantain peels fiber. *Arab. J. Chem.* **2023**, *16(2)*, 104443.
29. Ike, I.S.; Asadu, C.O.; Ezema, C.A.; Onah, T.O.; Ogbodo, N.O.; Godwin-Nwakwasi, E.U.; Onu, C.E. ANN-GA, ANFIS-GA and thermodynamics base modeling of crude oil removal from surface water using organic acid grafted banana pseudo stem fiber. *Appl. Surf. Sci.* **2022**, *9, 100259*.
30. Onu, C.E.; Nwabanne, J.T.; Ohale, P.E.; Asadu, C.O. Comparative analysis of RSM, ANN and ANFIS and the mechanistic modeling in eriochrome black-t dye adsorption using modified clay. *S. Afr. J. Chem. Eng.* **2021**, *36*, 24–42.
31. Hasanah, M.; Juleanti, N.; Priambodo, A.; Arsyad, F.; Lesbani, A.; Mohadi, R. Utilization of rambutan peel as s potential adsorbent for the adsorption of malachite

green, procion red, and congo red dyes. *Ecol. Eng. Environ. Technol.* **2022**, *23*, 148–157.

32. Njoku, V.O.; Foo, K.Y.; Asif, M.; Hameed, B.H. Preparation of activated carbons from rambutan (Nephelium lappaceum) peel by microwave-induced KOH activation for acid yellow 17 dye adsorption. *Chem. Eng. J.* **2014**, *250*, 198–204.
33. Aziz, A.; Mohamad Yusop M.F.; Ahmad, M.A. Removal of bisphenol S from aqueous solution using activated carbon derived from rambutan peel via microwave irradiation technique. *Sains Malays.* **2022**, *51(12)*, 3967–3980.
34. Tran, Q.T. et al. Experimental design, equilibrium modeling and kinetic studies on the adsorption of methylene blue by adsorbent: activated carbon from durian shell waste. *Materials (Basel)* **2022**, *15(23)*, 8566
35. Patmawati, Y.; Alwathan; Ramadani, N.H. Characterization of activated carbon prepare from low-rank coal of east kalimantan by using acid and base activation. Proceedings of the 8th Annual Southeast Asian International Seminar (ASAIS 2019) **2019**, 178–181.
36. Sokker, H.H.; El-Sawy, N.M.; Hassan, M.A.; El-Anadouli, B.E. Adsorption of crude oil from aqueous solution by hydrogel of chitosan based polyacrylamide prepared by radiation induced graft polymerization. *J. Hazard. Mater.* **2011**, *190(1)*, 359–365.
37. Cheu, S.C.; Kong, H.; Song, S.T.; Saman, N.; Johari, K.; Mat, H. High removal performance of dissolved oil from aqueous solution by sorption using fatty acid esterified pineapple leaves as novel sorbents. *RSC Adv.* **2016**, *6(17)*, 13710–13722.
38. Peng, D.; Lan, Z.; Guo, C.; Yang, C.; Dang, Z. Application of cellulase for the modification of corn stalk: leading to oil sorption. *Bioresour. Technol.* **2013**, *137*, 414–418.
39. Hussein, M.; Amer, A.A.; Sawsan, I.I. Oil spill sorption using carbonized pith bagasse: 1. preparation and characterization of carbonized pith bagasse. *J. Anal. Appl. Pyrolysis.* **2008**, *82(2)*, 205–211.
40. Tembhurkar, A.; Deshpande, R. Powdered activated lemon peels as adsorbent for removal of cutting oil from wastewater. *J. Hazard. Toxic Radioact. Waste* **2012**, *16(4)*, 311–315.
41. NwabuezeH, H.O.; Chiaha, P.N.; Ezekannagha, B.C.; Okoani, O.E. Acetylation of corn cobs using iodine catalyst, for oil spills remediation. *Int. J. Eng. Sci.* **2016**, *5(9)*, 53–59.
42. Wan Ngah, W.S.; Hanafiah, M.A.K.M. Adsorption of copper on rubber (hevea brasiliensis) leaf powder: kinetic, equilibrium and thermodynamic studies. *Biochem. Eng. J.* **2008**, *39(3)*, 521–530.
43. Olufemi, B.A.; Otolorin, F. Comparative adsorption of crude oil using mango (mangnifera indica) shell and mango shell activated carbon. *Environ. Eng. Res.* **2017**, *22(4)*, 384–392.
44. Vlaev, L.; Petkov, P.; Dimitrov, A.; Genieva, S. Cleanup of water polluted with crude oil or diesel fuel using rice husks ash. *J. Taiwan Inst. Chem. Eng.* **2011**, *42(6)*, 957–964.
45. Angelova, D.; Uzunova, S.; Uzunov, I.; Anchev, B. Dependence between oil sorption capacity of pyrolyzed rice husks and the composition and amount of fluids dispersed on their surface. *J. Chem. Technol. Metall.* **2012**, *47(2)*, 147–154.

JOURNAL OF
HYDRO-METEOROLOGY

Research Article

Development of an ensemble dynamic-probabilistic prediction model for tropical storm genesis in the Vietnam East Sea using the Logistic Regression approach

Dao Nguyen-Quynh Hoa[1]*, Tran Tan Tien[1]

[1] VNU University of Science, Vietnam National University; hoadao@vnu.edu.vn; tientt@vnu.edu.vn

***Corresponding author: hoadao@vnu.edu.vn; Tel.: +84–8857758771**

Received: 05 September 2023; Accepted: 23 October 2023; Published: 25 December 2023

Abstract: In this study, a logistic regression model is developed to forecast tropical storm (TS) genesis in the Vietnam East Sea from 2012 to 2019. The model incorporates seven potential predictors including dynamic and thermodynamic parameters at formation time retrieved from the WRF-LETKF outputs. After rigorous testing, six predictors are selected, excluding minimum sea-level pressure. In a broader context, the logistic regression model performs promisingly, generating forecast probabilities that enhance the accuracy of TS genesis predictions, particularly in early forecast cycles. The model's regression coefficients and forecast outcomes align well with test dataset results, affirming its stability and validity. As a result, the forecast probability from this model can be effectively employed as a probabilistic forecast value for predicting TS genesis status.

Keywords: Tropical cyclogenesis; WRF-LETKF; Ensemble prediction system; Logistic regression.

1. Introduction

Tropical storm (TS) genesis forecasting is recognized as one of the most challenging aspects of numerical weather prediction (NWP). The ability to accurately predict the formation of TSs is crucial for effectively managing and mitigating associated risks. With the significant advancements in computational capabilities, NWP models have started to assume a prominent role in forecasting TS genesis and tropical cyclone (TC) as a whole. Consequently, numerous studies have been undertaken to evaluate the reliability and accuracy of predictions regarding the formation and development of TCs to TS intensity as indicated by these models [1–3].

Beyond the utilization of a single-model deterministic forecasts, the adoption of ensemble forecasts for TS genesis prediction has gained popularity, showing superior predictive capabilities. Ensemble prediction systems are designed to address uncertainteis in initial condiitons and imperfections in model formulation, with the goal of providing a range of potential future atmospheric scenarios [4]. Despite the substantial volume of data required for processing, ensemble forecasting has demonstrated its effectiveness in enhancing the accuracy of tropical cyclones [5]. In a study assessing TS genesis from tropical cloud clusters using two global ensemble prediction systems, ECMWF-EPS and UKMO-EPS, conducted over the years 2018 to 2020 in the Northwestern Pacific, Northeastern Pacific, and Northern Atlantic regions. The author [6] reported relatively good forecasting skills. They found that the quality of probabilistic forecasts could be further enhanced by combining predictions from all multi-model ensemble members. Additionally, the authors observed that in cases

where on member did not provide an accurate prediction, favorable conditions for TC genesis were often present. This underscors the utility and value of ensemble forecasting systems in improving the prediction of TS genesis. The Local Ensemble Transform Kalman Filter (LETKF) scheme, pioneered at the University of Maryland [7, 8], has found widespread application in various numerical models, with notable use in the Weather Research and Forecasting (WRF) model [9–11]. Research by [9] highlighed the LETKF's exceptional utility in managing highly diverse data, such as satellite observations. The incorporation of multi-physics ensemble prediction of this scheme has demonstrated ist effectiveness in predicting the genesis of TC Wutip (2013). Building upon their case study, this study delves further into examining the efficacy of the WRF-LETKF when assimialting augmented observations in forecasting the likelihood of 45 TCs that occurred between 2012 and 2019 over the Vietnam East Sea.

Along with dynamical prediction model, numerous studies have delved into the application of statistical models for forecasting the development of TSs [12–15]. A common thread across these diverse researches is the incorporation of seasonal or climatological factors of various scales as predictors in their statistical models. A recent approach is applying these statistical model to the products of pre-existing NWP systems, so called the dynamical-probabilistic forecast models [16]. However, relatively few studies have concentrated on short-term events with lead times of up to 5 days. One of the most frequently employed techniques for probability prediction is the logistic regression model [17]. This model is capable of forecasting and discerning the likelihood of an event occurring, particularly the genesis of TCs [18].

Building upon the foundation of existing scientific knowledge, this study offers a dynamical-probabilistic forecast model to predict TS genesis over the Vietnam East Sea, expanding to ensemble forecast. The construction involves results from the WRF-LETKF forecasts and logistic regression model, with a specific focus on the likelihood of TS genesis from predefined events. This investigation sheds light on vital insights into the predicted TS genesis using products from ensemble prediction system. The subsequent sections of this paper are organized as follows: Section 2 outlines the experimental design. Section 3 provides a brief discussion of the results concerning the direct outputs to predict TC genesis various forecast cycles. Section 4 selects the associated predictors and section 5 constructs the probabilistic prediction model. Finally, Section 6 offers a summary and engages in further discussion.

2. Experiment settings

2.1. Ensemble prediction system

The Local Ensemble Transform Kalman Filter (LETKF) algorithm are applied to the non-hydrostatic version of the WRF model version 3.9.1 to create the ensemble-based data assimilation system (hereafter, WRF-LETKF) with variational data assimilation scheme 3DVAR at cold-start. Additionally, the algorithm's construction and its application in the case study of Wutip (2013) were thoroughly elucidated in the study of [19].

The model employs an ensemble size of 21 members, generating perturbations based on the atmospheric state from the previous cycle. These perturbations are integrated into the global deterministic analysis using, maintain uniform scaling with an inflation factor of 1,1 for assimilated variables to enhance the influence of ensemble noise. The multi-physic approach optimizes spread without requiring a larger number of members and combines various parameterization schemes. These schemes include 2 convective schemes, 3 boundary-layer schemes, 3 microphysics schemes, and 2 shortwave-longwave radiation schemes, resulting in a total of 36 potential combinations. From these, 21 combinations were selected for operational

purposes (refer to Table 1 [19]), considering model stability and minimizing internal conflicts during implementation.

The WRF model was configured with a domain that is large enough to cover the entire Vietnam East Sea and the surrounding waters of the Northwestern Pacific, an area with coordinates [95°E - 145°E; 0° - 30°N] (Figure 1). The domain has a spatial resolution of 27 km and includes 31 vertical levels and the forecast lead time is 5 days (120-hr forecasts) within a hourly interval.

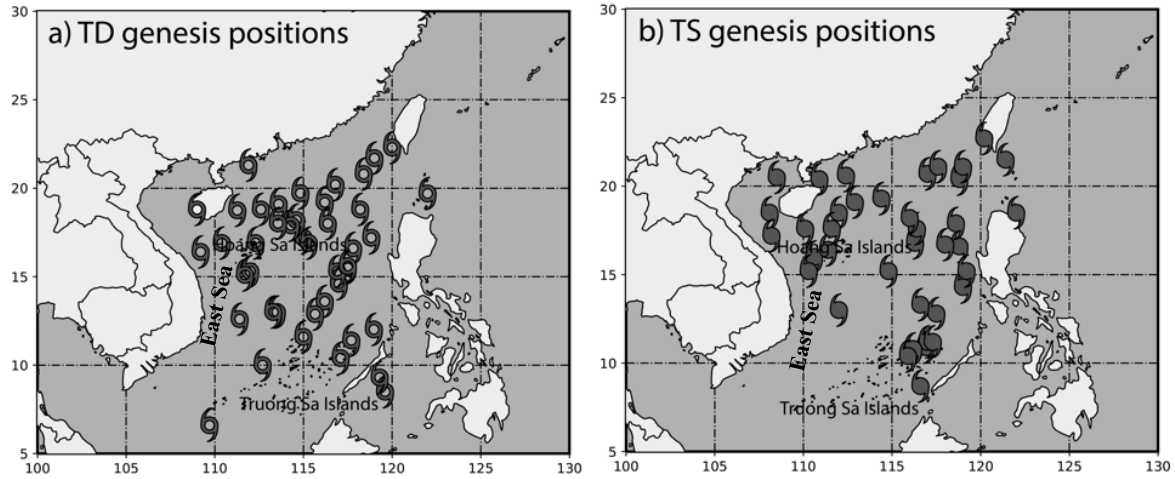


Figure 1. Illustration of the positions where 45 TCs formed (a) and developed into tropical storms (b) in the Vietnam East Sea during the period from 2012 to 2019 retrieved from IBTraCS best-track.

2.2. Data collection and TC tracking

The study relies on the International Best Track Archive for Climate Stewardship (IBTraCS) [20] to serve as the benchmark for TD formation ensemble prediction and for statistically verifying subsequent TS development. The investigation focuses exclusively on TCs in their early stages within the Vietnam East Sea. If TC maximum 10-m sustained wind speed reaches 20 kt (TD intensity), the location and timing of the cyclone in best-track is recorded as TD genesis (Figure 1a). Similarly, when the TC's intensity reaches or exceeds 34 kt (TS), it is considered as a development case, indicating that it has transitioned into a tropical storm (Figure 1b). The initial and boundary conditions for the numerical experiments were derived from the NCEP GFS analysis provided every 3-hr at 0.5° × 0.5° horizontal resolutions.

In this study, the forecasted timing and location of each TCs from TD formation to TS stage is important, as we examine the state of the atmosphere at the genesis of TD and its subsequent potential to reach TS intensity. We employ a straightforward yet essential tracking algorithm, with a specific emphasis on their early development stages over the ocean. The tracking algorithm follows the work made by [19]. The criteria used for selecting cyclones that develop into TSs involve local minimum sea-level pressure (P_{min}) having adjacent local maximum low-level vorticity (ζ_{low}) and minimum geopotential height at 700 hPa while scanning every grid point for each lead time of every forecast. All the extrema are required to exhibit at least 2 closed contours of 2 hPa; 10^{-5} s^{-1}; and 4 dams, respectively, to avoid noise. When TC's maximum 10-meter wind speed (V_{max}) nearest to the P_{min} location exceeds 20 kt, it is recorded as a TD formation. The same implementation has been used to select and verify TC development to TS stage, when V_{max} surpasses a threshold of 34 kt for the first time. In some cases, the phases of TC formation (reaching TD) and development (reaching TS) may overlap. Our definitions for successful TD and TS formation forecasts also consider the location predicted by the model in comparison to real storm's best-track. We do so by selecting TC centers that fall inside the Vietnam East Sea region, within 5° radius proximity to the best-track recorded within a 120-hr lead time.

Over the 2012-2019 period, we have collected a dataset comprising a total of 45 recorded TDs, while 35 among them continue to develop to TS stage and the rest are dissipated over the Vietnam East Sea.

2.3. Logistic regression model predicting the tropical storm genesis

2.3.1. Predictors

The purpose is to forecast the probability of TC development reaching TS intensity based on a pre-existing TC formation predicted in ensemble forecasts. The study employs a logistic regression model with the dependent variable as a quantitative variable predicting a probability with values within the range [0; 1]. The initial objective of this study is to establish the predictors for modeling the occurrence probability of TS genesis in the Vietnam East Sea region (refer to Table 1). As outlined in Table 1, these predictors consist of both dynamic parameters (minimum sea-level pressure, low-level relative vorticity, mid-level vertical velocity, and vertical wind shear) and thermodynamic parameters (moist static energy, surface latent heat flux, and low-level horizontal moisture convergence). Research conducted by several authors [21–23] indicates that these critical meteorological parameters can effectively differentiate between developing and non-developing disturbances, particularly in the context of TS formation. The local environment refers to atmospheric state centered on the location of a TC within 5° radius and standardized due to differences in dimension and scale.

Table 1. Descriptions of logistic regression model variables.

Categorization	Variable	Descriptions
Dynamic	P_{min}	Minimum sea-level pressure
	ζ_{low}	Average mid-to-low-level vertical vorticity $\zeta_{low} = \int_{850}^{500} \zeta \left(\frac{dp}{g}\right)$
	ω_{mid}	Average vertical velocity in 700 – 500hPa
	V_{sh}	Vertical shear between 200 and 850 hPa $V_{sh} = \sqrt{(u_{200} - u_{850})^2 + (v_{200} - v_{850})^2}$
Thermodynamic	MSE	Column-integrated moist static energy normalized by C_p $MSE = \frac{\int_{p_s}^{0} C_p T + gz + L_v q_v \left(\frac{dp}{g}\right)}{C_{pd}}$
	SLHF	Surface latent heat flux
	HMC_{low}	Low-level horizontal moisture convergence $HMC_{low} = -\int_{p_s}^{700hPa} \frac{\Delta \overline{uq_v}}{\Delta x} + \frac{\Delta \overline{vq_v}}{\Delta y} \left(\frac{dp}{g}\right)$

Student t-test is conducted on the probability distribution function of each predictor, based on 2 datasets categorized as having TC development reaching TS intensity (DEV) and not having such development (NON-DEV). Null hypothesis H_0: $\bar{x}_{i,1} = \bar{x}_{i,0}$; Alternative hypothesis H_1: $\bar{x}_{i,1} \neq \bar{x}_{i,0}$.

where $\bar{x}_{i,1}$ is the mean value of the i-th predictor in DEV group, and $\bar{x}_{i,0}$ is the corresponding mean value in the NON-DEV group. If the null hypothesis is rejected, and the alternative hypothesis is accepted with a 95% confidence level, meaning that there is a significant difference in the parameter's value between the two groups, then the parameter is selected as a predictor to be included in the logistic regression model.

2.3.2. Logistic regression model configurations

The logistic regression model is based on the concept of linear regression for classification problems. Starting with the output of a linear regression function, the logistic regression model uses the sigma function to find the probability distribution of data within

the range [0; 1]. Assume that we have a regression function with a set of n independent variables $x = (1, x_1, x_2, \dots, x_n)$:

$$\hat{y} = g(x) = w_0 + w_1 x_1 + \cdots + w_n x_n = w^T x \quad (1)$$

Here, represents the regression coefficients. To transform this equation using the sigma function to predict probabilities and introduce non-linearity into the regression model, we have the following expression:

$$p(y = 1|x; w) = \sigma(w^T x) = \frac{1}{1 + e^{-w^T x}} \quad (2)$$

$p(y = 1|x; w)$ is the conditional probability of the event y = 1 occurring based on the independent variables x and the regression coefficients w.

To assess the regression model's validity through cross-validation analysis, we conduct cross-validation of the regression model on a dataset consisting of 45 cases of TC genesis during the period 2012-2019, with 35 cases of TCs developing to TS and 10 of no development. This is done by dividing the dataset into 5 subsets (5-fold cross-validation). To ensure that the train and test datasets have a sufficient number of cases for both TC development and non-development, the number of TCs in each subset is established as follows:

- Train data: 28 developing TCs and 8 non-developing TCs.
- Test data: 7 developing TCs and 2 non-developing TCs.

In addition, to test the utility of the predictive equation and derive the final forecasting equations, the study employs the Wald test to assess the influence of the predictors on the regression model. The study uses $\alpha = 5\%$ as statistical significance threshold for the regression coefficients. The predicted results from ensemble dynamic-statistical model are determined from the average probability values of the ensemble members:

Here, M = 21 is the total number of ensemble members, and $p_{i,j}$ represents the probability of predicting TC development from the j-th ensemble member for the i-th forecast. $p_{i,j}$ is equal to 0 by default if the ensemble member does not predict TD formation in the Vietnam East Sea within the 120-hr forecast period based on the dynamical model's output.

2.3.3. Verification metrics

It is important to note that our study primarily centers around determining whether or not TS genesis has occurred within the Vietnam East Sea within a 120-hr lead time. In this context, the study does not specifically focus on the precise timing and location of the occurrence in comparison to observations. Consequently, the selection of verification metrics in this study is tailored to the specific goal of assessing the accuracy of forecasted probabilities (through Brier score) and the categorization of events versus non-events (through AUC-ROC).

- Brier score: $BS = \frac{1}{N}\sum_{i=1}^{N}(p_i - a_i)^2$ calculates the forecast accuracy by comparing the forecasted probability with the actual observation, indicating if the event occurred or not [24]. N is the total number of observed events. A lower BS indicates a better forecast alignment with reality (BS range: 0-1).

- AUC-ROC assesses the binary classification model performance, measuring a model's ability to distinguish positive and negative classes [25, 26]. A high AUC-ROC value (close to 1) suggests accuracy, close to 0 indicates inverse predictions, and 0.5 signifies poor classification.

3. Genesis forecasting in the WRF-LETKF

Figure 2 represents the verification scores of TD genesis and TS genesis over the Vietnam East Sea for 8 years during 2012-2019 from each ensemble member from WRF-LETKF. For TD genesis forecasts, the WRF-LETKF ensemble system demonstrates reasonably good forecasting skill up to a 5-day lead time, with the lowest BS occurring at

approximately 3.5 days before the formation (equivalent to 84-hr cycle). With the 48-hr forecast cycle, the ensemble system shows a stable increase in the probability of correctly predicted cases.

However, when it comes to forecasting the development of TCs reaching TS stage (TS genesis), the Brier score indicates a significant decrease compared to the TD genesis forecast, with most values exceeding 0.1 and a clear decreasing trend in the near forecast cycles. When combined with the AUC-ROC skill score, it is observed that the WRF-LETKF performs well in classifying the development of TCs to TS intensity in the forecast cycles starting from 96-hr, 84-hr, and 48-hr onward (with AUC-ROC values exceeding 0.6). In the forecast cycle of 120-hr and 108-hr, the model fails to classify the possibility of TD development into storms accurately.

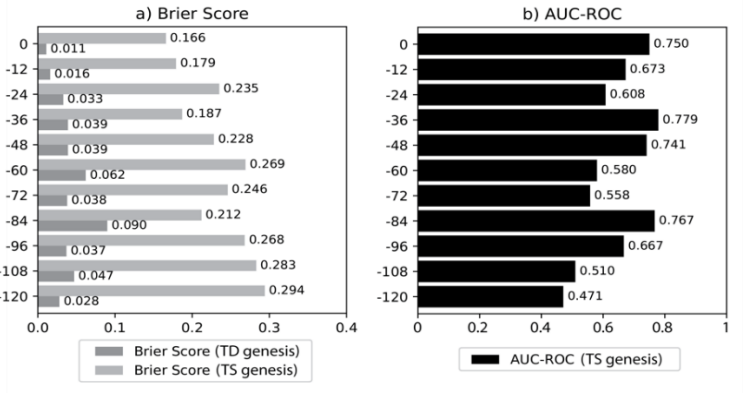


Figure 2. Brier score (a) and AUC-ROC (b) assessing the prediction of TD and TS genesis in the Vietnam East Sea using direct forecast products from WRF-LETKF.

In other words, an accurate forecast for the event of a TD genesis may not necessarily convey more accurate information about the TD's development into a full-fledged TS. It is the motivation of this study to use products in TD forecasted by the WRF-LETKF system to predict the probability of TS in a statistical combination.

4. Selection of predictors

In the context of multivariate regression model, it is crucial to avoid strong correlations among the independent variables. A high correlation suggests that the features of one independent variable closely coincide with those of another, potentially leading to a substantial reduction in the reliability of the model's regression coefficients. Hence, in the current study, we conducted an in-depth assessment aimed at evaluating multicollinearity among the independent variables and explore the relationships between these independent and dependent variables (Figure 3).

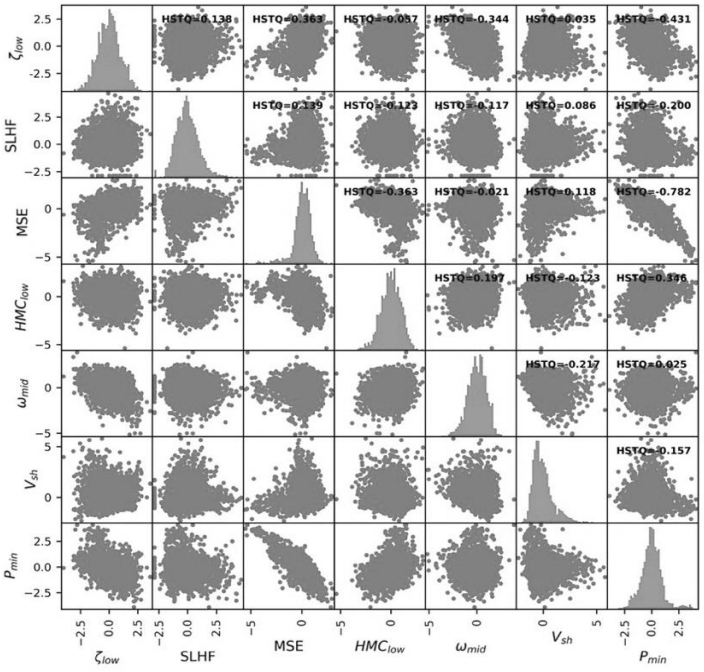


Figure 3. Correlation matrix between pairs of candidate predictors, specific correlation coefficients are highlighted in each subplot. The diagonal line represents probability distribution functions of individual parameters.

Results indicate that out of the total 7 candidate predictors selected to construct the logistic regression

equation from the forecasted product, P_{min} exhibits the highest correlation with the other variables. Specifically, the highest correlation coefficient value is observed between P_{min} and the normalized MSE (-0.782; inverse correlation). The findings also highlight the cause-and-effect relationship between the dynamic and thermodynamic features concerning the development of TCs, as manifested through the sea-level pressure at the cyclone center. Consequently, with the objective of building a logistic regression equation from independent predictors, P_{min} is not utilized as a candidate variable.

The predictor variables are directly chosen from the remaining candidate variables, which include Student t-tests. The selection results, as shown in Table 2, demonstrate that, at a statistical significance level of 5%, all candidate predictor variables meet the criteria and are selected as forecasting variables. Therefore, all dynamic and thermodynamic parameters selected for evaluation at the time of TD genesis exhibit distinct characteristics between developing and non-developing TC groups up to TS intensity.

5. Results

5.1. Ensemble dynamic - probabilistic model detecting TS genesis

Table 2. Regression coefficients and statistical test values in logistic regression model.

Eq.	Variable	Regression coefficient	Wald	P-value	Eq.	Variable	Regression coefficient	Wald	P-value
	Intercept*	1.6243	1343.6146	<0.01		intercept*	1.6661	1350.3482	<0.01
	$\zeta_{\lambda o \omega}$*	0.3576	50.2584	<0.01		$\zeta_{\lambda o \omega}$*	0.4036	61.2068	<0.01
1	*SLHF**	0.4075	78.5723	<0.01	2	*SLHF**	0.4268	82.0777	<0.01
	*MSE**	1.4384	404.3414	<0.01		*MSE**	1.4215	405.1923	<0.01
	HMC_{low}*	0.2898	36.3438	<0.01		HMC_{low}*	0.3384	47.6302	<0.01
	$\omega_{\mu\iota\delta}$	-0.0246	0.2533	0.6147		$\omega_{\mu\iota\delta}$	-0.0162	0.108	0.7425
	V_{sh}	-0.0444	1.1774	0.2779		V_{sh}	-0.076	3.2888	0.0698
	intercept*	1.615	1321.1707	<0.01		intercept*	1.5918	1322.6212	<0.01
	$\zeta_{\lambda o \omega}$*	0.3662	51.6439	<0.01		$\zeta_{\lambda o \omega}$*	0.3658	53.2859	<0.01
3	*SLHF**	0.4947	110.8745	<0.01	4	*SLHF**	0.4133	81.8117	<0.01
	*MSE**	1.4487	415.7788	<0.01		*MSE**	1.3863	400.4556	<0.01
	HMC_{low}*	0.2965	38.7205	<0.01		HMC_{low}*	0.2942	37.8519	<0.01
	$\omega_{\mu\iota\delta}$	-0.0137	0.0786	0.7792		$\omega_{\mu\iota\delta}$	-0.0123	0.0634	0.8012
	V_{sh}	-0.0813	3.7935	0.051		V_{sh}	-0.0415	1.0286	0.3105
	intercept*	1.6016	1318.983	<0.01					
	$\zeta_{\lambda o \omega}$*	0.3677	53.4547	<0.01		*Values are significant at 95% level.			
5	*SLHF**	0.4353	87.701	<0.01					
	*MSE**	1.4676	421.0099	<0.01					
	HMC_{low}*	0.3307	47.9665	<0.01					
	ω_{mid}	-0.0015	0.001	0.9752					
	V_{sh}	-0.0632	2.4137	0.1203					

From the results shown in Table 2, the regression coefficients of most predictors, except for ω_{mid} and V_{sh}, are positive, indicating a positive relationship between thermodynamic factors at the time of formation and the development of TCs to TS intensity. This implies that the likelihood of TCs reaching TS intensity increases with low-level vorticity and enhanced moisture convergence at the time of formation. The results clearly demonstrate the impact of vorticity, humidity, and low-level moisture convergence at formation on the forecast probability of TC development to TS, with statistical significance (p-value < 5%).

Conversely, the estimated regression coefficients for ω_{mid} and V_{sh}, are negative, suggesting that mid-level vorticity and decreased wind shear in the environment contribute

to the potential for TC development. However, the Wald test indicates that these parameters do not have a significant impact (ω_{mid} has a p-value > 0.6 in all prediction equations) because they do not meet the 95% confidence level of the test. With these results, the probabilistic equations for the development of TCs in the Vietnam East Sea from the logistic regression model as follows: $\mathrm{p(TS\ genesis)} = \frac{1}{1+e^{-z}}$ with:

Eq. 1: $z = 1.6243 + 1.4384 \times MSE + 0.4075 \times SLHF + 0.3576 \times \zeta_{low} + 0.2898 \times HMC_{low}$

Eq. 2: $z = 1.6661 + 1.4215 \times MSE + 0.4268 \times SLHF + 0.4036 \times \zeta_{low} + 0.3384 \times HMC_{low}$

Eq. 3: $z = 1.6150 + 1.4487 \times MSE + 0.4947 \times SLHF + 0.3662 \times \zeta_{low} + 0.2965 \times HMC_{low}$

Eq. 4: $z = 1.5918 + 1.3863 \times MSE + 0.4133 \times SLHF + 0.3658 \times \zeta_{low} + 0.2942 \times HMC_{low}$

Eq. 5: $z = 1.6016 + 1.4676 \times MSE + 0.4353 \times SLHF + 0.3677 \times \zeta_{low} + 0.3307 \times HMC_{low}$

5.2. Validation of the model predictability

The predictive outcomes of the logistic regression model concerning the potential development of TCs reaching TS stage in the Vietnam East Sea are assessed across test data reflecting variations between different forecasting cycles (Table 3). In general, BS values are relatively low, mostly below 0.2, in all forecasting cycles and prediction cases, demonstrating the model's high predictive accuracy. Specifically, BS is at its lowest value in the 60-hr forecast cycle (~2,5 days) when the logistic regression model is applied, especially in Eq. 3, with a value of 0.0893.

Table 3. Forecasting skill scores evaluating the probabilistic model on test data.

Fcst. cycles (hrs)	Ensemble dynamic-probabilistic model on test data									
	Eq. 1		Eq. 2		Eq. 3		Eq. 4		Eq.5	
	AUC-ROC	BS	AUC-ROC	BS	AUC-ROC	BS	AUC-ROC	BS	AUC-ROC	BS
-120	0.6538	0.1482	0.7083	0.1515	0.6731	0.1582	0.8846	0.1495	0.8077	0.1355
-108	0.7885	0.1369	0.8654	0.1391	0.8077	0.1344	0.8462	0.1337	0.5417	0.0972
-96	0.6667	0.1681	0.5417	0.1723	0.5833	0.1668	0.4792	0.1718	0.6875	0.1556
-84	0.7556	0.1073	0.6667	0.1145	0.6667	0.1248	0.6222	0.1170	0.7333	0.1033
-72	0.6667	0.1672	0.7308	0.1538	0.6731	0.1598	0.7692	0.1397	0.6346	0.1650
-60	0.7949	0.0893	0.7949	0.1063	0.7885	0.1353	0.8974	0.1008	0.8077	0.1377
-48	0.7115	0.1522	0.7292	0.1443	0.7436	0.1218	0.7179	0.1165	0.8205	0.1041
-36	0.7708	0.1504	0.7727	0.1590	0.6250	0.1539	0.8333	0.1466	0.7576	0.1360
-24	0.8974	0.0977	0.7500	0.1497	0.5385	0.1595	0.7500	0.1537	0.6923	0.1514
-12	0.8393	0.1317	0.6538	0.1472	0.7857	0.1412	0.7500	0.1522	0.7857	0.1393
0	0.7308	0.1527	0.7115	0.1593	0.8205	0.1211	0.8077	0.1435	0.6875	0.1630

Furthermore, AUC-ROC are relatively high, ranging between 0.8 to 0.9 from all equations. This indicates that the regression models effectively differentiate between cases of TC development and non-development based on the environmental variables described. It also suggests that the regression model performs most effectively during the 60-hr forecast cycle compared to other cycles. In contrast, during the 96-hr cycle, all the prediction equations yield BS values exceeding 0.16, and AUC-ROC hovers around 0.5.

An overall assessment reveals that among the proposed prediction equations, Eq. 1 demonstrates the highest predictive capability across the test dataset, with relatively high AUC-ROC (the lowest being 0.6538 at the 120-hr cycle) and BS values below 0.1 in the 60-hr and 24-hr cycles before the formation.

Table 4. Forecasting skill scores evaluating the probabilistic model on test data.

Fcst. cycles (hrs)	Ensemble dynamic-probabilistic model in 2012-2019 period									
	Eq. 1		Eq. 2		Eq. 3		Eq. 4		Eq. 5	
	AUC-ROC	BS	AUC-ROC	BS	AUC-ROC	BS	AUC-ROC	BS	AUC-ROC	BS
-120	0.7885	0.1507	0.7692	0.1504	0.8077	0.1497	0.8077	0.1517	0.8077	0.1508
-108	0.8846	0.1198	0.8846	0.1192	0.8846	0.1171	0.8846	0.1207	0.8846	0.1192
-96	0.7500	0.1725	0.7500	0.1732	0.7500	0.1719	0.7500	0.1737	0.7500	0.1729
-84	0.8333	0.1256	0.8333	0.1256	0.8333	0.1253	0.8333	0.1271	0.8333	0.1263
-72	0.6731	0.1926	0.6731	0.1921	0.6731	0.1918	0.6731	0.1926	0.6731	0.1918
-60	0.8462	0.1459	0.8462	0.1454	0.8462	0.1442	0.8462	0.1463	0.8462	0.1453
-48	0.7500	0.1464	0.7500	0.1464	0.7500	0.1447	0.7500	0.1464	0.7500	0.1457
-36	0.7292	0.1602	0.7083	0.1603	0.7083	0.1601	0.7292	0.1609	0.7083	0.1598
-24	0.6923	0.1483	0.6923	0.1489	0.6923	0.1481	0.6923	0.1490	0.6923	0.1483
-12	0.7143	0.1964	0.7143	0.1977	0.7143	0.1971	0.7143	0.1969	0.7143	0.1970
0	0.6923	0.1987	0.6731	0.1998	0.6923	0.1979	0.6923	0.1988	0.6923	0.1984

By examining BS and AUC-ROC, it is evident that the dynamical-statistical hybrid model has seen a significant improvement compared to the forecast results analyzed in Figure 2, especially in the early forecast cycles before actual formation time. In these cycles, all prediction equations exhibit notably reduced BS values, ranging from over 0.23-0.3 (Figure 2) down to 0.11-0.19 (Table 4). This reduction indicates that the forecast model provides a more accurate description of whether or not a TC will develop to TS intensity. The lowest BS values are achieved in the 120-hr and 108-hr forecast cycles preceding the actual formation, with BS decreasing from 0.294 to an average of 0.15 and from 0.283 to an average of 0.12, respectively. Although the results do not significantly differ between the various forecast cycles, the analysis shows that BS increases in the 12-hr preceding the formation compared to the original ensemble forecasting products, suggesting that the potential development of at TD to TS intensity can be better identified when the initial conditions incorporate basic information about atmospheric circulation and the nearby TC structure at the time of formation.

Observing BS and AUC-ROC, no significant differences are found among the logistic regression equations, indicating similarity in performance across the cross-validated train and test datasets. The verification metrics for each prediction equation on the cross-validated test dataset (Table 3) and the entire dataset (Table 4) yield similar results, ensuring the model's consistent forecasting capabilities. Therefore, the optimized logistic regression model, based on the proposed ensemble forecasting results, can be considered an effective tool for forecasting the likelihood of TC development to TS intensity in the Vietnam East Sea.

6. Summary and conclusion

While the general performance of the assimilated multi-physics WRF-LETKF in short-term forecasting, regarding tropical disturbances evolving into TD, is quite accurate when assessing probabilities and errors. However, the pure dynamical model's forecast for the likelihood of further development into TS exhibits lower accuracy. In the study, we tried to develop a simple statistical model from these products and discussed the capability of forecasting the TS genesis of TCs in the SCS using the logistic regression model. We develop an ensemble dynamic-probabilistic forecast model aiming at forecasting a probability of the higher or lower level of TS genesis frequency of TC in each of the 45 TC formation events during the period 2012-2019.

The forecast model incorporated a total of seven candidate predictors, which encompassed both dynamic parameters (minimum sea-level pressure, relative vorticity of the

lower-level, vertical velocity of the middle-level, and vertical wind shear) and thermodynamic parameters (moist static energy, surface latent heat flux, accumulated moisture convergence of the lower-level). Out of these seven variables, six were selected as predictors for the logistic regression model after multicollinearity verification and a significance test, except for minimum sea-level pressure. We developed a combination of equations for each train-test datasets in cross-validation procedure and concluded that vertical velocity at lower level and vertical wind shear have a minor impact to the probabilistic forecast, therefore excluding them from our newly developed logistic regression model.

In a broader context, when considering the performance of the logistic regression equations derived from the ensemble prediction products, the results are promising. The model generates a forecast probability between 0 and 1 for each member forecast. The combination of probabilities between ensemble members for a forecast lead time determines the overall likelihood of TS genesis. The forecasting skill using this approach has improved compared to direct forecast products, especially in the early forecast cycles before the actual formation occurs. Additionally, the results of the regression coefficients and forecasts produced by the model align closely with the outcomes from the utilization of test dataset, thus indicating the stability and validity of the newly developed model as an accurate statistical tool. Consequently, the forecast probability generated by this model can be effectively utilized as a probabilistic forecast value in predicting the status of TS genesis.

The logistic regression model offers the advantage of providing rapid forecasting information while accommodating the non-linear nature of variables. In the future, our focus will be on extending the forecast time range of the model and incorporating a more comprehensive set of variables for a predictive model of TS genesis.

Author contributions: Conceptualization: T.T.T., D.N.Q.H.; Methodology: D.N.Q.H.; Data curation: D.N.Q.H.; Writing-original draft: D.N.Q.H.; Writing-editing: T.T.T.

Acknowledgement: The authors are grateful for the time and effort given by the anonymous reviewers whose contributions greatly strengthened this manuscript.

Statement: The authors collectively declare that this article represents our research work, has not been previously published elsewhere, and is not copied from any previous studies. There are no conflicting interests among the group of authors.

References

1. Halperin, D.J.; Fuelberg, H.E.; Hart, R.E.; Cossuth, J.H. Verification of tropical cyclone genesis forecasts from global numerical models: Comparisons between the North Atlantic and eastern North Pacific basins. *Wea. Forecasting* **2016**, *31*, 947–955.
2. Liang, M.; Chan, J.C.L.; Xu, J.; Yamaguchi, M. Numerical prediction of tropical cyclogenesis Part I: Evaluation of model performance. *Quart. J. Roy. Meteor. Soc.* **2021**, *147*, 1626–1641.
3. Jaiswal, N.; Kishtawal, C.M.; Bhomia, S.; Pal, P.K. Multi-model ensemble-based probabilistic prediction of tropical cyclogenesis using TIGGE model forecasts. *Meteor. Atmos. Phys.* **2016**, *128*, 601–611.
4. Pedlosky, J. Geophysical fluid dynamics. 1979.
5. Zhang, X.; Yu, H. A probabilistic tropical cyclone track forecast scheme based on the selective consensus of ensemble prediction systems. *Wea. Forecasting* **2017**, *32*, 2143–2157.
6. Zhang, X.; Fang, J.; Yu, Z. The forecast skill of tropical cyclone genesis in two global ensembles. *Wea. Forecasting* **2023**, *38(1)*, 83–97. https://doi.org/10.1175/WAF-D-22-0145.1.

7. Hunt, B.R.; Kostelich, E.J.; Szunyogh, I. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena* **2007**, *230*, 112–126.
8. Miyoshi, T.; Kunii, M. Using AIRS retrievals in the WRF-LETKF system to improve regional numerical weather prediction. *Tellus A* **2012**, *64(1)*, 18408. https://doi.org/10.3402/tellusa.v64i0.18408.
9. Kieu, C.Q.; Truong, N.M.; Mai, H.T.; Ngo-Duc, T. Sensitivity of the track and intensity forecasts of Typhoon Megi (2010) to satellite-derived atmospheric motion vectors with the ensemble Kalman filter. *J. Atmos. Oceanic Technol.* **2012**, *29(12)*, 1794–1810. https://doi.org/10.1175/jtech-d-12-00020.1.
10. Liu, J.; Fertig, E.; Li, H.; Kalnay, E.; Hunt, B.; Kostelich, E.; Szunyogh, I.; Todling, R. Comparison between local ensemble transform Kalman filter and PSAS in the NASA finite volume GCM - Perfect model experiments. *Nonlinear Processes Geophys.* **2007**, *15*, 645–659. https://doi.org/10.5194/npg-15-645-2008.
11. Kwon, H.; Lee, W.; Won, S.H.; Cha, E.J. Statistical ensemble prediction of the tropical cyclone activity over the western North Pacific. *Geophys. Res. Lett.* **2007**, *34*, 24805. https://doi.org/10.1029/2007GL032308.
12. Leroy, A.; Wheeler, M. Statistical prediction of weekly tropical cyclone activity in the Southern Hemisphere. *Mon. Weather Rev.* **2008**, *136*, 3637–3654. https://doi.org/10.1175/2008MWR2426.1.
13. Mestre, O.; Hallegatte, S. Predictors of Tropical Cyclone Numbers and Extreme Hurricane Intensities over the North Atlantic Using Generalized Additive and Linear Models. J. Clim. **2009**, *22*, 633–648. https://doi.org/10.1175/2008JCLI2318.1.
14. Chan, J.; Shi, J.E.; Lam, C.M. Seasonal forecasting of tropical cyclone activity over the western North Pacific and the South China Sea. *Wea. Forecasting* **1998**, *13*, 997–1004. https://doi.org/10.1175/1520-0434(1998)013<0997:SFOTCA>2.0.CO;2.
15. Wijnands, J.; Qian, G.; Kuleshov, Y. Variable selection for tropical cyclogenesis predictive modeling. *Mon. Weather Rev.* **2016**, *144*, 4605–4619. https://doi.org/10.1175/MWR-D-16-0166.1.
16. Wilks, D.S. Statistical methods in the atmospheric sciences. *Int. Geophys. Series* **2006**, 59, xi. https://doi.org/10.1016/S0074-6142(06)80036-7.
17. Choi, J.W.; Kang, K.R.; Kim, D.W.; Kim, T.R. Development of a probability prediction model for tropical cyclone genesis in the northwestern pacific using the logistic regression method. *J. Korean Earth Sci. Soc.* **2010**, *31(5)*, 454–464. https://doi.org/10.5467/JKESS.2010.31.5.454.
18. Tien, T.T.; Hoa, D.N.Q.; Thanh, C.; Kieu, C. Assessing the impacts of augmented observations on the forecast of Typhoon Wutip (2013)'s formation using the ensemble Kalman filter. *Wea. Forecasting* **2020**, *35(4)*, 1483–1503. https://doi.org/10.1175/waf-d-20-0001.1.
19. Knapp, K.R.; Kruk, M.C.; Levinson, D.H.; Diamond, H.J.; Neumann, C.J. The international best track archive for climate stewardship (IBTrACS): Unifying tropical cyclone data. *Bull. Am. Meteorol. Soc.* **2010**, *91(3)*, 363–376.
20. Fu, B.; Peng, M.S.; Li, T.; Stevens, D.E. Developing versus nondeveloping disturbances for tropical cyclone formation. Part II: Western North Pacific. *Mon. Weather Rev.* **2012**, *140(4)*, 1067–1080. https://doi.org/10.1175/2011MWR3618.1.
21. Kerns, B.W.; Chen, S.S. Cloud clusters and tropical cyclogenesis: Developing and nondeveloping systems and their large-scale environment. *Mon. Weather Rev.* **2013**, *141(1)*, 192–210. https://doi.org/10.1175/MWR-D-11-00239.1.
22. Gray, W. Environmental influences on tropical cyclones. *Aust. Meteorol. Mag.* **1988**, *36*, 127–139.

23. Brier, G.W. Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* **1950**, *78(1)*, 1–3.
24. Swets, J.A. The relative operating characteristic in psychology. *Science* **1973**, *182*, 990–1000.
25. Buizza, R.; Miller, M.; Palmer, T.N. Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. 1999, pp. 2887-2908.

JOURNAL OF
HYDRO-METEOROLOGY

Research Article

Landslide susceptibility mapping based on the Weights of Evidence model for mountainous areas of Quang Nam province, Vietnam

Tran Anh Tuan [1,2]*, Tran Thi Tam [3], Pham Viet Hong [1], Nguyen Thi Anh Nguyet [1]

[1] Institute of Marine Geology and Geophysics, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam; tatuan@imgg.vast.vn; hc18052001@yahoo.com; nguyet.imgg@gmail.com

[2] Graduate University of Science and Technology, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam; tatuan@imgg.vast.vn

[3] Center for Agricultural Meteorological Research, Vietnam Institute of Meteorology, Hydrology and Climate change, No.23 - 62 Alley, Nguyen Chi Thanh Road, Dong Da District, Hanoi Vietnam; tam.tranthi@imh.ac.vn

*Corresponding author: tatuan@imgg.vast.vn; Tel.: +84–985150307

Received: 5 September 2023; Accepted: 27 October 2023; Published: 25 December 2023

Abstract: This study shows the results of landslide susceptibility mapping for the southwest region of Quang Nam province using the Weights of Evidence (WoE) model. The input data consists of a landslide inventory and ten influencing factors, i.e., geology, distance to fault, elevation, relief amplitude, slope, aspect, rainfall, soil type, land use, and distance to road. The landslide inventory was constructed from three principal sources: fieldwork survey, legacy data from previous studies, and additional analytical data from high-resolution Google Earth satellite imagery. The landslide locations were randomly categorized into two parts in the ratio 70/30: 70% (811 landslides) for modeling and 30% (348 landslides) for verification. All input data are normalized and constructed into the GIS landslide database. The results of the multicollinearity test show that no collinearity existed between ten input variables. The computation of the weights for classes of influencing factors from 70% of the landslide data using the WoE model has allowed the establishment of the landslide susceptibility map. The model performance was evaluated by using the receiver operating characteristic (ROC) analysis. The area under the curve (AUC) was computed for the success rate curve (using 70% landslide data) and the prediction rate curve (using 30% landslide data) at 0.855 and 0.844, respectively. Thus, it can be confirmed that the landslide susceptibility mapping based on the WoE model was very reliable in the study area.

Keywords: Landslide susceptibility; Weights of Evidence; GIS; Quang Nam province.

1. Introduction

Vietnam is heavily affected by the negative impacts of global climate change [1]. Heavy and irregular rainfall was one of the consequences of climate change that caused natural disasters to occur with increasing intensity and frequency. The estimated damage caused by natural disasters was about 1.5% of Vietnam's GDP per year [2]. Notably, landslides have been a dangerous type of natural disaster that has caused damage in the mountainous areas of North and Central Vietnam during the rainy season [3–7]. Landslides

and flash floods caused 46 deaths, 17 missing, and total economic losses of 11 trillion VND in Quang Nam province in 2020 [8].

Nowadays, many solutions to prevent landslides have been implemented such as engineering, non-engineering, and adaptive solutions. Landslide susceptibility mapping is one of the necessary adaptive solutions for disaster damage reduction. Landslide susceptibility mapping methods have been developed widely with increasing accuracy, in which GIS-based landslide susceptibility mapping was an effective method for identifying and zoning landslide-prone areas [9] such as geomorphic and landslide inventory techniques, multi-criteria analysis, statistically based models, deterministic, and machine learning approaches [10]. Each method group had advantages and disadvantages and was suitable for different scales [11]. The statistical methods constructed based on the framework of statistical science have been widely used in landslide susceptibility assessment. They were divided into two main groups: bivariate statistical methods and multivariate statistical methods. The commonly used bivariate statistical methods were Frequency Ratio/Likelihood Ratio [12, 13], Weights of Evidence [14–16], and Information Value/Statistics Index [3, 17] methods. Multivariate statistical methods determine the weight of each input variable to the total landslide susceptibility instead of assessing the single relationship of each influencing factor to landslide occurrence as known in bivariate statistical approaches [10, 18]. Multivariate statistical methods have been used frequently as logistic regression [17, 19] and discriminant [20, 21] methods.

Generally, the statistical methods give good predictive results. Their accuracy increases as past landslide events are investigated in more detail. Although statistical approaches have been used relatively commonly in the world in landslide susceptibility mapping, they are still quite limited in Vietnam. In this group of methods, selecting the appropriate research scale and level of data detail is very significant because it affects the accuracy of the research results. This study has specified a model suitable for applying landslide susceptibility mapping, which is the WoE model. The study area is the mountainous district of southwestern Quang Nam province, Vietnam, where landslides occur frequently, causing significant loss of life and property.

2. Data used and methods

2.1. Study area and data used

2.1.1. The study area description

The study area is the mountainous region of Quang Nam province with a total area of about 2838 km^2 (Figure 1). This site includes poor districts but is home to unique cultural values of ethnic minorities such as Co Tu, Ca Dong, Xo Dang, etc. These are favorable conditions for the development of ethnic and cultural tourism. In addition, the natural conditions are suitable for forestry and hydropower development, notably the Song Tranh and Dak Mi terraced hydroelectric systems. However, the natural conditions here are also favorable for various natural disasters, including landslides and debris flows.

The study area has a complex geological structure, thick weathered crust, intense destructive fault zones, and is located in a relatively strong earthquake zone. The terrain conditions are high mountainous, steep slopes, complex cleavage, and many large mountain systems with mountains over 2000 m in height. Notably, Ngoc Linh is the highest peak of the Truong Son range (2598 m) between Quang Nam and Kon Tum provinces. The study area belongs to the influence of one of the heavy rain centers in Vietnam. Monitoring data for the three years 2019, 2020, and 2021 shows that the total annual rainfalls are 3209 mm, 5541 mm, and 5201 mm, respectively. The rainy season usually starts in September of the previous year and lasts until January of the following year. October 2020 was a period of

record heavy rain, especially when Typhoon Molave made landfall in Vietnam (October 28, 2020). The total rainfall of this month measured at Tra My station was up to 1880mm (Figure 2). The landslides occurred widely in central Vietnam at that time. Many landslides have caused terrific damage to people and property.

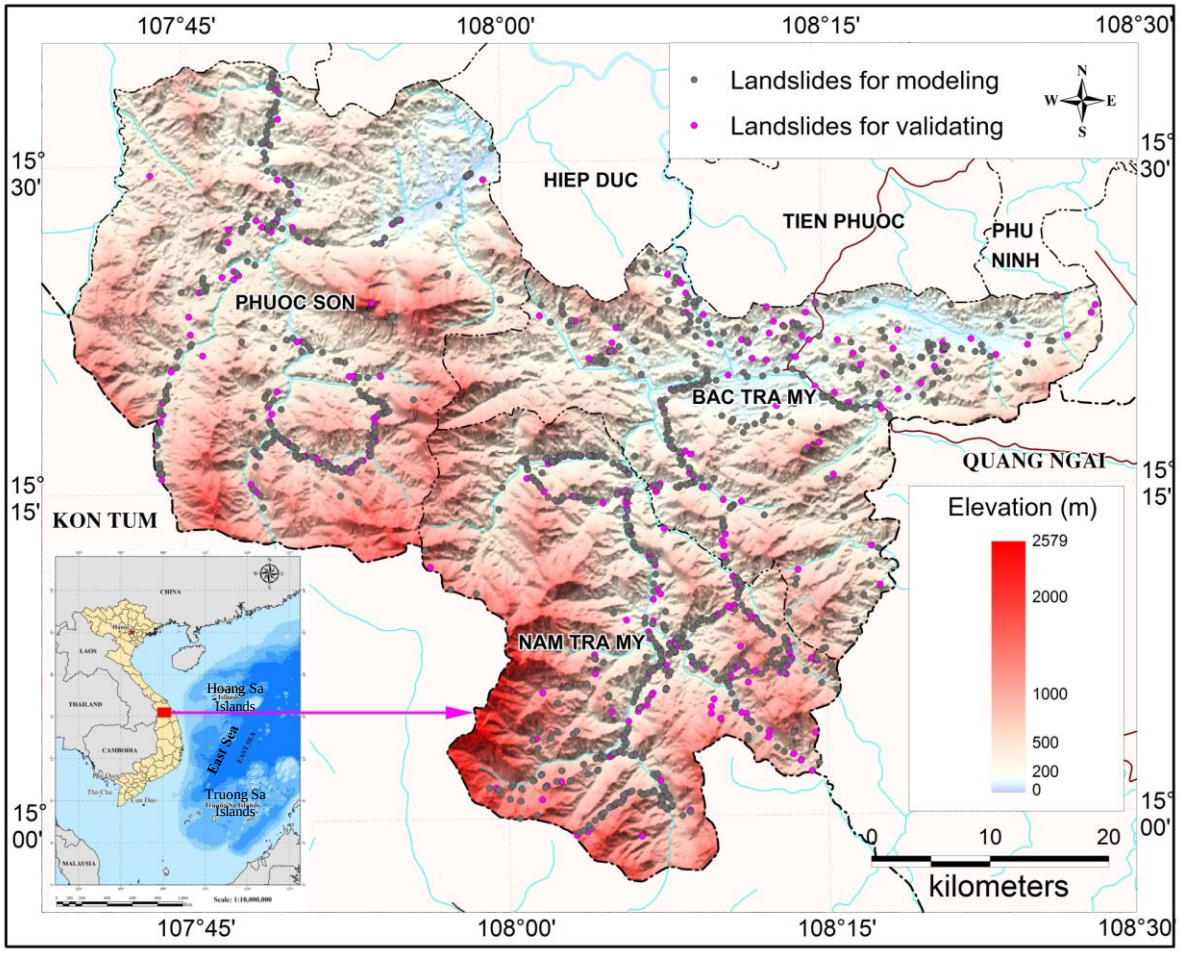


Figure 1. Study area and landslide locations.

The above-mentioned natural conditions have strongly influenced the common occurrence of landslides. Besides, economic activities such as the development of roads, the operation of terraced hydropower reservoirs, infrastructure construction, residential development planning, rudimentary and backward agricultural practices on sloping land, and burning forests have directed to negatively affected land use, forest cover that caused landslides in the study area.

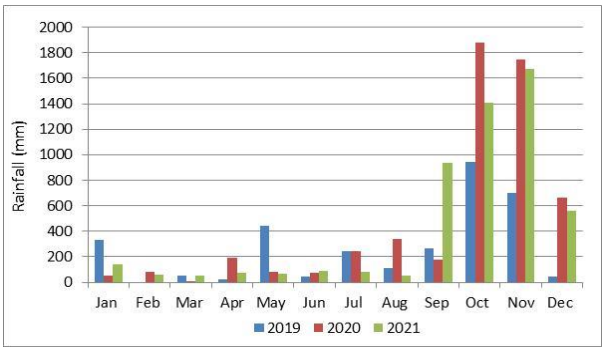


Figure 2. Distribution of total monthly rainfall at Tra My station for 2019, 2020 and 2021.

2.1.2. Data used

The landslide inventory data was one of the necessary data inputs for the landslide susceptibility models. In this study, 1159 landslide locations (Figure 1) were collected from three principal sources: Field survey data from March to April 2021 (414) under the project code VAST05.03/21-22; additional analytical data from Google Earth satellite imagery

(250) using the visual interpretation method; and inherited data from previous studies related to the study area (495) [22–25]. The landslide data were randomly divided into two parts, of which 70% (811 landslides) were used for modeling and 30% (348 landslides) for validation. The field survey results show that landslides have widely occurred in the rainy seasons. In the rainy season of 2020, three typical landslides buried 18 houses and killed 32 people in the Nam Tra My and Phuoc Son districts on October 28, 2020 (Figure 3).

Figure 3. Some pictures of typical landslides: a) A landslide region in Phuoc Loc commune, Phuoc Son district (Google Earth image in October 2021); b) The landslide occurred in Tra Leng commune, Nam Tra My district; and c) The landslide occurred in Tra Van commune, Nam Tra My district. Photographs were taken in April 2021 by Tran Anh Tuan.

The landslide-related factors include natural factors and human activities. They are input-independent variables in landslide susceptibility models. Ten factors were selected and constructed into a GIS database from the primary data in the study area (Table 1).

Table 1. Data used in the study.

Factors	Data source	Scale
Geology	Geological and Mineral Resources maps	1:200.000
Distance to fault		
Elevation	National topographic maps	1:50.000
Relief amplitude		
Slope		
Aspect		
Rainfall	Precipitation data at rain gauges in the study area and nearby	1:50.000
Soil type	Soil map of Quang Nam province	1:50.000
Land use	Land use map of Quang Nam province	1:50.000
Distance to road	National topographic maps	1:50.000

The geological conditions that affect the landslide process are lithological composition and tectonic activity. Accordingly, two influencing factors were selected as geology showing stratigraphies with various types of lithology, and distances to faults. The

geological factor consists of fourteen geological units (Figure 4a), in which Kham Duc formation (PR_{2-3}-$Є_1$ kd) and Tac Po formation (PR_1 tp) are the two units with the most extensive distribution in the study area, 35.767% and 29.412%, respectively. The distance to faults was classified into five classes (Figure 4b) from the fault system. Based on national topographic maps, a digital elevation model (DEM) with a spatial resolution of 20x20m was generated. From the DEM, four geomorphometric factors were extracted, including elevation with ten classes (Figure 4c), relief amplitude with seven categories (Figure 4d), and slope and aspect were both classified into nine categories (Figures 4e, 4f).

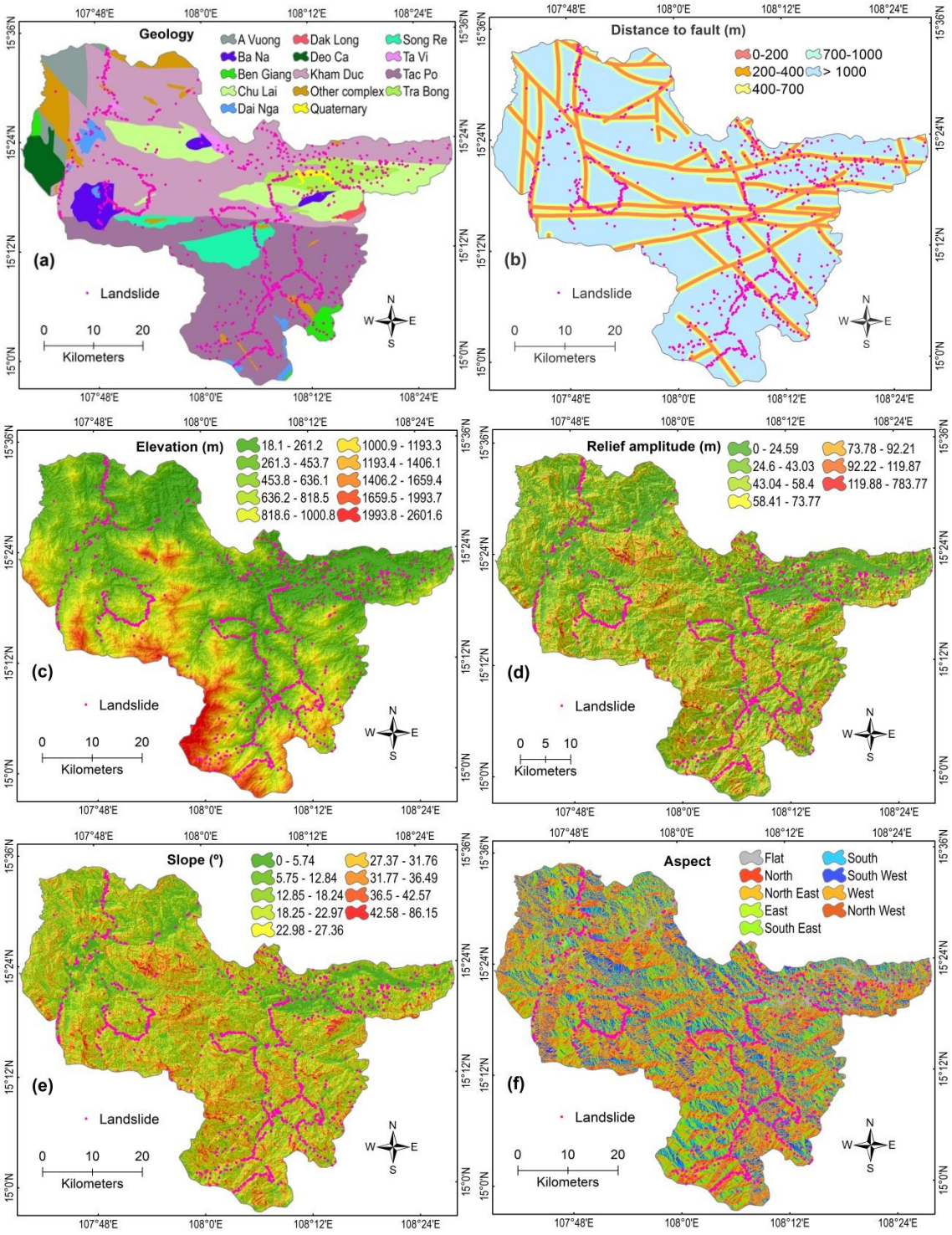


Figure 4. Landslide influencing factors: (a) Geology; (b) Distance to fault (m); (c) Elevation (m); (d) Relief amplitude (m); (e) Slope (o); (f) Aspect.

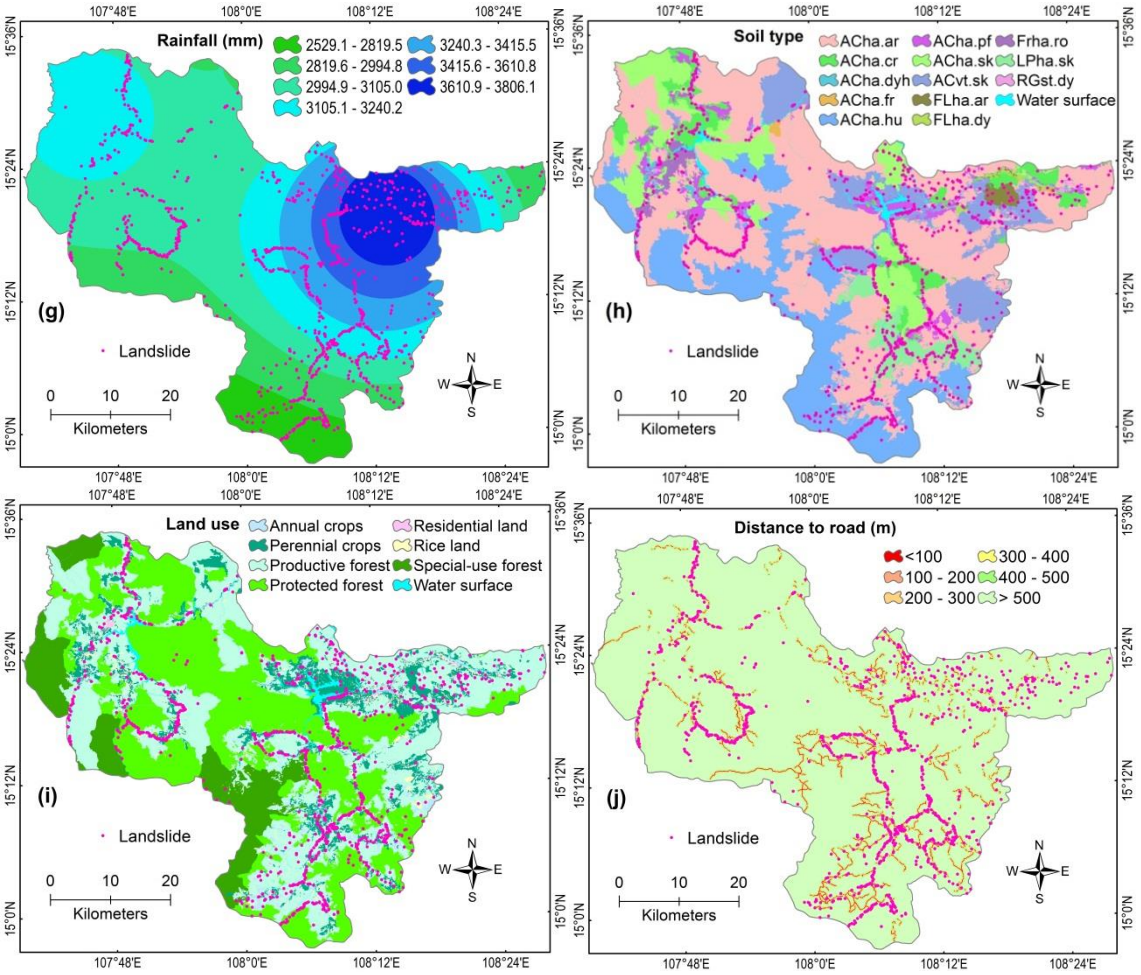


Figure 4. (continue): (g) rainfall (mm); h) Soil type; i) Land use; and j) Distance to road (m).

Precipitation is a landslide-triggering factor. The rainfall factor was constructed using total precipitation data from September to November 2020, collected from rain gauges in the study area and nearby. It was classified into seven classes (Figure 4g). The soil type map consists of 13 various soil types (excluding water surface) (Figure 4h). The Acha.ar is a common soil type, accounting for 43.446% of the total area, followed by ACha.hu (18.363%), ACvt.sk (10.815%), and ACha.sk (10.159%). The remaining soil types have small areas (less than 5% for each type).

Two factors related to human activities are land use and distance to roads. The land use factor was categorized into eight classes (Figure 4i). Forest types occupied most of the study area, of which productive forest and protected forest were the most extensive area, accounting for 37.458% and 35.281% of total area, respectively. Residential land had the smallest area, accounting for only 0.972%. The construction of the road system in mountainous areas has changed the slope leading to instability of the slope where the road crosses. The distance to road factor was constructed with five classes (Figure 4i) from a road system. However, only roads that cut through areas with slopes higher than five degrees were identified in this study, to avoid increasing the weight in flat areas where landslides rarely or not likely to occur.

2.2. Methods

2.2.1. Weights of Evidence

The WoE method was first developed for mapping the mineral potential [26]. Afterward, this method was applied in several studies in landslide susceptibility mapping

[12, 14, 27–30]. In the WoE method, a pair of weights for each class of each related factor, W^+ and W^-, is defined as:

$$W^+ = ln\frac{P\{X|L\}}{P\{X|\bar{L}\}} \quad (1) \qquad W^- = ln\frac{P\{\bar{X}|L\}}{P\{\bar{X}|\bar{L}\}} \quad (2)$$

where P is the probability, X is the class in a landslide-related factor, L signifies the existence of landslides, and the dash above ($\bar{}$) indicates the absence of variables under consideration. From there, the WoE can be written in numbers of pixels as follows:

$$W^+ = ln\frac{N\{X \cap L\}/N\{L\}}{N\{X \cap \bar{L}\}/N\{\bar{L}\}} \quad (3) \qquad W^- = ln\frac{N\{\bar{X} \cap L\}/N\{L\}}{N\{\bar{X} \cap \bar{L}\}/N\{\bar{L}\}} \quad (4)$$

where $N\{A\}$ is the pixel number of A present on a map and ∩ is intersection math.

Variances of weights (S^2) can approximately be estimated as [31]:

$$S^2(W^+) = \frac{1}{N\{X \cap L\}} + \frac{1}{N\{X \cap \bar{L}\}} \quad (5) \qquad S^2(W^-) = \frac{1}{N\{\bar{X} \cap L\}} + \frac{1}{N\{\bar{X} \cap \bar{L}\}} \quad (6)$$

The weight contrast value (C) presents the spatial relationship between the related factor and the occurrence of landslides [26]. C is positive for a favorable spatial relationship and negative for an unfavorable one [32]. The formula is as below:

$$\mathrm{C} = \mathrm{W}^+ - \mathrm{W}^- \quad (7)$$

The standard deviation of contrast ($S(C)$) is given by the formula:

$$\mathrm{S(C)} = \sqrt{\mathrm{S}^2(\mathrm{W}^+) + \mathrm{S}^2(\mathrm{W}^-)} \quad (8)$$

The standardized final weight (W) gives a measure of confidence [28]:

$$\mathrm{W} = \frac{\mathrm{C}}{\mathrm{S(C)}} \quad (9)$$

In the WoE model, the natural logarithm (ln) of zero is not defined, so the weight value under consideration is assigned the minimum weight value. Finally, the calculating of landslide susceptibility index (LSI) was done by overlaying the weighting values of related factors together in GIS using the equation below:

$$LSI = \sum_{i=1}^{n} W_{ij} \quad (10)$$

where W_{ij} is the final weight of the j-th class in the factor X_i (i = 1,2,…, n, n is the number of related factors).

2.2.2. Model verification

The significance of the statistical model in the susceptibility analysis was confirmed by testing the model using the validating dataset. One of the popular methods is the ROC curve analysis to verify the performance and compare different models [17]. The ROC curve is generated by representing the true positive rate (on the Y-axis) based on the false positive rate (on the X-axis) at the threshold varies. The AUC is the area underneath the entire ROC curve and was used to estimate the accuracy of a model. The AUC value ranges from 0.5 to 1.0. When AUC is 0.5, it means the model has no sense in terms of application. An AUC value close to 1 would indicate a high-performance prediction model.

3. Results and discussions

3.1. Relationship between landslide occurrence and related factors

In this study, the multicollinearity test method is performed to confirm that the input variables are independent of each other. The tolerance (TOL) index and the variance inflation factor (VIF) were used for multicollinearity checking. TOL is $1\text{-}R^2$ for the regression of one independent variable against all the other independents, while VIF is the reciprocal of TOL. The multicollinearity almost certainly occurs when TOL values are less than 0.1 or VIF values exceed 10 [33]. The results of the multicollinearity test show that the relief amplitude factor presents the minimum TOL and maximum VIF values are 0.435 and 2.297, respectively (Table 2). It confirms that no collinearity was observed among the ten selected variables.

Table 2. The result of the multicollinearity test.

Variables	Collinearity Statistics	
	Tolerance	**VIF**
Geology	0.708	1.413
Distance to fault	0.805	1.242
Elevation	0.633	1.580
Relief amplitude	0.435	2.297
Slope	0.441	2.266
Aspect	0.982	1.019
Rainfall	0.809	1.236
Soil type	0.645	1.551
Land use	0.679	1.473
Distance to road	0.792	1.263

The weighted values calculated using the WoE model indicated the importance of each influencing factor (Table 3). The higher values of the final weight showed a higher significance level for a factor class [27]. For the group of geological conditions factors, the highest weight value of 4.562 belongs to Tac Po formation (PR_1 *tp*), followed by Tra Bong Complex (O-S *tb*) has a weight value of 3.280. These two geological units indicate a significant positive correlation with the landslides. Some geological units with positive weight values but smaller than the two above-mentioned geological units are Ba Na Complex (K- E *bn*), Kham Duc Formation (PR_{2-3}-$Є_1$ *kd*), Undivided Quaternary (edQ) and Ta Vi Complex (PR_3 *tv*). The remaining units are all negatively weighted indicating a weak association for landslides. For the distance to fault, the weight value is inversely proportional to the distance to the fault. Within the distance < 1000 m from the fault location, all positive weights are given, in which the highest weight value belongs to the < 200m class (3.667), followed by the 200-400 m class (3.277). Meanwhile, the distance >1000m gives a negative weight value (-5.434).

The weighted value of the elevation factor shows the frequency of landslides is pretty high in the altitude range of < 453.7m, in which the weights for the 18.1-261.2 m and 261.3-453.7 m classes are 8.276 and 6.372, respectively. The negative weight value for the elevation > 453.7 m indicates low landslide frequency. In the case of relief amplitude, the final weight is positive in the range between 43.04 m and 119.87, with the highest weight value being 3.945 obtained at the amplitude values from 58.41 m to 73.77 m. The negative weight values are for the relief amplitude < 43.04 m or > 119.87 m. For the slope factor, the weight value increases gradually for the slope between 0^o and 36.49^o and decreases when the slope value is higher than 36.49^o. Positive weights are obtained from slopes $>18.24^o$. The maximum value is 5.64^o in the 31.77^o - 36.49^o class, while slopes from 0^o to 18.24^o give negative weight values. The results of the weight calculation for the aspect factor show that there is no significant difference. The correlation between aspect and landslide occurrence is positive for the aspect in North East, East, South East, South, and South West directions, while the North, West, and North West directions all give negative weights. The Flat class does not have landslides, so the weight is indeterminate. Thus, it gets the minimum value calculated from all factor classes (-14.646).

For the rainfall factor, the weights are positive for classes with rainfall > 3105.0 mm and negative values for the rainfall classes ≤ 3105.0 mm, except for the smallest rainfall range (2529.1-2819.5 mm), the weights obtained are positive and higher than the weights of other classes (4.037). Perhaps this class is distributed at the edge of the study area with a small area. In the case of soil type, positive weights appear in soil units such as ACha.cr, ACha.fr, ACha.pf, ACvt.sk, FLha.ar, Frha.ro, and Lpha.sk and negative for remaining soil types. Notably, the high landslide probability occurs in the soil types of ACha.cr (5.101), ACha.pf (4.742), and ACvt.sk (8.007). Lpha.sk is the soil unit with the highest weight (1.803). The water surface class was assigned the minimum weight (-14.646) because no landslides occurred.

Assessment of land use showed high weights for annual crops (3.425), productive forest (4.567), residential land (9.694), and the highest for perennial crops (13.618). Meanwhile, low (negative) weight values were protected forest (-9.162), special-use forest (-7.018), and rice land (-2.402). The water surface class has a weighted value of -14.646. For the distance to the road, the weight value decreases as the distance are further away from the position of the road. The weighted values are positive within a distance < 200 m, and negative for classes with a distance > 200 m. Notably, the maximum landslide probability with calculated weight is 26.849 within the distance < 100 m from the position of the road, and the minimum value (-14.646) when the distance is > 500 m. All influencing factors were stored in raster format (4367 columns, 3456 rows) with a pixel size of 20x20 meters.

Table 3. Computed weights for classes of influencing factors using the WoE model.

Classes of influencing factors	Nº. of landslide pixels	% of landslide	Nº. of pixels in class	% of class	Weights of Evidence (WoE)						
					W_{ij}^+	W_{ij}^-	C	$S^2(W^+)$	$S^2(W^-)$	S(C)	W_{ij}
Geology											
A Vuong	3	0.370	223212	3.146	-2.141	0.028	-2.169	0.333	0.001	0.578	-3.750
Ba Na	27	3.329	206784	2.915	0.133	-0.004	0.137	0.037	0.001	0.196	0.702
Ben Giang	5	0.617	96947	1.366	-0.796	0.008	-0.803	0.200	0.001	0.449	-1.791
Chu Lai	67	8.261	744722	10.497	-0.239	0.025	-0.264	0.015	0.001	0.128	-2.071
Dai Nga	7	0.863	99041	1.396	-0.481	0.005	-0.486	0.143	0.001	0.380	-1.281
Dak Long	1	0.123	22295	0.314	-0.936	0.002	-0.937	1.000	0.001	1.001	-0.937
Deo Ca	1	0.123	164373	2.317	-2.966	0.022	-2.988	1.000	0.001	1.001	-2.986
Kham Duc	301	37.115	2537624	35.767	0.037	-0.021	0.058	0.003	0.002	0.073	0.801
Other complex	27	3.329	308289	4.345	-0.266	0.011	-0.277	0.037	0.001	0.196	-1.415
Quaternary	6	0.740	27954	0.394	0.630	-0.003	0.634	0.167	0.001	0.410	1.546
Song Re	17	2.096	291974	4.115	-0.675	0.021	-0.695	0.059	0.001	0.245	-2.837
Ta Vi	7	0.863	48365	0.682	0.236	-0.002	0.238	0.143	0.001	0.380	0.627
Tac Po	298	36.745	2086743	29.412	0.223	-0.110	0.332	0.003	0.002	0.073	4.562
Tra Bong	44	5.425	236632	3.335	0.487	-0.022	0.508	0.023	0.001	0.155	3.280
Distance to fault (m)											
0-200	111	13.687	697477	9.831	0.331	-0.044	0.375	0.009	0.001	0.102	3.667
200-400	103	12.700	662378	9.336	0.308	-0.038	0.346	0.010	0.001	0.105	3.277
400-700	120	14.797	913547	12.876	0.139	-0.022	0.161	0.008	0.001	0.099	1.631
700-1000	94	11.591	796158	11.221	0.032	-0.004	0.037	0.011	0.001	0.110	0.333
>1000	383	47.226	4025395	56.736	-0.183	0.199	-0.382	0.003	0.002	0.070	-5.434
Elevation (m)											
18.1 - 261.2	227	27.990	1199533	16.907	0.504	-0.143	0.647	0.004	0.002	0.078	8.276
261.3 - 453.7	253	31.196	1550869	21.859	0.356	-0.127	0.483	0.004	0.002	0.076	6.372
453.8 - 636.1	128	15.783	1271616	17.923	-0.127	0.026	-0.153	0.008	0.001	0.096	-1.587
636.2 - 818.5	95	11.714	997247	14.056	-0.182	0.027	-0.209	0.011	0.001	0.109	-1.915
818.6 - 1000.8	38	4.686	744981	10.500	-0.807	0.063	-0.870	0.026	0.001	0.166	-5.235
1000.9 - 1193.3	37	4.562	534790	7.538	-0.502	0.032	-0.534	0.027	0.001	0.168	-3.172
1193.4 - 1406.1	20	2.466	359063	5.061	-0.719	0.027	-0.746	0.050	0.001	0.226	-3.294
1406.2 - 1659.4	10	1.233	247113	3.483	-1.038	0.023	-1.062	0.100	0.001	0.318	-3.336
1659.5 - 1993.7	2	0.247	127751	1.801	-1.988	0.016	-2.004	0.500	0.001	0.708	-2.830
1993.8 - 2601.6	1	0.123	61992	0.874	-1.958	-0.001	-1.957	1.000	0.001	1.001	-1.956
Relief amplitude (m)											
0 - 24.59	73	9.001	1133248	15.973	-0.574	0.080	-0.653	0.014	0.001	0.123	-5.324
24.6 - 43.03	195	24.044	1928529	27.182	-0.123	0.042	-0.165	0.005	0.002	0.082	-2.006
43.04 - 58.4	220	27.127	1795134	25.302	0.070	-0.025	0.094	0.005	0.002	0.079	1.195
58.41 - 73.77	187	23.058	1258840	17.743	0.262	-0.067	0.329	0.005	0.002	0.083	3.945
73.78 - 92.21	94	11.591	692509	9.761	0.172	-0.020	0.192	0.011	0.001	0.110	1.753
92.22 - 119.87	38	4.686	249199	3.512	0.288	-0.012	0.300	0.026	0.001	0.166	1.808
119.88 - 783.77	4	0.493	37496	0.528	-0.069	0.000	-0.069	0.250	0.001	0.501	-0.139
Slope (º)											
0-5.74	44	5.425	810704	11.426	-0.745	0.066	-0.810	0.023	0.001	0.155	-5.228
5.75-12.84	57	7.028	896153	12.631	-0.586	0.062	-0.648	0.018	0.001	0.137	-4.720
12.85-18.24	106	13.070	1203263	16.959	-0.261	0.046	-0.306	0.009	0.001	0.104	-2.940

Classes of influencing factors	Nº. of landslide pixels	% of landslide	Nº. of pixels in class	% of class	Weights of Evidence (WoE)						
					W_{ij}^+	W_{ij}^-	C	$S^2(W^+)$	$S^2(W^-)$	S(C)	W_{ij}
18.25-22.97	148	18.249	1209946	17.054	0.068	-0.015	0.082	0.007	0.002	0.091	0.905
22.98-27.36	138	17.016	1055878	14.882	0.134	-0.025	0.159	0.007	0.001	0.093	1.706
27.37-31.76	133	16.400	852734	12.019	0.311	-0.051	0.362	0.008	0.001	0.095	3.816
31.77-36.49	115	14.180	607596	8.564	0.504	-0.063	0.568	0.009	0.001	0.101	5.640
36.50-42.57	55	6.782	353686	4.985	0.308	-0.019	0.327	0.018	0.001	0.140	2.341
42.58-86.15	15	1.850	104995	1.480	0.223	-0.004	0.227	0.067	0.001	0.261	0.870
Aspect											
Flat	0	0.000	68966	0.972	NA	0.010	NA	NA	0.001	NA	-14.646
Noth	101	12.454	985860	13.895	-0.110	0.017	-0.126	0.010	0.001	0.106	-1.186
North East	133	16.400	1039771	14.655	0.112	-0.021	0.133	0.008	0.001	0.095	1.404
East	116	14.303	892796	12.584	0.128	-0.020	0.148	0.009	0.001	0.100	1.475
South East	94	11.591	803074	11.319	0.024	-0.003	0.027	0.011	0.001	0.110	0.244
South	109	13.440	831475	11.719	0.137	-0.020	0.157	0.009	0.001	0.103	1.522
South West	101	12.454	865449	12.198	0.021	-0.003	0.024	0.010	0.001	0.106	0.222
West	77	9.494	808365	11.394	-0.182	0.021	-0.204	0.013	0.001	0.120	-1.699
North West	80	9.864	799199	11.264	-0.133	0.016	-0.148	0.013	0.001	0.118	-1.260
Rainfall (mm)											
2529.1 - 2819.5	75	9.248	417156	5.880	0.453	-0.036	0.489	0.013	0.001	0.121	4.037
2819.6 - 2994.8	62	7.645	880229	12.406	-0.484	0.053	-0.537	0.016	0.001	0.132	-4.065
2994.9 - 3105.0	204	25.154	2444094	34.448	-0.314	0.133	-0.447	0.005	0.002	0.081	-5.524
3105.1 - 3240.2	204	25.154	1612013	22.721	0.102	-0.032	0.134	0.005	0.002	0.081	1.653
3240.3 - 3415.5	107	13.194	710273	10.011	0.276	-0.036	0.312	0.009	0.001	0.104	3.008
3415.6 - 3610.8	87	10.727	530396	7.476	0.361	-0.036	0.397	0.011	0.001	0.113	3.498
3610.9 - 3806.1	72	8.878	500794	7.058	0.229	-0.020	0.249	0.014	0.001	0.123	2.018
Soil type											
ACha.ar	238	29.346	3082502	43.446	-0.392	0.223	-0.615	0.004	0.002	0.077	-7.975
ACha.cr	73	9.001	356473	5.024	0.583	-0.043	0.626	0.014	0.001	0.123	5.101
ACha.dyh	3	0.370	39701	0.560	-0.414	0.002	-0.416	0.333	0.001	0.578	-0.719
ACha.fr	4	0.493	20119	0.284	0.554	-0.002	0.556	0.250	0.001	0.501	1.109
ACha.hu	87	10.727	1302826	18.363	-0.538	0.089	-0.627	0.011	0.001	0.113	-5.525
ACha.pf	55	6.782	256568	3.616	0.629	-0.033	0.662	0.018	0.001	0.140	4.742
ACha.sk	51	6.289	720747	10.159	-0.480	0.042	-0.522	0.020	0.001	0.145	-3.607
ACvt.sk	160	19.729	767292	10.815	0.601	-0.105	0.707	0.006	0.002	0.088	8.007
FLha.ar	10	1.233	50502	0.712	0.550	-0.005	0.555	0.100	0.001	0.318	1.743
FLha.dy	1	0.123	26005	0.367	-1.089	0.002	-1.092	1.000	0.001	1.001	-1.091
Frha.ro	19	2.343	74645	1.052	0.801	-0.013	0.814	0.053	0.001	0.232	3.505
LPha.sk	108	13.317	252436	3.558	1.320	-0.107	1.427	0.009	0.001	0.103	13.803
RGst.dy	2	0.247	36446	0.514	-0.734	0.003	-0.737	0.500	0.001	0.708	-1.040
Water surface	0	0.000	108693	1.532	NA	0.015	NA	NA	0.001	NA	-14.646
Land use											
Annual crops	37	4.562	185598	2.616	0.556	-0.020	0.576	0.027	0.001	0.168	3.425
Perennial crops	198	24.414	680688	9.594	0.934	-0.179	1.113	0.005	0.002	0.082	13.618
Productive forest	367	45.253	2657640	37.458	0.189	-0.133	0.322	0.003	0.002	0.071	4.567
Protected forest	158	19.482	2503186	35.281	-0.594	0.218	-0.812	0.006	0.002	0.089	-9.162
Special-use forest	8	0.986	763352	10.759	-2.390	0.104	-2.493	0.125	0.001	0.355	-7.018
Residential land	38	4.686	68988	0.972	1.573	-0.038	1.611	0.026	0.001	0.166	9.694
Rice land	5	0.617	126961	1.789	-1.066	0.012	-1.078	0.200	0.001	0.449	-2.402
Water surface	0	0.000	108542	1.530	NA	0.015	NA	NA	0.001	NA	-14.646
Distance to road (m)											
0-100	300	36.991	545541	7.689	1.571	-0.382	1.953	0.003	0.002	0.073	26.849
100-200	61	7.522	482483	6.800	0.101	-0.008	0.109	0.016	0.001	0.133	0.815
200-300	27	3.329	431201	6.078	-0.602	0.029	-0.631	0.037	0.001	0.196	-3.222
300-400	41	5.055	382875	5.396	-0.065	0.004	-0.069	0.024	0.001	0.160	-0.430
400-500	22	2.713	342878	4.833	-0.578	0.022	-0.600	0.045	0.001	0.216	-2.774
>500	360	44.390	4909977	69.204	-0.444	0.591	-1.035	0.003	0.002	0.071	-14.646

3.2. Validation of the model

The LSI was computed by summed final weights of factor classes, as shown in Eq. (10). Then, the WoE model performance was validated using the ROC curve and AUC (Figure 5). In this study, the landslides in the training dataset (811 landslides) were used to compute the success rate, and the landslides (348 landslides) in the test dataset to calculate the prediction rate. The results for the success rate indicated that an AUC of 0.855 showed the fitness of the WoE model because the training dataset was used in building the model. However, the success rate curve is not meaningful in evaluating the performance of prediction models [34]. Meanwhile, the prediction rate curve showed a value of 0.844, which proves well predictive power of the model.

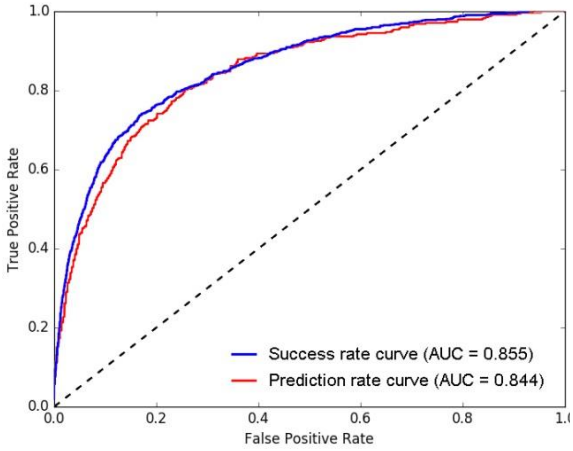


Figure 5. Model validation with success rate and prediction rate.

3.3. Landslide susceptibility mapping

The LSI of the study area ranges from -81.65 to 81.24. LSI values represent different levels of susceptibility to landslides [32]. The landslide susceptibility map (Figure 6) shows five susceptibility levels as very low with the LSI values between -81.65 and -34.38, low (-34.38 to -19.50), moderate (-19.50 to -2.44), high (-2.44 to 17.36) and very high (17.36 to 81.24) using Jenks Natural Breaks method. This method is based on the division into natural groups inherent in the data set such that the variance within each class and between classes are minimum and maximum, respectively [35].

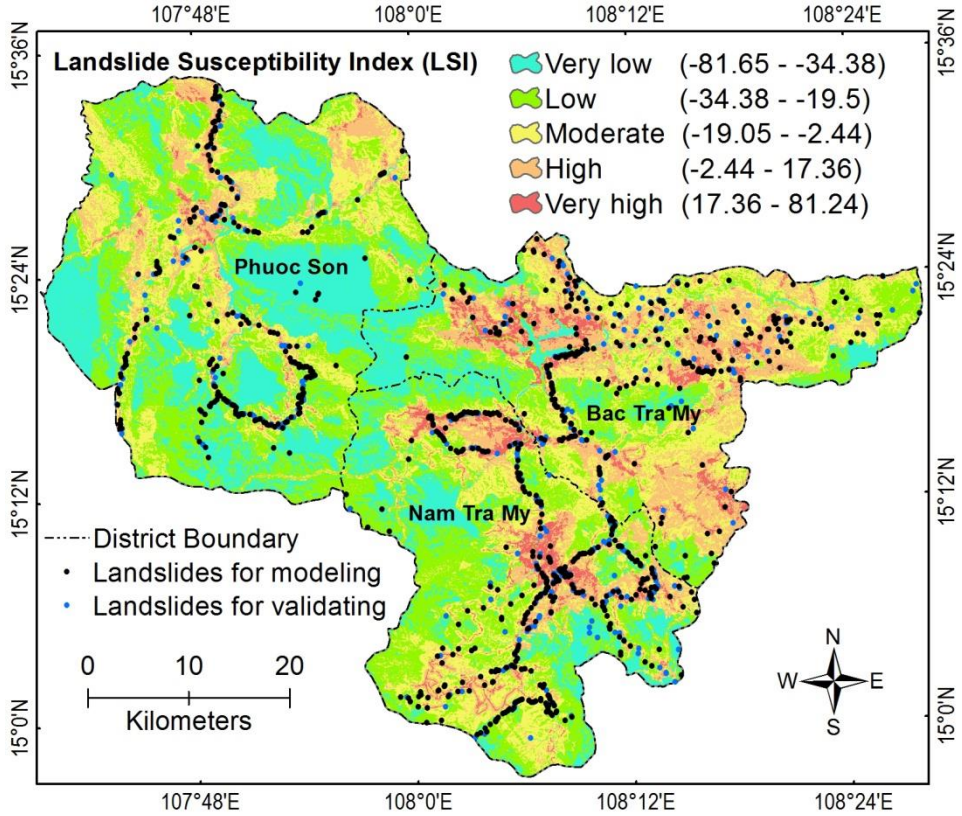


Figure 6. Landslide susceptibility map of the study area.

To determine landslides that occur in each susceptibility class, the landslide inventory and susceptibility index maps were overlaid together, then the percentage of landslides in each susceptibility level was calculated. The result of assessing the reliability is shown in Figure 7. Accordingly, the very high class with only 5.27% of the study area has 49.96 % landslide occurrence, while only 2.16% of the landslide location occurs in the very low class, with 22.13% area. The other susceptibility levels were high class with 31.76% of the total area, moderate class (24.46%), and low class (16.38%) with percentages of landslide occurrences of 6.90%, 16.74%, and 24.25%, respectively. This result once again confirms that the landslide susceptibility mapping using the WoE model is very reliable.

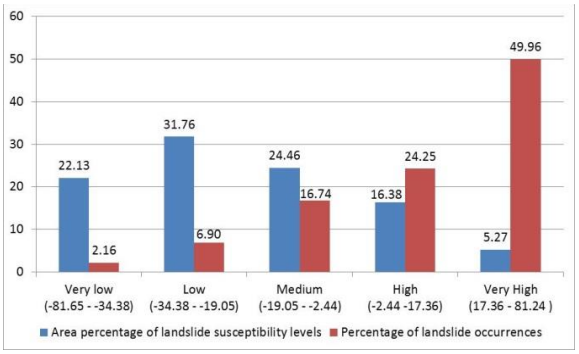


Figure 7. The results of assessing the reliability of the landslide susceptibility map.

3.4. Discussions

The review assessment of landslide susceptibility models shows that the group of bivariate statistical methods was only appropriate for research scales from medium (1:250.000–1:25.000) to detailed (>1:5000) and is not suitable for small scale (<1:250.000) [11]. In addition, bivariate statistical methods depended on the detail of inventory landslide data. According to the map generalization rule, the smaller the map scale, the lower the level of data detail, so it can be confirmed that the more the map scale decreases, the lower the model performance is. A landslide susceptibility mapping study for Quang Nam province using the WoE model [16] indicated that with a smaller map scale and a lower level of data detail, the model performance was significantly reduced to only 0.735 for success rate and 0.707 for prediction rate. Thus, the results of verifying the WoE model performance and assessing the reliability of the landslide susceptibility map in this study show that selecting a research method appropriate to the scale of the study area has delivered better research results.

4. Conclusions

Based on the research results, some conclusions are delineated as follows:

- Ten influencing factors were checked for multicollinearity showing independence from each other to landslide. All of these factors contributed to the landslide process in the study area, among which human activities were the intensely dominant factors such as land use and distance to road. Following were natural factors like slope and soil type.

- The results of testing the model performance using the ROC curve and AUC confirmed that the model capacity for landslide susceptibility mapping was very satisfactory. The AUC for the success rate curve and prediction rate curve were 0.855 and 0.844, respectively.

- The landslide susceptibility map showed 49.96 % of landslides occurred in the very high susceptibility class, 24.25% of landslides in the high class, the medium level contains 16.74% of landslides, the low and very low categories accounting for 6.90% and 2.16% of landslides, respectively. It is confirmed that the landslide susceptibility map generated from the Weights of Evidence model ensures reliability in land use planning and prevention and mitigation of landslide damage in the study area.

- Although the WoE method is quantitative, this method needs to be carefully selected to appropriate the research scale. Besides, due to the nature of a statistical method, the accuracy of the research results depends on the detailed landslide inventory data. The accuracy achieved from this study shows that the WoE method is suitable for landslide susceptibility mapping at a scale of 1:50.000.

Contributions of authors: Methodology, Investigation, Data analysis, Manuscript preparation, Manuscript editing: T.A.T.; Data analysis, Manuscript preparation: T.T.T.; Investigation, Data analysis: P.V.H.; Methodology, Data analysis: N.T.A.N. All authors have read and agreed to the published version of the manuscript.

Acknowledgement: This research was financially supported by the Vietnam Academy of Science and Technology (VAST), the grant project code is VAST05.03/21-22.

Declaration of competing interest: The authors declare that this article was the work of the authors, has not been published elsewhere, and has not been copied from previous research; there was no conflict of interest within the author group.

References

1. World Bank Group. Climate-resilient development in Vietnam: strategic directions for the World Bank. Washington, D.C, 2011. Available online: http://documents.worldbank.org/curated/en/348491468128389806/Climate-resilient-development-in-Vietnam-strategic-directions-for-the-World-Bank (accessed on 12 May 2023).
2. Ministry of Natural Resources and Environment of The Socialist Republic of Vietnam. National Disaster Risk in Viet Nam in the Period 2006-2016 and Forecasting and Warning System. 11th Emergency Preparedness Working Group Meeting Nha Trang, Viet Nam, 2017. Available online: https://www.apec-epwg.org/media/2309/f15e3a390421e8a5719bb2c859049604.pdf (accessed on 12 May 2023).
3. Meinhardt, M.; Fink, M.; Tünschel, H. Landslide susceptibility analysis in central Vietnam based on an incomplete landslide inventory: Comparison of a new method to calculate weighting factors by means of bivariate statistics. *Geomorphology* **2015**, *234*, 80-97. https://doi.org/10.1016/j.geomorph.2014.12.042.
4. Tien Bui, D.; Tuan, T.A.; Klempe, H.; Pradhan, B.; Revhaug, I. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* **2016**, *13*, 361–378. https://doi.org/10.1007/s10346-015-0557-6.
5. Tien Bui, D.; Tuan, T.A.; Hoang, N.D.; Thanh, N.Q.; Nguyen, D.B.; Van Liem, N.; Pradhan, B. Spatial prediction of rainfall-induced landslides for the Lao Cai area (Vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization. *Landslides* **2017**, *14*, 447–458. https://doi.org/10.1007/s10346-016-0711-9.
6. Tuan, T.A.; Dan, N.T. Research the landslide susceptibility and zoning in the Son La hydroelectricity area by the Saaty's Analytical Hiearchy Process (AHP). *J. Sci. Earth* **2012**, *3*, 223–232. https://doi.org/10.15625/0866-7187/34/3/2538.
7. Tuan, T.A.; Pha, P.D.; Tam, T.T.; Dui, D.T. A New Approach Based on Balancing Composite Motion Optimization and Deep Neural Networks for Spatial Prediction of Landslides at Tropical Cyclone Areas. *IEEE Access.* **2023**, *11*, 69495-69511. https://doi.org/10.1109/ACCESS.2023.3291411.
8. Provincial People's Committee of Quang Nam. The report of plans responds to natural hazards according to risk levels in the context of the COVID-19 epidemic in Quang Nam province (in Vietnamese), 2021. Available online: https://quangnam.gov.vn/webcenter/portal/ubnd/pages_tin-tuc/chi-tiet?dDocName=PORTAL174284 (accessed on 22 May 2023).
9. Merghadi, A.; Yunus, A.P.; Dou, J.; Whiteley, J.; Thai Pham, B.; Bui, D.T.; Avtar, R.; Abderrahmane, B. Machine learning methods for landslide susceptibility

studies: A comparative overview of algorithm performance. *Earth-Science Rev.* **2020**, *207*, 103225. https://doi.org/10.1016/j.earscirev.2020.103225.

10. Shano, L.; Raghuvanshi, T.K.; Meten, M. Landslide susceptibility evaluation and hazard zonation techniques–a review. *Geoenviron. Disasters* **2020**, *7(1)*, 1-19. https://doi.org/10.1186/s40677-020-00152-0.
11. Yong, C.; Jinlong, D.; Fei, G.; Bin, T.; Tao, Z.; Hao, F.; Li, W.; Qinghua, Z. Review of landslide susceptibility assessment based on knowledge mapping. *Stochastic Environ. Res. Risk Assess.* **2022**, *36*, 2399–2417. https://doi.org/10.1007/s00477-021-02165-z.
12. Saro, L.; Min, K. Statistical analysis of landslide susceptibility at Yongin, Korea. *Environ. Geology* **2001**, *40(9)*, 1095–1113. https://doi.org/10.1007/s002540100310.
13. Mind'je, R.; Li, L.; Nsengiyumva, J.B.; Mupenzi, C.; Nyesheja, E.M.; Kayumba, P. M.; Gasirabo, A.; Hakorimana, E. Landslide susceptibility and influencing factors analysis in Rwanda. *Environ. Dev. Sustainability* **2020**, *22(8)*, 7985–8012. https://doi.org/10.1007/s10668-019-00557-4.
14. Dahal, R.K.; Hasegawa, S.; Nonomura, A.; Yamanaka, M.; Dhakal, S.; Paudyal, P. Predictive modeling of rainfall-induced landslide hazard in the Lesser Himalaya of Nepal based on weights-of evidence. *Geomorphology* **2008**, *102(3–4)*, 496–510. https://doi.org/10.1016/j.geomorph.2008.05.041.
15. Pradhan, B.; Oh, H.J.; Buchroithner, M. Weights-of-evidence model applied to landslide susceptibility mapping in a tropical hilly area. Geomat. *Nat. Hazards Risk* **2010**, *1*, 199–223. https://doi.org/10.1080/19475705.2010.498151.
16. Thanh, D.C.; Binh, P.T.; Dam, N.D. Using weights of evidence (WOE) for landslide susceptibility mapping in Quang Nam province. *J. Sci. Technol. Civil Eng.* **2022**, *16(2V)*, 139–152. https://doi.org/10.31814/stce.huce(nuce)2022-16(2V)-12.
17. Bui, D.T.; Lofman, O.; Revhaug, I.; Dick, O. Landslide susceptibility analysis in the Hoa Binh province of Vietnam using statistical index and logistic regression. *Nat. Hazards* **2011**, *59*, 1413–1444. https://doi.org/10.1007/s11069-011-9844-2.
18. Schicker, R.; Moon, V. Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. *Geomorphology* **2012**, *161*, 40–57. https://doi.org/10.1016/j.geomorph.2012.03.036.
19. Ayalew, L.; Yamagishi, H. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* **2005**, *65(1-2)*, 15–31.
20. Baeza, C.; Corominas, J. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group* **2001**, *26(12)*, 1251–1263. https://doi.org/10.1002/esp.263.
21. Pham, B.T.; Prakash, I. Evaluation and comparison of LogitBoost Ensemble, Fisher's Linear Discriminant Analysis, logistic regression and support vector machines methods for landslide susceptibility mapping. *Geocarto Int.* **2019**, *34(3)*, 316–333. https://doi.org/10.1080/10106049.2017.1404141.
22. Duc, D.M.; Lieu, T.M.; Binh, T.Q.; Hang, V.T.; Van, H.P.; Vinh, H.D.; Tan, T.D.; Anh, G.Q.; Ngoc, D.M.; Duc, D.M. Landslide hazard prediction along the mountainous transport arteries in Quang Nam province and the adaptation measures (Vietnamese). Hanoi University of Science, Vietnam National University, Hanoi, Hanoi, Rep. ĐTĐL.CN-23/17, 2020.
23. Hung, L.Q.; Van, N.T.H.; Van, S.P.; Ninh, N.H.; Tam, N.; Huyen, N.T. Landslide inventory mapping in the fourteen Northern provinces of Vietnam: Achievements

and difficulties. In Advancing Culture of Living With Landslides, Sassa, K.; Mikoš, M.; Yin, Y. Eds.; Springer: Cham, Switzerland, 2017, pp. 501–510. https://doi.org/10.1007/978-3-319-59469-9_44.

24. Tan, M.T.; Van, H.V.; Tan, N.T.; Vinh, H.Q.; Van, L.N.; Luong, L.D.; Ha, T.T.; Van, T.N.; Thuy, H. L.Th.; Anh, L.T.; Van, T.T.T. Landslide hazard assessment by geological and geomorphological methods integrated with the GIS optimal weighting model in river basins in Thua Thien Hue, Quang Nam, and Da Nang areas, proposing solutions prevent (Vietnamese). *Inst. Geol. Sci. Vietnam Acad. Sci. Technol.* Hanoi, Vietnam, Rep. VAST 09.01/11-12, 2014, 2014.
25. Thanh, N. Q.; Yem, N. T.; Anh, T. T.; Phuong, N. T.; Cau, N. T.; Ngu, N. D.; Hieu, N. T.; Dai, H. Van; Thái, T. H.; Cong, N. T.; Minh, N. Le; Hoang, N. Van; Lien, V. T. H.; Tien, N. V.; Tuan, T. A.; Tai, N. T.; Kien, N. T.; Hung, N. Van; Thom, B. Van; Hau, D. T. To study, supplement and develop a map of natural disasters in Vietnam's mainland based on research results from 2000 up to now (Vietnamese). Inst. Geological Sci., Vietnam Acad. Sci. Technol., Hanoi, Vietnam, Rep. KC.08.28/11-15, 2015.
26. Bonham-Carter, G.F.; Agterberg, F.P.; Wright, D.F. Weights of evidence modeling: a new approach to mapping mineral potential. In *Statistical Applications in the Earth Sciences*, Agterberg, F.P., Bonham-Carter G.F., Eds.; Geol. Survey Canada Paper, 1989, 89-9, pp. 171–183.
27. Kayastha, P.; Dhital. MR.; De Smedt, F. Landslide susceptibility mapping using the weight of evidence method in the Tinau watershed, Nepal. *Nat. Hazards* **2012**, *63*, 479–498. https://doi.org/10.1007/s11069-012-0163-z.
28. Neuhäuser, B.; Terhorst, B. Landslide susceptibility assessment using "weights-of-evidence" applied to a study area at the Jurassic Escarpment (SW-Germany). *Geomorphology* **2007**, *86*, 12–24. https://doi.org/10.1016/j.geomorph.2006.08.002.
29. Polykretis, C.; Chalkias, C. Comparison and evaluation of landslide susceptibility maps obtained from weight of evidence, logistic regression, and artificial neural network models. *Nat. Hazards* **2018**, *93*, 249–274. https://doi.org/10.1007/s11069-018-3299-7.
30. van Westen, C.J.; Rengers, N.; Soeters, R. Use of Geomorphological Information in Indirect Landslide Susceptibility Assessment. *Nat. Hazards* **2003**, *30*, 399–419. https://doi.org/10.1023/B:NHAZ.0000007097.42735.9e.
31. Bonham-Carter, G.F. Geographic information systems for geoscientists, modeling with GIS, Pergamon, Press, Ontario, 1994, pp. 398.
32. Lee, S.; Choi, J.; Min, K. Landslide susceptibility analysis and verification using the Bayesian probability model. *Env. Geol*. **2002**, *43*, 120–131. https://doi.org/10.1007/s00254-002-0616-x.
33. Menard, S. Applied Logistic Regression Analysis (Sage University Paper Series on Quantitative Applications in the Social Sciences), series no. 106, 2nd ed.; ThousandOaks, CA: Sage, 1995.
34. Brenning, A. Spatial prediction models for landslide hazards: review, comparison and evaluation. *Nat. Hazards Earth Syst. Sci.* **2005**, *5(6)*, 853–862. https://doi.org/10.5194/nhess-5-853-2005.
35. Jenks, G.F. The data model concept in statistical mapping. *Int. Yearb Carto* **1967**, *7*, 186–190.

JOURNAL OF
HYDRO-METEOROLOGY

Research Article

Assessment of heavy metal pollution and ecological risk in the sediment of Cua Luc Bay, Quang Ninh Province

Huu Tuan Do[1]*

[1] Faculty of Environmental Sciences, VNU University of Science, Vietnam National University, Hanoi, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam; tuandh@vnu.edu.vn

*Correspondence: tuandh@vnu.edu.vn; Tel.: +84–2438584995

Received: 5 October 2023; Accepted: 12 November 2023; Published: 25 December 2023

Abstract: Heavy metals have big impacts on environmental quality and ecological risk. In this research, the I_{geo} accumulation index and ecological risk index (ERI) were used to assess the heavy metal pollution and ecological risk in the sediment of Cua Luc Bay, Quang Ninh Province. The results indicated that heavy metal concentrations in the study were almost lower than QCVN 43:2017/BTNMT. Only 3 monitored points, S1 (172.87 mg/kg), S3 (147.12 mg/kg), and S4 (115.54 mg/kg), have a light pollution level of Cu (108 mg/kg). The I_{geo} heavy metal accumulation index indicates that As (1.85), Cu (2.34), and Sb (1.42) are at a moderate level. While Se (3.65), Ag (3.58), and Sn (3.09) are at a high heavy metal level. As and Cu are two major factors posing ecological risk in the region (66.5%). However, the average ERI of all monitored points is 138, indicating that the ecological risk for heavy metals in the region is low. Despite having a low ecological risk for heavy metals in the region, it is necessary to strictly monitor heavy metals in sediment to manage them better in the future.

Keywords: Heavy metals; Ecological risk; Sediment; Heavy metals in sediment; Cua Luc Bay.

1. Introduction

Coastal bays play an important role in coastal ecosystems. With a typical mangrove forest ecosystem, the coastal bay is home to many species of marine animals that serve as food for humans. Coastal bays are also places where sediment is deposited and stored in mangrove forests. Therefore, sediments in coastal bays often contain heavy metals released from the mainland by rivers. The concern about heavy metals in sediments has been studied by many scientists for a long time. With their bioaccumulation in the food chain, heavy metals pose many potential risks to the ecosystem and human health. To assess heavy metal pollution in sediment, there are numerous methods such as Geo accumulation index, Ecological risk index, Enrichment Factor, Average pollution index, Pollution load index, Nemerow pollution index, Ecological risk factor [1]. In which, Geo accumulation index (I_{geo}), Ecological risk index (ERI) are more widely used [1]. There are numerous studies using Igeo to assess heavy metals in sediment such as in river sediment [2–5], urban sediment [6], coastal bays [7], reservoirs and dams [8, 9]. ERI is also an important method used to assess ecological risk for rivers [10], coastal [11], and wetland [12].

In Vietnam, heavy metals in sediment are also interesting research topics. The researches mainly focused on rivers such as rivers in Ho Chi Minh City [13], the Cau River [14], the Mekong River [15], the Nhue - Day River [16], lakes and dams [17]. However, research in coastal bays is still limited. Cua Luc Bay in Ha Long City is a place with an interaction between sea water flowing in from Ha Long Bay and river water from regional rivers such

as the Troi River and Dien Vong River. It also has mangrove forests that create an ideal environment for sediment deposition. Besides, the forest has sea foods as important sources of income for local people, especially species such as mangrove clams (Mud clam, Corrugate lucine), octopus, shrimp, crab, fish, etc. Therefore, assessing the level of heavy metal pollution and ecological risks in the region is important. The objectives of this study include assessing the level of heavy metal pollution and the ecological risk of heavy metals to the sedimentary environment in the area.

2. Methods

2.1. Samples collection

In the study, a suite of heavy metals was scrutinized, encompassing Arsenic (As), Selenium (Se), Chromium (Cr), Cobalt (Co), Nickel (Ni), Copper (Cu), Zinc (Zn), Molybdenum (Mo), Silver (Ag), Cadmium (Cd), Tin (Sn), Antimony (Sb), and Lead (Pb). Collection of sediment samples was executed in September 2023. Subsequent analytical procedures were conducted in strict adherence to the National Technical Regulation on Sediment Quality QCVN 43:2017/BTNMT.

Spatial delineation of sample provenance is elucidated in Table 1, while Figure 1 provides a graphical representation of the sampling sites, facilitating a comprehensive visual and quantitative assessment of the geographical distribution of the samples under investigation.

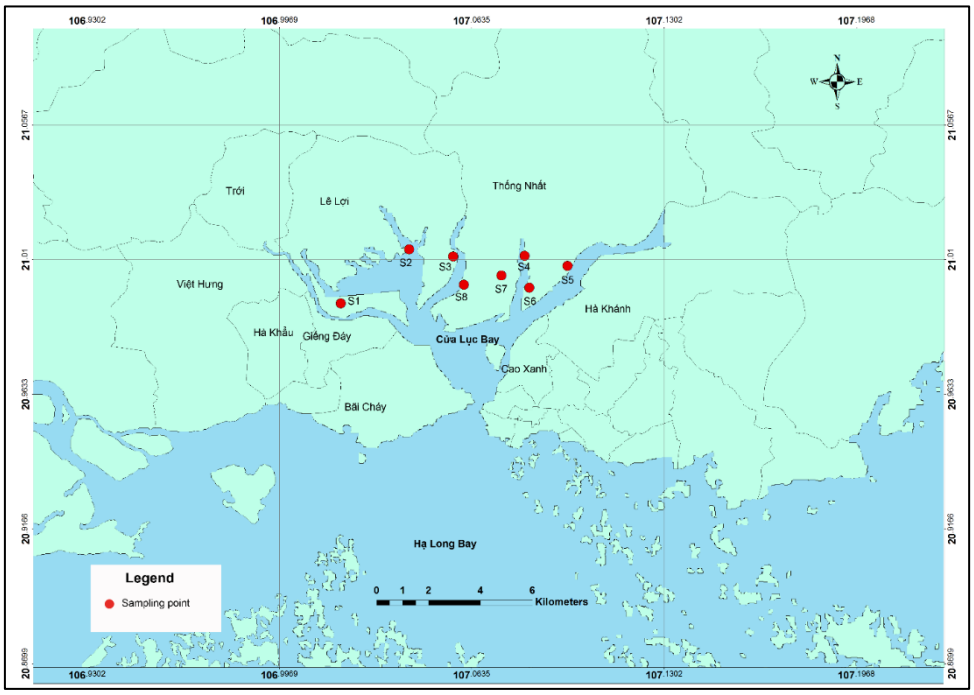


Figure 1. Location of sampling points.

Table 1. Coordinates of sediment sampling points.

Ord.	Sample Code	Longitude	Latitude
1.	S1	107.018116	20.994858
2.	S2	107.041819	21.013584
3.	S3	107.057094	21.011129
4.	S4	107.081856	21.011316
5.	S5	107.096709	21.007830
6.	S6	107.083500	21.000280
7.	S7	107.073800	21.004530
8.	S8	107.060800	21.001310

2.2. Methods

2.2.1. Ecological risk index (ERI)

To evaluate the potential ecological risk of heavy metal, ERI is used as an index method [18].

$$ERI = \sum RI = \sum Ti \times PI \quad (1)$$

where $PI = C_s/C_b$; PI is the pollution index, Cs is the observed metal concentration, Cb denotes the corresponding background values [19], Ti denotes the heavy metal's toxic-response factor, and RI stands for each heavy metal's potential ecological risk factor. The values of Cb and Ti were shown in the Table 2.

Table 2. Values of Cb and Ti [18–20].

Factor	As	Cr	Co	Ni	Cu	Zn	Cd	Pb
Cb (ppm)	2	35	11.6	18.6	14.3	52	0.102	17
Ti	10	2	5	5	5	1	30	5

After calculating, the ERI values were compared to the ERI levels: ERI < 150: low ecological risk; 150 ≤ ERI <300: moderate ecological risk; 300 ≤ ERI <600: considerable ecological risk; ERI ≥ 600 high ecological risk [18, 21]. Potential ecological risk factor for each metal: RI < 40: low risk; 40 ≤ RI < 80: moderate risk; 80 ≤ RI < 160 considerable; 160 ≤ RI < 320: high potential ecological risk; 320 ≤ RI: very high risk [21].

2.2.2. Geo Accumulation index (I_{geo})

Heavy metal pollution in soil and water sediments can be evaluated using the geo-accumulation index by [22]. The base two logarithms of the background metal concentration and the total metal concentration can be multiplied to find I_{geo} using the equation below.

$$I_{geo} = \log_2 \frac{C_n}{1.5\ B_n} \quad (2)$$

where C_n is observed heavy metal concentration of samples; B_n is geo-chemical background concentration of heavy metals [19]; I_{geo} = 0-1: Low; I_{geo} = 1-3: Moderate; I_{geo} = 3-5: High [1].

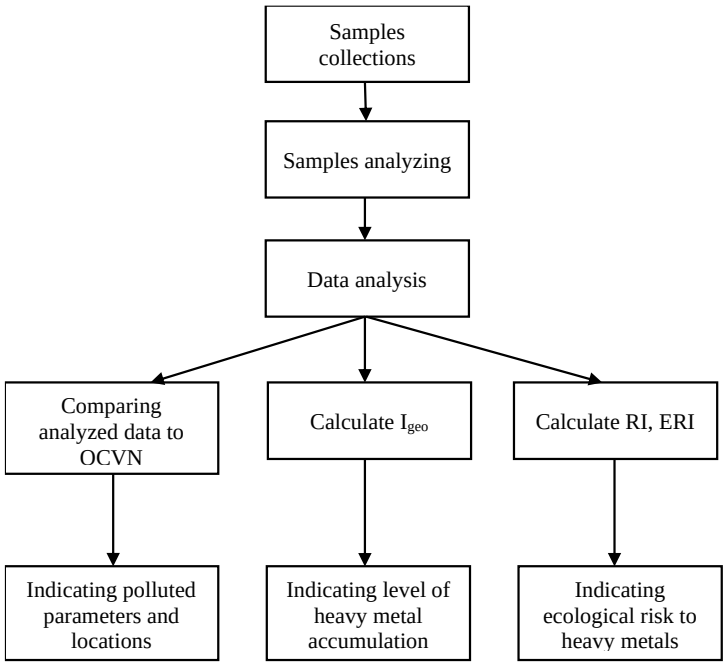


Figure 2. Research flowchart.

2.3. Statistical method

After analyzing samples, the results are subjected to basic and in-depth statistical processing using R statistical software (Figure 2).

3. Results and Discussion

3.1. Heavy metals in sediment

Analytical data presented in Table 3 indicate that arsenic (As) concentrations in the studied samples span a range from 8.38 to 11.73 mg/kg, mean concentration of 10.79 mg/kg. Notably, these concentrations remain below the permissible limit set by QCVN

43:2017/BTNMT, which is 41.6 mg/kg. Selenium (Se) exhibits a mean concentration of 1.57 mg/kg, with the observed concentrations varying between a minimum of 0.85 mg/kg and a maximum of 2.34 mg/kg. Chromium (Cr) concentrations in the environmental matrix exhibit a peak concentration of 28.05 mg/kg at sampling location S3, contrasting with the nadir of 10.02 mg/kg observed at point S4. All measured Cr concentrations were beneath the regulatory threshold of 160 mg/kg as prescribed by QCVN 43:2017/BTNMT.

Concentrations of cobalt (Co) and nickel (Ni) extend from 2.74 to 23.12 mg/kg and 9.17 to 45.96 mg/kg, respectively. The apex of Co and Ni concentrations was detected at sampling site S5, documenting concentrations of 23.12 mg/kg for Co and 45.96 mg/kg for Ni. Point S5's heightened metal concentrations are attributed to the influence of contaminants from the upstream sector of the Dien Vong River, an area that has been historically impacted by extensive coal mining activities.

Table 3. Desciptive statistics of heavy metals concentration (mg/kg).

Symbol	Mean	Standard Deviation	Minimum	Median	Maximum	QCVN 43:2017/BTNMT
As	10.79	1.17	8.38	11.39	11.73	41.6
Se	1.57	0.63	0.85	1.44	2.34	-
Cr	22.19	5.91	10.02	24.31	28.05	160
Co	8.70	7.20	2.74	5.87	23.12	-
Ni	20.41	13.40	9.17	15.01	45.96	-
Cu	108.27	46.35	38.96	119.03	172.87	108
Zn	62.26	11.32	48.78	61.16	83.06	271
Mo	2.24	0.81	0.96	2.40	3.08	-
Ag	0.99	0.72	0.23	0.73	2.19	-
Sn	32.04	8.26	19.84	34.47	44.31	-
Sb	1.24	0.34	0.85	1.20	1.66	-
Pb	17.42	4.06	9.59	16.99	22.35	112

In the context of this investigation, copper (Cu) has emerged as the most contaminated metal within the study area, with concentrations ranging from 38.96 mg/kg to an apex of 172.87 mg/kg. Notably, sampling points S1, S3, and S4 revealed Cu concentrations that exceed the regulatory standard of 108 mg/kg, with measured concentrations of 172.87 mg/kg, 147.12 mg/kg, and 115.54 mg/kg, respectively (Figure 3). These points of elevated Cu pollution are proximal to the industrial regions encompassing the communes of Gieng Day, Viet Hung, Le Loi, and Thong Nhat.

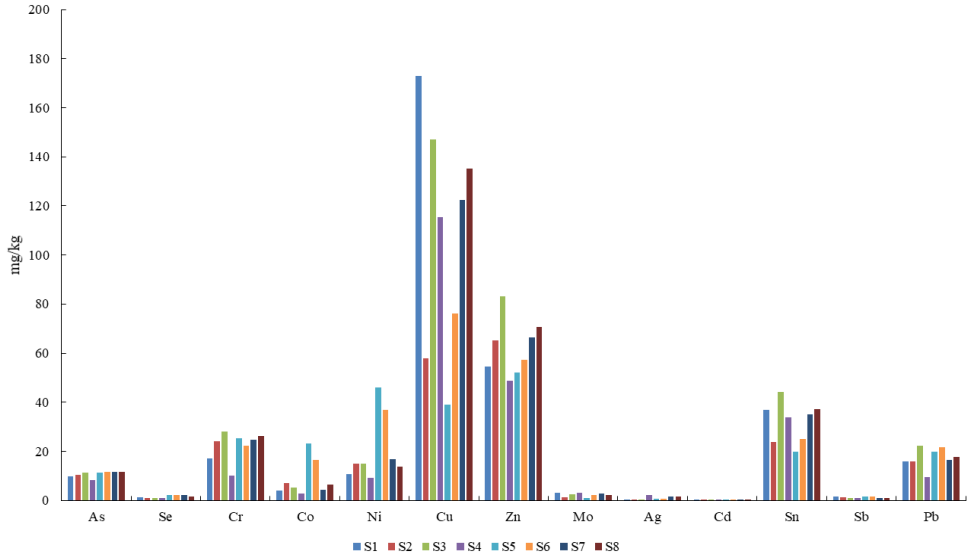


Figure 3. Concentration of heavy metals at sampling points in the study area.

Contrastingly, the concentrations of zinc (Zn) and lead (Pb) have been determined to be within acceptable limits. The average Zn concentration has been quantified at 62.26 mg/kg, which is considerably lower than the regulatory of 271 mg/kg as delineated by QCVN 43:2017/BTNMT. Similarly, the mean Pb concentration stands at 17.42 mg/kg, which is also below the prescribed standard of 172 mg/kg (Figure 4). This denotes a relatively lower level of pollution for these metals in the study areas.

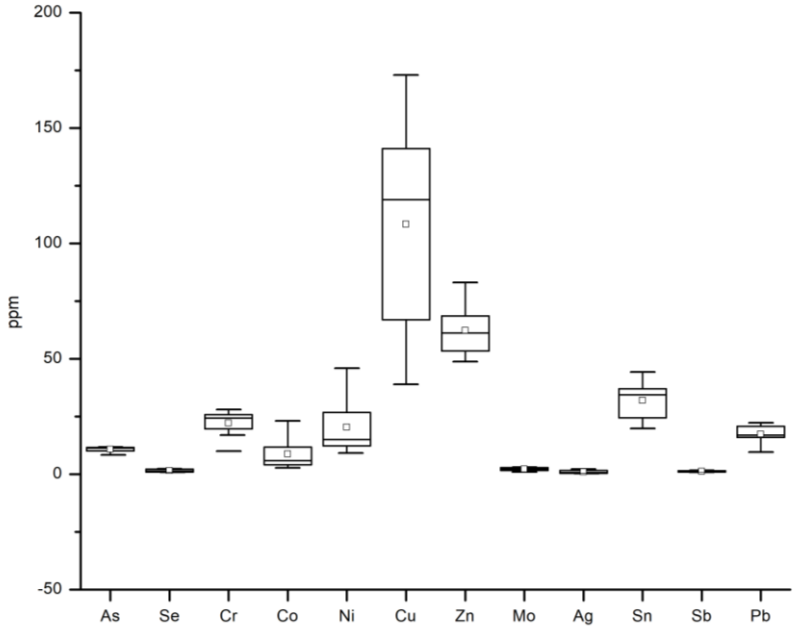


Figure 4. Concentration of heavy metals in sediment of the study area.

Comparing the monitored results with QCVN 43:2017/BTNMT reveals that most heavy metal parameters in the study area are below the regulated standards. Such findings suggest that despite the prevalence of industrial zones within the vicinity, the accumulation of heavy metals in the sediments of Cua Luc Bay remains at a low level.

3.2. Geo Accumulation index (I_{geo})

The geo-accumulation index (I_{geo}) was calculated to assess the pollution levels of various heavy metals in the study area, with the findings depicted in Figure 5. The I_{geo} values for Cr, Co, Ni, Zn, Mo, Cd, and Pb are indicative of low levels. Specifically, I_{geo} calculations for As spanned from 1.48 to 1.95, with an average of 1.85, positioning As within the moderate level category. Similarly, Cu presented I_{geo} values ranging from 0.86 to 3.01 with a mean value of 2.34, and Sb exhibited a range from 0.86 to 1.84 and an average of 1.42, both also classified as moderate levels.

Conversely, Se recorded higher I_{geo} values, from 2.77 to 4.23, with a mean of 3.65, indicating a high level of accumulation. Ag and Sn also fell into the high level category, with I_{geo} values for Ag ranging from 1.48 to 4.73 (mean 3.58) and for Sn ranging from 2.4 to 3.56 (mean 3.09). These elevated indices underscore a significant concern for the enrichment of these metals within the sedimentary environment of the study area.

Upon evaluating the concentrations of heavy metals within the scope of the regulatory QCVN 43:2017/BTNMT, it has been observed that only Cu exhibits a high I_{geo} value, indicating a considerable level of pollution. In contrast, other metals that are regulated under this standard, such as As, Cr, Zn, Cd, and Pb, are all characterized by low I_{geo} values, suggesting minimal geo-accumulation in the study area. However, for metals not explicitly regulated by QCVN 43:2017/BTNMT, such as Sb, Se, Ag, and Sn, there is a range from moderate to high I_{geo} values.

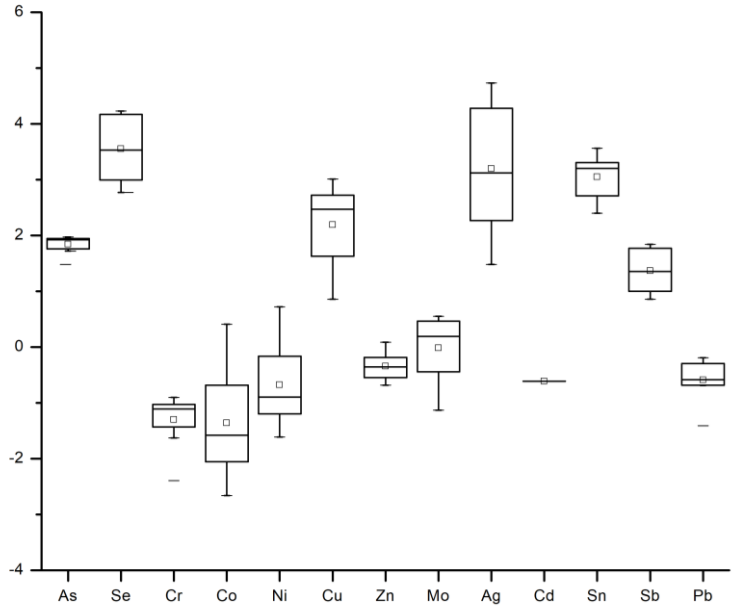


Figure 5. I_{geo} values of heavy metals.

3.3. Ecological risk of heavy metals

The potential ecological risk factor (RI), as tabulated in Table 4, indicates that arsenic (As) and copper (Cu) are the predominant contributors to the ecological risk in the region, together accounting for 66.5% of the total potential risk. The individual contributions of the heavy metals to the ecological risk of the region are: As constitutes 39% of the overall potential ecological risk, indicating it as the single most significant contributor. Cr represents a relatively minor portion, contributing 1% to the potential risk. Co accounts for 3% of the ecological risk, reflecting a lower level of concern in relation to As and Cu. Ni's contribution at 4%, to the risk. Cu is the second-largest contributor after As, responsible for 27% of the potential ecological risk. Zn with a contribution of 1%, Zn poses a relatively low risk under the current analysis. Cd often being present in lower concentrations, Cd represents a significant risk factor, contributing 21% due to its high toxicity. Pb contributes 4%, which is similar to Ni. These data underscore As and Cu as the primary metals of concern with respect to ecological risk within the study area. This suggests that remedial attention and mitigation strategies should prioritize the management of these metals to reduce potential ecological impacts.

Table 4. Heavy metal potential ecological risk factor (RI).

Symbol	S1	S2	S3	S4	S5	S6	S7	S8	Mean	Risk Factor
As	49.3	52.3	56.7	41.9	57.1	58.1	57.6	58.7	54.0	Moderate risk
Cr	1.0	1.4	1.6	0.6	1.4	1.3	1.4	1.5	1.3	Low risk
Co	1.8	3.1	2.2	1.2	10.0	7.1	1.8	2.8	3.7	Low risk
Ni	2.9	4.0	4.0	2.5	12.4	9.9	4.5	3.7	5.5	Low risk
Cu	60.4	20.2	51.4	40.4	13.6	26.7	42.8	47.2	37.9	Moderate risk
Zn	1.1	1.3	1.6	0.9	1.0	1.1	1.3	1.4	1.2	Low risk
Cd	<29.4	<29.4	<29.4	<29.4	<29.4	<29.4	<29.4	<29.4	<29.4	Low risk
Pb	4.7	4.7	6.6	2.8	5.8	6.4	4.8	5.2	5.1	Low risk

The assessment of the potential ecological risk factor (RI) for arsenic (As) across all monitored locations indicates that there is a uniform moderate risk level, with the RI values oscillating between 49.3 and 58.7. In terms of copper (Cu), sampling points S1, S3, S4, S7, and S8 are classified under a moderate risk category, while the remaining points, namely S2,

S5, and S6, are identified as having a low potential risk. For the other metals evaluated, namely Cr, Co, Ni, Zn, Cd, and Pb, the potential ecological risk is categorized as low across the board.

When the risk levels for Cua Luc Bay are placed in a comparative context with other coastal bays in the countries near Vietnam, as detailed in Table 5, a notable pattern emerges: The elements Cr, Co, Ni, Zn, and Pb are at a low risk in Cua Luc Bay as well as in the other surveyed coastal bays. Cua Luc Bay is distinguished by its low risk for cadmium Cd, whereas the comparative bays exhibit a range of potential risks, stretching from moderate to very high level.

Table 5. Heavy metal potential ecological risk in some coastal bays.

Symbol	Cua Luc Bay	Hangzhou Bay, China [23]	Ulsan Bay, South Korea [24]	Daya Bay, China [25]	Bay of Bengal Coast, Bangladesh [26]
As	Moderate risk	Low risk	Moderate risk	Low risk	Low risk
Cr	Low risk	Low risk	Low risk	Low risk	Low risk
Co	Low risk	Low risk	Low risk	-	-
Ni	Low risk	Low risk	Low risk	Low risk	-
Cu	Moderate risk	Low risk	Moderate risk	Low risk	-
Zn	Low risk	Low risk	Low risk	Low risk	-
Cd	Low risk	Moderate risk	High risk	Moderate risk	Very high risk
Pb	Low risk	Low risk	Low risk	Low risk	Low risk

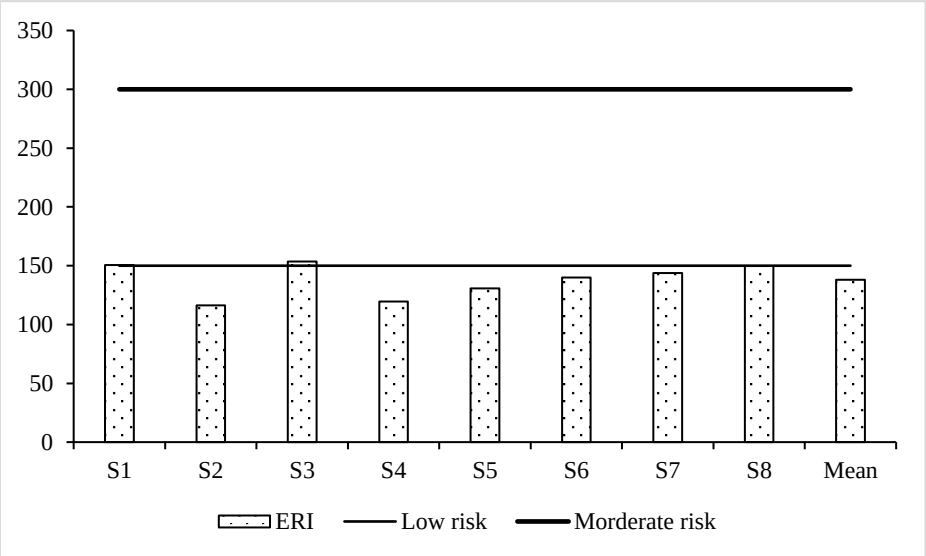


Figure 6. ERI values of sampling points.

The Ecological Risk Index (ERI) values delineated in Figure 6, spanning from 116.3 to 153.6, suggest variability in the potential ecological risk across different sampling points. Two specific points, S1 and S3, are classified as having moderate ecological risk, which is defined by an ERI between 150 and 300. All other points fall into the low ecological risk category, which indicates that their ERI values are below 150.

The mean ERI for all monitored points is calculated to be 138. This average positions the study area as a whole within the low ecological risk level for heavy metals. This assessment suggests that while there are specific areas of heightened concern (namely S1 and S3), the overall heavy metal contamination across the sampled locations does not pose a high ecological risk at this juncture. However, the areas with moderate risk warrant closer scrutiny and possible intervention to prevent any escalation in risk levels.

4. Conclusion

The study reveals that most heavy metal levels are below the threshold of QCVN 43:2017/BTNMT, with the exception of copper (Cu) at three points (S1, S3, and S4), which exceed the permissible levels. The Index of Geoaccumulation (I_{geo}) shows that As, Cu, and Sb have moderate accumulation levels, while selenium Se, Ag, and Sn are classified as having high accumulation levels. Arsenic and copper are the primary contributors to ecological risks in the area, accounting for 66.5% of the concern. However, the average Ecological Risk Index (ERI) for the area is 138, suggesting a low ecological risk from heavy metals overall.

Although the current ecological risk from heavy metals is low, the swift expansion of industrial areas could lead to greater risks in the future. It's crucial to maintain vigilant surveillance of heavy metals in the sediment. This is important to provide early warnings to local communities about the safety of consuming seafood from the mangrove forest.

The monitored samples for this research are limited. Deeper research with more samples is recommended to discover the potential risk of heavy metals in the region.

Author contribution statement: Constructing research idea: H.T.D.; Select research methods: H.T.D.; Data processing: H.T.D.; Sample analysis: H.T.D.; Take samples: H.T.D.; Writing original draft preparation: H.T.D.; Writing review and editing: H.T.D.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Prasad Ahirvar, B.; Das, P.; Srivastava, V.; Kumar, M. Perspectives of heavy metal pollution indices for soil, sediment, and water pollution evaluation: An insight. *Total Environ. Res. Themes*. **2023**, *6*, 100039.
2. Haris, H.; Looi, L.J.; Aris, A.Z.; Mokhtar, N.F.; Ayob, N.A.A.; Yusoff, F.M.; Salleh, A.B.; Praveena, S.M. Geo-accumulation index and contamination factors of heavy metals (Zn and Pb) in urban river sediment. *Environ. Geochem. Health* **2017**, *39(6)*, 1259–1271.
3. Adel Mashaan Rabee, Y.F.A.F.; Abd-Al-Husain, N.A.; Nameer, M. Using Pollution Load Index (PLI) and Geoaccumulation Index (I-Geo) for the Assessment of Heavy Metals Pollution in Tigris River Sediment in Baghdad Region. *J. Al-Nahrain Univ*. **2011**, *14(4)*, 108–114.
4. Andem, A.B.; Okorafor.; Ama, K.; Oku.; Esien, E.; Ugwumba.; Alex, A. Evaluation and characterization of trace metals contamination in the surface sediment using pollution load index (pli) and geo-accumulation index (igeo) of Ona River, Western Nigeria. *Int. J. Sci. Technology Res.* **2015**, *4(1)*, 29–34.
5. Hanif, N.; Eqani, S.A.M.A.S.; Ali, S.M.; Cincinelli, A.; Ali, N.; Katsoyiannis, I.A.; Tanveer, Z.I.; Bokhari, H. Geo-accumulation and enrichment of trace metals in sediments and their associated risks in the Chenab River, Pakistan. *J. Geochem. Explor*. **2016**, *165*, 62–70.
6. Martínez, L.L.G.; Poleto, C. Assessment of diffuse pollution associated with metals in urban sediments using the geoaccumulation index (Igeo). *J. Soils Sediments*. **2014,** *14(7)*, 1251–1257.
7. Wardani, N.K.; Prartono, T.; Sulistiono, S. Sediments quality based on geo-accumulation index in heavy metals (Pb, Cu, and Cd) of Cengkok Coastal waters, Banten Bay. *Jurnal Pendidikan IPA Indonesia*. **2020**, *9(4)*, 574–582.
8. Ghrefat, H.A.; Abu-Rukah, Y.; Rosen, M.A. Application of geoaccumulation index and enrichment factor for assessing metal contamination in the sediments of Kafrain Dam, Jordan. *Environ. Monit. Assess.* **2011**, *178(1)*, 95–109.
9. Zahra, A.; Hashmi, M.Z.; Malik, R.N.; Ahmed, Z. Enrichment and geo-accumulation of heavy metals and risk assessment of sediments of the Kurang Nallah–Feeding

tributary of the Rawal Lake Reservoir, Pakistan. *Sci. Total Environ.* **2014**, *470-471*, 925–933.
10. Bao, K.; Liu, J.L.; You, X.G.; Shi, X.; Meng, B. A new comprehensive ecological risk index for risk assessment on Luanhe River, China. *Environ. Geochem. Health.* **2018**, *40(5)*, 1965–1978.
11. Lin, G.; Xu, X.; Wang, P.; Liang, S.; Li, Y.; Su, Y.; Li, K.; Wang, X. Methodology for forecast and control of coastal harmful algal blooms by embedding a compound eutrophication index into the ecological risk index. *Sci. Total Environ*. **2020**, *735*, 139404.
12. Zhang, H.; Chen, L. Using the Ecological Risk Index Based on Combined Watershed and Administrative Boundaries to Assess Human Disturbances on River Ecosystems. Hum. Ecol. *Risk Assess.: Int. J.* **2014**, *20(6)*, 1590–1607.
13. Thủy, H.T.T.; T.T.C.L.; Vy, N.H.N. Geochemical study of selected heavy metal in the aquatic sediments of Hochiminh City. *Sci. Technol. Develop., Vol 10, No.01 - 2007* **2007**, *10(1).* (In Vietnamese)
14. Anh, D.T.T.; C.V.H. Study on the distribution of heavy metals in sediment under the Cau River Basin. *J. Anal. Sci.* **2015,** *20(4).* (In Vietnamese)
15. Duong, P.T.; Tram, H.T.K. Research and evaluate contents of heavy metal in river sediment in the estuary of the Mekong river. *J. Sci. Ho Chi Minh Univ. Edu.* **2015**, *9(75)*, 119–129. (In Vietnamese)
16. Trinh, L.T.; Trang, K.T.T.; Trung, N.T.; Linh, N.K.; Tham, T.T. Heavy Metal Accumulation and Potential Ecological Risk Assessment of Surface Sediments from Day River Downstream. *VNU J. Sci.: Earth Environ. Sci.* **2018**, *34(4)*. (In Vietnamese)
17. Lợi, V.Đ.; N.V.; Quân, T.H.; Thuận, Đ.V.; Hà, P.T.T. Speciation of heavy metals in sediment of Tri An lake. *J. Anal. Sci.*. **2015**, *20(3).* (In Vietnamese)
18. Taiwo, A.M.; Michael, J.O.; Gbadebo, A.M.; Oladoyinbo, F.O. Pollution and health risk assessment of road dust from Osogbo metropolis, Osun state, Southwestern Nigeria. *Hum. Ecol. Risk Assess.: Int. J.* **2020**, *26(5)*, 1254–1269.
19. Wedepohl, H.K. The composition of the continental crust. Geochim. *Cosmochim. Acta.* **1995**, *59(7)*, 1217–1232.
20. Xu, Z.Q.; Ni, S.; Tuo, X.G.; Zhang, C.J. Calculation of heavy metal's toxicity coefficient in the evaluation of potential ecological risk index. *Environ. Sci. Technol.* **2008,** *31*, 112–115.
21. Hakanson, L. An ecological risk index for aquatic pollution control. A sedimentological approach. *Water Res.* **1980**, *14(8)*, 975–1001.
22. Muller.G. Index of geoaccumulation in sediments of the Rhine River. *Geo. J.* **1969**, *2*, 108–118.
23. Li, R.; Yuan, Y.; Li, C.; Sun, W.; Yang, M.; Wang, X. Environmental health and ecological risk assessment of soil heavy metal pollution in the coastal cities of estuarine bay–A case study of Hangzhou Bay, China. *Toxics* **2020**, *8(3)*, 75.
24. Ra, K.; Kim, J.K.; Hong, S.H.; Yim, U.H.; Shim, W.J.; Lee, S.Y.; Kim, Y.O.; Lim, J.; Kim, E.S.; Kim, K.T. Assessment of pollution and ecological risk of heavy metals in the surface sediments of Ulsan Bay, Korea. *Ocean Sci. J.* **2014**, *49(3)*, 279–289.
25. Tang, H.; Ke, Z.; Yan, M.; Wang, W.; Nie, H.; Li, B.; Zhang, J.; Xu, X.; Wang, J. Concentrations, Distribution, and Ecological Risk Assessment of Heavy Metals in Daya Bay, China. *Water*, **2018**, *10(6)*, 780.
26. Ali, M.M.; Islam, M.S.; Islam, A.R.M.T.; Bhuyan, M.S.; Ahmed, A.S.S.; Rahman, M.Z.; Rahman, M.M. Toxic metal pollution and ecological risk assessment in water and sediment at ship breaking sites in the Bay of Bengal Coast, Bangladesh. *Mar. Pollut. Bull.* **2022**, *175*, 113274.

JOURNAL OF
HYDRO-METEOROLOGY

Research Article

Evaluation of the effects of surfactants in the water of Kim Nguu River, Hanoi

To Xuan Quynh[1]*, Vu Duc Toan[2]*

[1] Trade Union University; toxuanquynh1304@gmail.com
[2] ROOM strong research, Environmental, and life science research laboratory, Thuyloi University; vuductoan@tlu.edu.vn

*Correspondence: toxuanquynh1304@gmail.com; vuductoan@tlu.edu.vn;
Tel.: +84–945801989

Received: 8 October 2023; Accepted: 13 November 2023; Published: 25 December 2023

Abstract: Surfactants are one of the new pollutants of current concern worldwide. Many studies show the toxicity of surfactants to human health and microorganisms. Kim Nguu River is one of Hanoi's wastewater drainage rivers, and the number of residents in the river area is quite dense. In addition, around the river area, there are many hospitals, factories, and industrial parks in operation. Especially the 8/3 textile factory where a lot of detergents are produced due to the bleaching and dyeing processes in the river. In this study, the author conducted sampling and analyzed the concentration of pollutants in the Kim Nguu River according to the method of Kadokami. This is a comprehensive analysis method that is both qualitative and quantitative According to analysis, there are currently two substances belonging to the group of surfactants: Nonylphenol and 4-tert-Octyphenol with concentrations measured in 2018 and 2019 of 4.1-12.45 µg/L, 11.4 to 29.56 µg/ L and 2.02-3.21 µg/L and 3.95-6.7 µg/L. The concentration of surfactants in the Kim Nguu River tends to increase sharply over time, and the level of impact of the concentration of surfactants in the Kim Nguu River on the ecosystem is at a high level. Kim Ngu River is just one of Hanoi's wastewater drainage rivers. Further research is needed on other wastewater drainage rivers to have clearer conclusions about the appearance of surfactants. in rivers draining wastewater.

Keywords: Surfactants; Nonylphenol; 4-tert-Octyphenol; Pollution.

1. Introduction

There have been several studies around the world showing the toxicity of surfactants to the ecosystem [1]. We can mention research on crustaceans, yeast, and algae to find out the effects of surfactants on organisms [2]. Some studies on the toxicity of Nonylphenol to the environment include research by Shaukat [3] about the effects of Nonylphenol on Brine shrimp: Artemia sinica and the results obtained are that compared to n-heptyl phenol, Nonylphenol, t-butyl phenol, 2,4-dichlorophenol, and bisphenol A, Nonylphenol has a strong harmful effect on artemia. second only to n-heptyl phenol. Another study by [4] with Fish: Pseudochromis fridamani and Algae: Selenastrum capricornutum showed surprising results, all fish died after only 8 days of exposure to Nonylphenol, the ability to produce and form cells in green algae decreased significantly [5].

In Vietnam, there are very few studies on surfactants in water. One of the few studies showed that Nonylphenol compounds have caused serious effects on the vitality and reproduction of crustacean species C. cornuta, D. lumholtzi, and D. magna. This can lead to a decline in the number of zooplankton and the loss of ecological diversity in water bodies

with high Nonylphenol content [6]. This study can be a premise for further research on surfactants in the river and also partly warns about the currently quite high concentration of surfactants in the river.

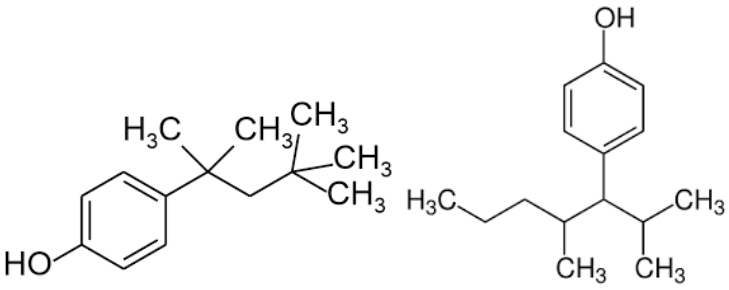


Figure 1. Structure of Nonylphenol and 4-tert-Octyphenol.

Surfactants are amphiphiles consisting of both hydrophilic and hydrophobic parts, which can participate in intermolecular interactions that contribute to their surface activity. When the hydrophilic parts are arranged together on the outside and the hydrophobic parts are arranged together in the center, it is called a Micelle. These micelles increase the solubility of pollutants while increasing the toxicity of the surfactant.

The mechanism of action of surfactants is to reduce the surface tension of the solvent they contain, making the solvents easily disperse into each other. Surfactants can remove dirt and grease but cannot treat them and carry them out into the environment. When exposed to the environment with a mechanism of action to reduce surface tension, surfactants carry impurities that make it difficult for oxygen and sunlight to penetrate, leading to a serious decrease in dissolved oxygen. This is the cause of disturbance in the activities of microorganisms [7].

Nonylphenol is a pale-yellow liquid that is sparingly soluble in water but soluble in organic solvents. Nonylphenol is used to produce antioxidants, detergents, and cleaners. Since Nonylphenol was first synthesized in 1940 its use and production have increased rapidly [8, 9].

4-tert-Octyphenol is also a surfactant. 4t-OP is very toxic to aquatic organisms. In rainbow trout, juvenile salmon growth was significantly reduced when exposed to 4t-OP. 4-tert-Octyphenol and Nonylphenol are both endocrine disruptors for humans and organisms [9–11].

2. Materials and Methods

2.1 Methods of investigation and data collection

The section of the Kim Nguu River studied is 4km long, from the beginning of Kim Nguu Street to the end of the Yen So wastewater treatment plant. Through the process of field investigation, the author selected 6 sampling locations, each location characteristic of the waste source discharged into the Kim Nguu River. M1 is the starting point for research, determining the initial concentration of the Kim Nguu River; M2 is the point with waste sources from residential areas and 3 hospitals; M3 is the concentration point of waste sources in a new urban area; M4 is the point with waste sources from residential areas and a hospital; M5 is the point with waste sources from Vinh Tuy Industrial Park, wholesale markets, and residential areas; M6 is the treated wastewater point of Yen So wastewater treatment plant.

Table 1. Location of sampling sites in the Kim Nguu River.

Numerical order	Location Description	Longitude	Latitude
1	M1: The beginning of Kim Nguu Street	21°00’25.7”N	105°51’39.7”E
2	M2: Lac Trung Bridge	21°00’10.8”N	105°51’41.5”E

Numerical order	Location Description	Longitude	Latitude
3	M3: Mai Dong Bridge	20°59’46.5”N	105°51’44.0”E
4	M4: KuO Bridge	20°59’17.0”N	105°51’47.3”E
5	M5: In front of the Yen So wastewater treatment plant	20°58’33.6”N	105°51’55.3”E
6	M6: Yen So regulating lake	20°58’13.5”N	105°51’36.8”E

2.2. Sampling method

The six water samples were all taken in April of 2018 and 2019 (Figure 2). All water samples were taken according to TCVN 6663-3-2008 standards. After taking water samples, they were refrigerated and brought back for analysis at the Laboratory of the Institute of Environmental.

Figure 2. Location of the study area.

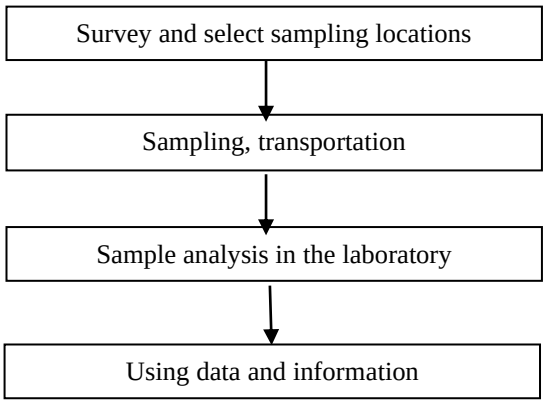


Figure 3. Research structure diagram.

2.3. Methods of sample analysis

Bring the water sample to room temperature. Take a 500ml water sample, add 30g of anhydrous NaCl, and shake well. Add 100ml of dichloromethane to the solution and continue shaking for 10 minutes, Let the shaking solution rest for 10 minutes, and then filter the extract through a funnel containing anhydrous Na_2SO_4. Repeat the above process two more times, each time with 50ml of dichloromethane. Use a rotary vacuum evaporator to concentrate the

obtained extract. Clean the newly concentrated solution using an activated Silicagel column to obtain the cleaned solution. Continue to concentrate the cleaned solution using a rotary vacuum evaporator until the solution remains about 5ml. Concentrate the solution to 1ml by blowing N_2 gas. The solution, after being blown with N_2 gas, was stored in vials, and then put into a gas chromatograph-mass spectrometer for analysis [12].

3. Results and discussion

3.1 Evaluation of insecticide contamination in the water of the Kim Nguu River

After the sample is extracted and put into 1ml vials, it will be put into a GC-MS-MS-SRM 9TSQ Quantum XLS machine, Thermo Fisher Scientific, USA for analysis. The machine has a library of more than 900 substances, however, that library does not contain some typical surfactants such as Sulfate-based substances and Sodium-based substances, these substances are found in many domestic wastewaters. Kim Nguu is the recipient of domestic wastewater from a densely populated area. It is possible that these substances will not be detected by this analytical procedure. Surfactants in Vietnam are still limited in research on their accumulation in organisms, and only a few substances have been studied.

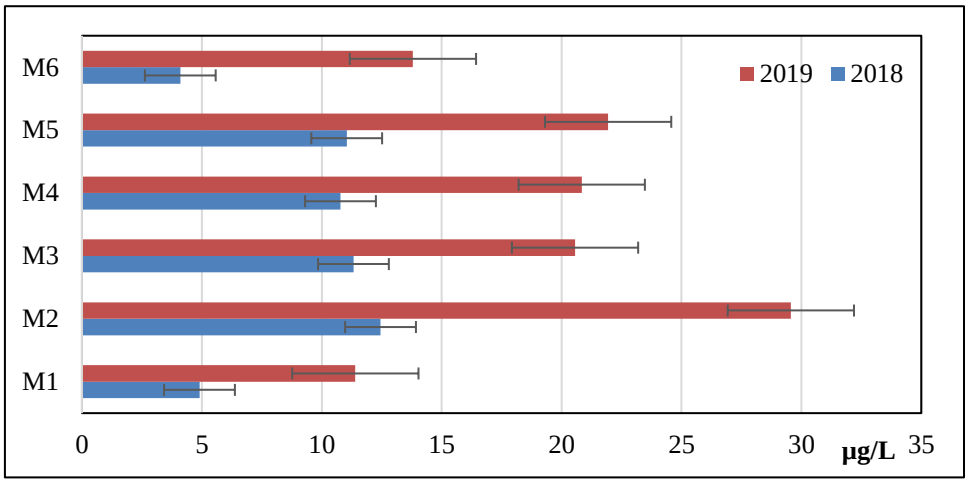


Figure 4. Nonylphenol concentration in Kim Nguu River.

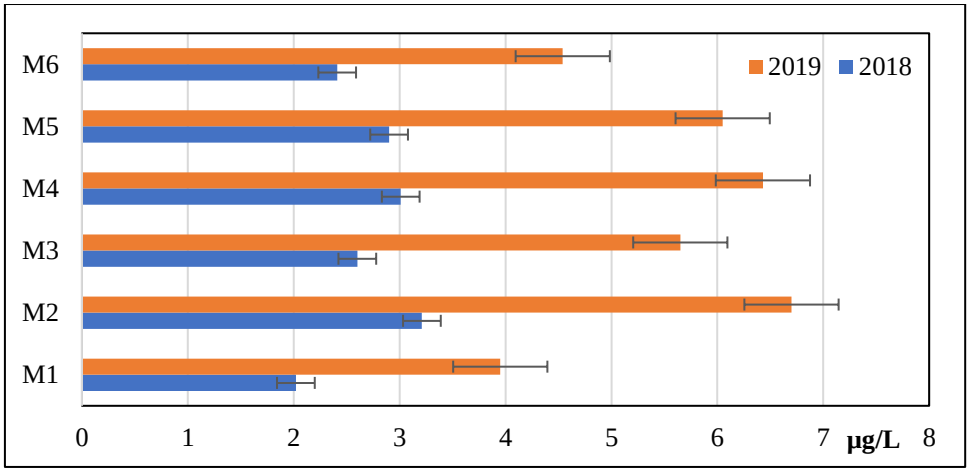


Figure 5. 4-tert-Octyphenol concentration in Kim Nguu River.

4-tert-octyphenol concentrations recorded in 2018 and 2019 were 2.02-3.21µg/L and 3.95-6.7 µg/L, respectively. Like the concentration of Nonylphenol, the concentration of 4-tert-octyphenol also increases over time for the same reason. Looking at Figure 5, it can be seen that the concentration trends at the sampling locations of Nonylphenol and 4-tert-

octyphenol are almost the same, the only difference is the concentration of 4-tert-octyphenol at the location. M5 is not the position with the second highest concentration. This may be explained by the fact that 4-tert-octyphenol is not a substance widely used in the textile dyeing industry.

The concentration of Nonylphenol in Kim Nguu River water in 2019 ranged from 11.4 to 29.56 µg/L, with an average concentration of 19.68µg/L. With this concentration, the Nonylphenol concentration in Kim Nguu River water is much higher than the average of 24 surface water samples in Yellow River China (average concentration 0.805µg/L) [13]. Some other countries such as Germany (average 0.13µg/L) [14], Switzerland (average 0.48µg/L) [15], and Austria (average 0.89µg/L) [16] are all much higher.

3.2 Ecological risk assessment of Nonylphenol in Kim Nguu River water

There are many studies around the world on the effects of Nanolphenol concentrations on aquatic organisms. As for 4-tert-octyphenol, the toxicity is added to the same Phenol groups. Within the scope of this study, only the ecological effects of Nonylphenol will be presented.

To evaluate the ecological effects of Nonylphenol, studies have used several parameters: Criteriton maximum concentration- CMC, and Criteriton continuous concentration - CCC. The values of the CMC and CCC parameters are specified in the National Recommended Aquatic Life Criteria table [17–19].

From Figure 6, it can be easily seen that the Criterion continuous concentration - CCC of Nonylphenol in 2018 has 2 positions M1 and M6 are at the chronic threshold and the rest are above the acute threshold. By 2019, concentrations at all locations had exceeded the acute threshold, no location in 2 years was below the chronic level.

From Figure 7, it can be seen that the Criteriton maximum concentration - CMC of Nonylphenol in 2018 was all below the acute level, positions M1, and M6 were below the chronic level, and the remaining positions were above the chronic level. In 2019, at location M2, the concentration exceeded the acute level and the remaining locations exceeded the chronic level.

Table 2. National Recommended Aquatic Life Criteria.

Pollution	Freshwater CMC (Acute µg/L)	Freshwater CMC (Chronic µg/L)	Freshwater CCC (Acute µg/L)	Freshwater CCC (Chronic µg/L)
Nonylphenol	28	7	6.6	1.7

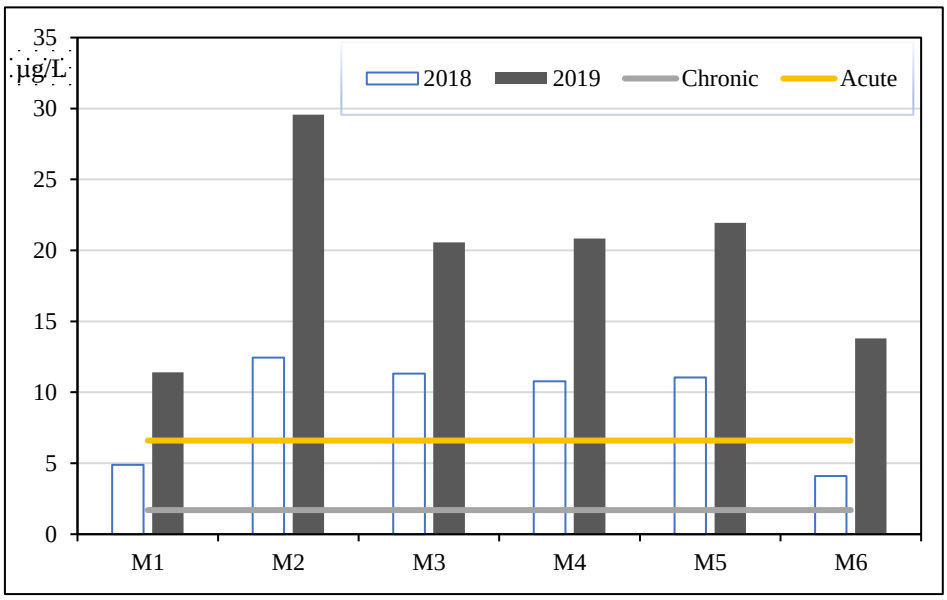


Figure 6. Criterion continuous concentration of Nonylphenol.

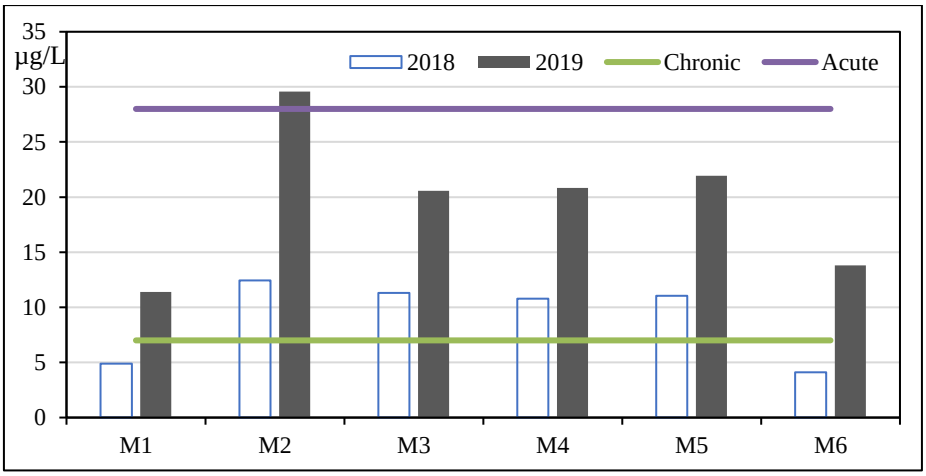


Figure 7. Criterion maximum concentration of Nonylphenol.

In general, the concentration of Nonylphenol in KimNguu River water is at an alarming level for the ecosystem, with an increasing trend each year, there is a high possibility that the level of impact will be even higher.

4. Conclusion

The analytical method of Kadokami et al can detect Nonylphenol and 4-tert-Octyphenol in water of Kim Nguu River, Hanoi. Research is needed to select appropriate analysis methods for surfactants so that more substances can be analyzed.

Through the analysis process, it was discovered that there are 2 surfactants in the Kim Nguu River. Nonylphenol and 4-tert-Octyphenol with concentrations measured in 2018 and 2019 of 4.1-12.45 µg/L, 11.4 to 29.56 µg/L and 2.02-3.21µg/L and 3.95-6.7 µg/L. The concentration of surfactants tends to increase each year.

Nonylphenol concentration in Kim Nguu River water is at a level that adversely affects the ecosystem at all locations. Nonylphenol concentration in Kim Nguu River water is also much higher than Nonylphenol concentration in water of some rivers in the world.

Author contribution statement: Constructing research idea: T.X.Q., V.D.T.; Select research methods: T.X.Q., V.D.T.; Data processing: T.X.Q., V.D.T; Sample analysis: T.X.Q.; Take samples: T.X.Q; Writing original draft preparation: T.X.Q., V.D.T.; Writing review and editing: T.X.Q., V.D.T.

Acknowledgments: The authors would like to thank the strong research group ROOM, Environmental, and Life Science Research Laboratory, Thuyloi University for their support during the research.

Competing interest statement: The authors declare that this article was the work of the authors, has not been published elsewhere, and has not been copied from previous research; there was no conflict of interest within the author group.

References

1. Kadokami, K.; Jinya, D.; Iwamura, T. Survey on 882 Organic Micro-Pollutants in Rivers throughout Japan by Automated Identification and Quantification System with a Gass Chromatography-Mass Spectrometry Database. *J. Environ. Chem.* **2009**, *19(3)*, 351–360.
2. Jena, G.; Dutta, K.; Daverey, A. Surfactants in water and wastewater (greywater): Environmental toxicity and treatment options. *Chemosphere* **2023**, *341*, 140082. https://doi.org/10.1016/j.chemosphere.2023.140082.

3. Shaukat, A.; Liu, G.; Li, Z.; Xu, D.; Huang, Y.; Chen, H. Toxicity of five phenolic compounds to brine shrimp Artemia sinica. *J. Oncean Univ. China.* **2014**, *13*, 141–145.
4. Hamlin, H.J.; Marciano, K.; Downs, C.A. Migration of nonylphenol from food- grade plastic is toXic to the coral reef fish species Pseudochromis fridmani. *Chemosphere* **2015**, *139*, 223–228.
5. Spehar, R.; Brooke, L.T.; Markee, T.P.; Kahl, M.D. Comparative to Xi city and bioconcentration of nonylphenol in fresh water organisms. *Environ. Toxicol. Chem.* **2010**, *29*, 2104–2111.
6. Chi, V.T.M. Ảnh hưởng của hợp chất gây rối loạn nội tiết nonylphenol lên sức sống và sinh sản của ba loài vi giáp xác, Ceriodaphnia cornuta, Daphnia lumholtzi và Daphnia magna. *Tạp chí Khoa học Trường Đại học Cần Thơ* **2016**, *43*, 34–41. (In Vietnamese)
7. Nhan, L.T.H. Công nghệ Chất hoạt đông bề mặt, Hồ Chí Minh. Đại học Bách Khoa - Đại học Quốc gia HCM, 2012. (In Vietnamese)
8. Agency, U.E.P. Aquatic life Ambient Water Quality Criteria - Nonylphenol. 2005.
9. Bina, B. Determination of 4-nonylphenol and 4-tert-octylphenol compounds in various types of wastewater and their removal rates in different treatment processes in nine wastewater treatment plants of Iran. *Chin. J. Chem. Eng.* **2018**, *26*, 183–190.
10. Pan Jiangqing, L.J.W.D. Study of the Mechanism of the Synergistic Action of Nickel Stearate and 2,2'-Thiobis-(-4-Tert-octylphenol) in the Photo-oxidation of Polypropylene. *Polym. Degrad. Stab.* **1991**, *32*, 313–320.
11. Environmental Protection Agency, U.S. Phenol, 4-(1,1,3,3-tetramethylbutyl), 2014.
12. Quynh, T.X.; Toan, V.D. Endocrine Disrupting Compounds (EDCs) in surface waters of the KimNguu River, Vietnam. *Bull. Environ. Contam. Toxicol.* **2019**, *103*, 734–738.
13. Wang, L.; Ying, G.; Chen, F.; Zang, I. Monitoring of selected estrogenic compounds and estrogenic activity in surface water and sediment of the Yellow River in China using combined chemical and biological tools. *Environ. Pollut.* **2012**, 241–249.
14. Kuch, M.H.; Ballschimer, K. Determination of endocrine disrupting phenolic compounds and estrogens in surface and drinking water by HRGC-NCL-MS in the program per litter ranger. *Environ. Sci. Technol*. **2001**, *35*, 3201–3206.
15. Fenner, K.; Kooijman, C.; Scheriger, M. Including transformation products into the risk assessement for chemicals; the case of nonylphenol ethoxylate usager in Switzerland. *Environ. Sci. Technol*. **2001**, *36*, 1147–1154.
16. Hohenblum, P.; Gans, O.; Moche, W. Monitoring of selected estrogenic hormones and industrial chemicals in groundwaters and surface water in Austria. *Sci. Total Environ.* **2004**, *333*, 185.
17. Zhang, J.; Shi, J.; Ge, H.; Tao, H.; Guo, W.; Yu, X.; Zhang, M.; Li, B.; Xiao, R.; Xu, Z.; Li, X. Tiered ecological risk assessment of nonylphenol and tetrabromobisphenol A in the surface water of China bases on the augmented species sensitivity distribution models. *Ecotoxicol. Environ. Saf.* **2022**, *236*, 113446.
18. EPA, U. Aquatic Life Ambient Water Quality Criteria - Npnylphenol. EPA, Washington DC, 2005.
19. Zang, L.; Wei, C. Criteria for assessing the ecological risk of nonylphenol for quatic life in Chinese surface fresh water. *Chemosphe* **2017**, *184*, 569–574.

JOURNAL OF
HYDRO-METEOROLOGY

Research Article

Assessment of the influence of urban flood in Thu Duc City in the period of planning

Thinh Nguyen Tran Phu[1,2], Kim Tran Thi[3], Phung Nguyen Ky[2]*

[1] Institute of Environment and Natural Resources, Vietnam National University Ho Chi Minh City, 142 To Hien Thanh Street, Ward 14, District 10, Ho Chi Minh City, Vietnam; thinhnguyenputi@gmail.com

[2] People's Committee of Thu Duc City, 168 Truong Van Bang, Thanh My Loi Ward, Thu Duc City, Ho Chi Minh City, Vietnam; kyphungng@gmail.com

[3] Faculty of Marine Resource Management, Ho Chi Minh City University of Natural Resources and Environment, Vietnam, 236B Le Van Sy Street, Ward 1, Tan Binh District, Ho Chi Minh City, Vietnam; ttkim@hcmunre.edu.vn

*Corresponding author: kyphungng@gmail.com; Tel.: +84–908275939

Received: 8 October 2023; Accepted: 14 November 2023; Published: 25 December 2023

Abstract: Ho Chi Minh City is a low-elevation coastal city with and rapidly increasing population. Thu Duc, the city in Ho Chi Minh City, is the first city in Vietnam that belongs to the type of city administrative unit directly under the central city. This area is greatly affected by flooding, classified into three groups: heavy precipitation, tidal current, and combined tidal current and heavy precipitation. In this study, the flooding situation caused by the groups in the planning period (from 2010 to 2020) will be assessed and mapped by using GIS. The results show that, in the period from 2016 to 2021, flooding by heavy precipitation was more serious than that by tidal current. Compared to the drainage planning map in District 2, District 9, and Thu Duc District by 2020, some flooded locations due to no drainage system, lack of drainage systems, and unresponsive drainage systems have not yet been planned. In District 2 and District 9, some locations experienced flooding attributed to heavy precipitation, yet there is currently no drainage plan in place until 2020. In Thu Duc district, the majority of flooding incidents caused by precipitation can be attributed to the degradation of the existing drainage system, compounded by the absence of a sewer system. Furthermore, there were more points of re-flooding and more severe flooding compared to the same time in the previous period. This research is the scientific basis for flood control planning for the city as well as socioeconomic planning.

Keywords: Thu Duc City; Urban flood; Planning; Heavy precipitation; Tidal current.

1. Introduction

Urban flooding is the combined result of natural factors (precipitation and tidal current) and humans including drainage systems and river networks and other drainage structures [1]. Urban flooding causes serious damage to infrastructure, the economy, and people's lives. Cities with a high population density and high urbanization rate have many potential risks, including the problem of urban flooding [2–6].

In a recent report on natural disasters by the United Nations Office for Disaster Risk Reduction (UNISDR) and the Centre for Research on the Epidemiology of Disasters (CRED) in 2015, it was highlighted that 43% of such disasters occurring between 1995 and 2015 were attributed to floods. These flood-related events had a profound impact, affecting

over half (56%) of all individuals affected by any category of natural disaster, and tragically resulting in the loss of life for approximately a quarter (26%) of those affected [7]. In numerous Chinese cities, urban flooding has emerged as a significant peril [8]. This issue is exacerbated by the substantial lag in the development of underground stormwater infrastructure, including drainage facilities, which struggle to keep pace with the rapid urban population growth and economic expansion. Consequently, instances of severe urban flooding have become increasingly frequent [8]. In 2017, it was estimated that urban flooding resulted in a direct economic loss of approximately 214 billion RMB, equivalent to approximately US$ 31.7 billion [9].

Ho Chi Minh City is a large metropolis in Vietnam, which has been severely flooded in recent years [10]. It is ranked in the top 20 cities with the highest population size affected by coastal flooding [11, 12]. The topography of the city is relatively low (nearly 65% of the area has a natural elevation of < 1.5 m), influenced by tides from the East Sea, limiting the ability to drain water [13]. During the rainy season, flooding in low-lying areas and also in central districts occurs frequently between August and December. According to Circular No.338/BXD-KTQH dated March 10, 2003, of the Ministry of Construction, Vietnam on the urban water drainage program, urban flooding is a flood situation in which the flooding points in the inner city are determined according to the following parameters: the volume of water in the area must be greater than 1,000m^3, equivalent to an inundation range of 500m long, 20m wide, and 0 deep. 1m, flooding time is 30' after rain.

The causes of flooding in the city are summarized as: (1) runoff from heavy precipitation is generated that exceeds the drainage capacity of the sewer system [14, 15]; (2) the tidal increases in the river and canal system [16, 17]; (3) The influence of flood discharge from upstream reservoirs [18, 19]; (4) the infrastructure failure to respond to rapid urbanization [20]; (5) the concreting leads to increased runoff [21]; (6) the subsidence of ground level [22]; (7) the lakes, rivers, canals, natural low-lying areas have been leveled to serve socio-economic development [23]; (8) the planning and development of residential areas are not reasonable [15]; (9) the progress in implementing the plans is still slow [24]; (10) people lack awareness in protecting and maintaining the drainage system [25].

Thu Duc City, Ho Chi Minh City, Vietnam was established under Resolution 1111/NQ-UBTVQH14 of the UBTVQH14 (effective from January 1, 2021); is a city directly under Ho Chi Minh City (HCMC), which was merged by 3 districts: District 2, District 9 and Thu Duc District (natural area of nearly 212 km^2). The city is a low-lying coastal city has a dense population, with a total population of more than 1 million people [26]. Furthermore, it has a dense system of rivers and canals, strongly influenced by the hydrological regime of the Saigon River, when high tides and heavy precipitation often cause flooding in areas with topography below 2.0 m [13]. Although the drainage system is regularly maintained and upgraded, floods due to heavy precipitation, tidal current, and combined tidal current and precipitation occur on some of the city’s main streets which have seriously been affecting the people’s lives in the area [15, 27, 28]. Currently, Thu Duc City still has many flooded points, especially, some of these are major streets with high traffic volume, seriously affecting people's lives and the economy. Compared with the number of regular flooding points reported by the steering center of urban flood control in 2015 of 18 points, the number of statistical flooding points in 2022 has nearly doubled to 37 points in the whole city. Newly the points appear mainly in Thu Duc District and a few points in the west of District 9.

Specifically, for Thu Duc City, a very young city newly established in January 2021 through the merger of three districts: District 2, District 9, and Thu Duc District - the issue of flooding and the impact of climate change become even more severe without a coordinated development direction.

To adapt and prevent flooding, there are various computational methods such as mathematical modeling [29–31], Geographic Information System (GIS) [32, 33], ecological reservoirs, urban reservoirs, and machine learning, serving comprehensive planning. In practice, many strategic plans have been implemented, leveraging national resources to promote sustainable development trends in the region and control urban flooding. These include urban and rural system development planning, functional area development planning, technical infrastructure development planning, land use planning, 5-year land use plans, district and inter-district development planning, water resource exploitation and utilization planning, mitigation of water-related impacts, disaster prevention and response to climate change planning.

In this study, the flooding situation caused by tides, heavy precipitation, and a combination of these factors during the planning period (from 2010 to 2021) with 63 flooded points will be evaluated and mapped using GIS. Subsequently, a comparison will be made with the urban planning to assess 37 heavily street flooding locations. This assessment will aid in identifying the causes of flooding and proposing suitable solutions. The results will serve as the fundamental basis for designing future drainage systems and for the assessment and management of flood risks.

2. Materials and methods

2.1. Study area

Thu Duc city is located in the east of Ho Chi Minh City, in the southern key economic region, and is the focal point of the main traffic routes between Ho Chi Minh City and the Southeast provinces. It is contiguous as follows: The East is Bien Hoa City and Long Thanh District, Dong Nai Province with the border of Dong Nai River. To the west are District 12, Tan Binh District, Binh Thanh District, District 1, and District 4 with the boundary of the Saigon River. To the south are Nhon Trach district, Dong Nai province, and District 7 (through the Saigon River), and to the north is Binh Duong province (Figure 1).

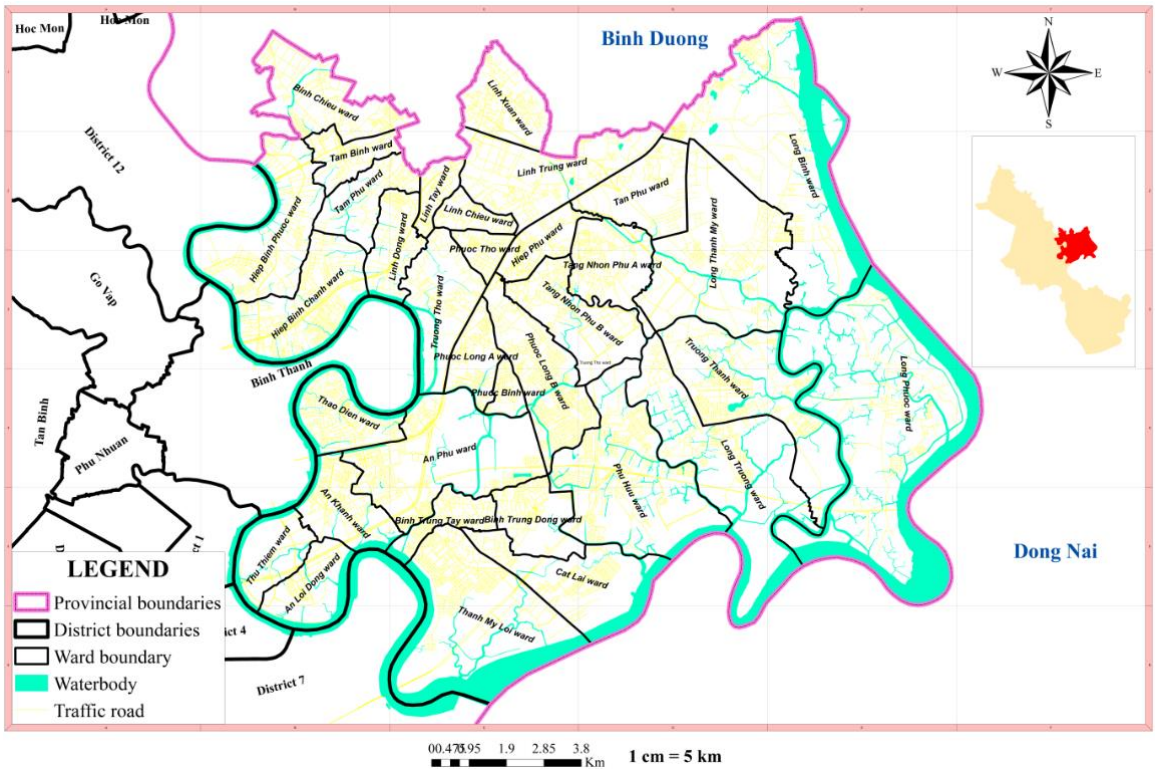


Figure 1. Study area.

Low-lying topography accounts for over 60% of the city's land with elevations below 2m (Figure 2). Besides, the geological foundation is weak and prone to subsidence, compression, and frequent bank failure. The drainage system, which doesn't meet the population growth rate in the area, is not continuous due to investment by many separate and asynchronous projects.

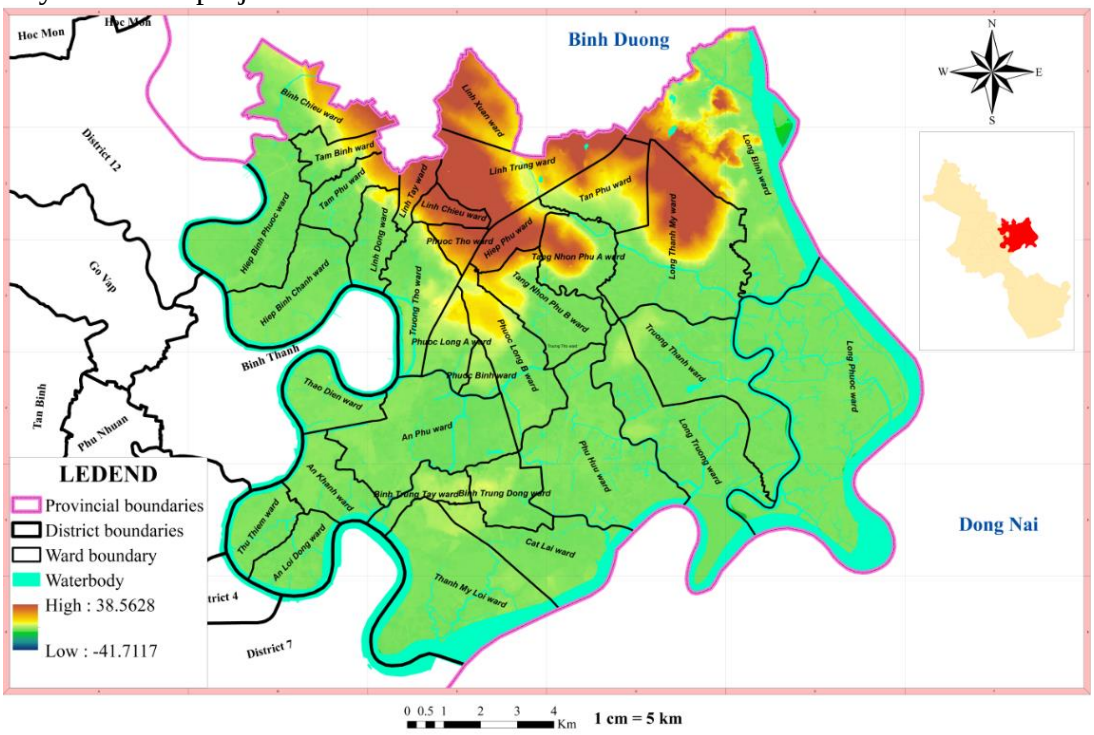


Figure 2. Digital elevation model.

2.2. Data collection

In the study, the number of times flooded and the average depth of flooding in the period of 2010-2021 of 63 flooded points in Thu Duc City, which was formerly Districts 2, 9, and Thu Duc District, is collected from Ho Chi Minh City Infrastructure Management Center and Urban Management Office of Thu Duc City. In which, the number of occurrences of flooding and the average depth of flooding are entirely sufficient for all 37 heavily flooded street locations.

The maps collected when merging the three districts are (1) The Administrative Map of Thu Duc City from the People's committee Thu Duc City, and (2) the Digital Elevation Map (DEM) from the Center for Applied GIS of Ho Chi Minh City (HCMGIS).

The other maps are collected individually for each old district at different times including: (1) Drainage planning map in District 2, District 9, and Thu Duc District by 2020; (2) Drainage map in District 2 in 2012, in District 9 in 2009, and Thu Duc District in 2011 (3) Environmental Assessment and Strategic Environmental Map in District 2 in 2012 from Ho Chi Minh City Department of Planning and Architecture.

2.3. Methodology

The study is structured according to the following framework (Figure 3).

There are three steps in this study:

Step 1: Data collection and mapping of flooding.

Step 2: Statistical analysis of flooding depth and area at 63 flooding points in the city, with detailed information on flooding depth and area for 37 heavily flooded streets during the periods 2010-2015 and 2016-2021.

Step 3: Analysis of the causes of flooding and comparison with urban planning.

Collecting and synthesizing documents

The method of collecting and synthesizing documents is carried out based on the inheritance, analysis, and synthesis of relevant sources of documents and information, selectively chosen from data such as reports, international and domestic documents, and from published studies/reports.

Geographic Information System

GIS - Geographic Information System, is an organized collection, including computer hardware, software, geographic data, and people, designed to capture, store, update, control, analyze, and display all forms of geo-location-related information. It allows the construction of spatial analysis, management, integration of information layers.

Applying GIS, remote sensing, and specialized software to systematize and map data and calculation results. In this study, the GIS method is used to visualize images, map flood points and process planning maps.

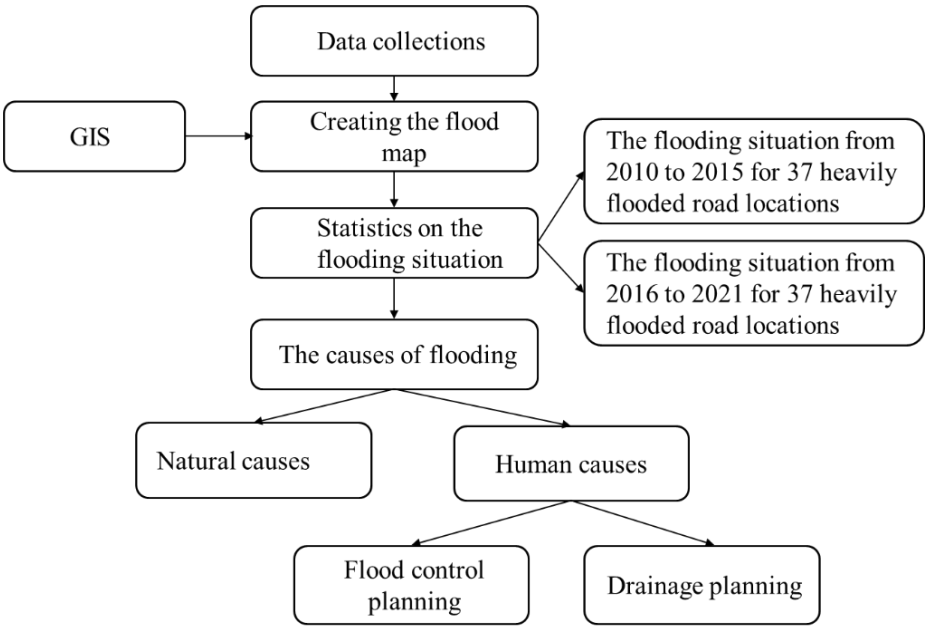


Figure 3. Implement framework.

3. Results

3.1. The flooding situation

Generally, the flood situation in Thu Duc City has been happening quite complicatedly. There is an increasing trend in both the number of flooded points and flooding times. Notably, heavy precipitation combined with tidal currents caused some streets in the central area of District 2 and Thu Duc where there is a low-lying and low foundation with low-lying foundations to be regularly flooded. The location of flood points is described in Figure 4 and the current flooding situation of all flooding points is described in Table 1.

a) Flooding due to heavy precipitation:

In the precipitation with a volume of 30 mm, mild flooding conditions begin to appear. When the precipitation volume reaches 40-50 mm or more, moderate to heavy flooding conditions occur.

In the precipitation of 60-70 mm in volume, about 50% of the existing flooded locations are moderately flooded. Meanwhile, all flooding locations in the city are flooded with moderate to severe precipitation with a volume of 80-100 mm or more. The areas heavily flooded by precipitation include National Highway 13 (No. 31, No. 32), Road No.

10 (No. 39) in Linh Dong Ward; Road 26 (No. 49) in Linh Dong Ward, Alley 789 (No. 42) in Tam Binh Ward; Cau Xay Street (No. 57) in Tan Phu Ward, etc.

The number of flooded points and the level of flooding will increase significantly when heavy precipitation coincides with tidal current. Flooding due to heavy precipitation not only impacts the living environment and people's lives but also directly affects transportation, causing severe traffic congestion.

According to a study by Huong (2022), in District 9, there are 350 streets, 74 traffic bridges, and 940 alleys. Waterlogging on the streets during heavy precipitation makes transportation challenging, leading to potential engine failures and accidents such as vehicle overturns and people being swept away. The streets are prone to congestion and flooding during precipitation, resulting in extended travel times, property damage, increased transportation costs, and elevated environmental pollution [34].

b) Flooding due to tidal currents:

The river and canal system of Thu Duc City is influenced by the semi-diurnal tidal regime from the East Sea through major rivers such as the Saigon River and Dong Nai River. The water level varies seasonally and at different locations due to the tidal changes downstream and the flood discharge regime of upstream structures. The annual water level on the Saigon - Dong Nai - Nha Be River fluctuates from +1.15 m to +1.70 m, with the difference between the highest and lowest water levels ranging from 3.5 m to 4.0 m.

According to survey results, tidal flooding in low-lying areas of the city begins to occur at tide levels from +1.0 m and can currently reach up to +1.7 m.

Regarding flooding by tidal currents, heavily flooded areas are concentrated in District 2 and Thu Duc District. Thu Duc District has heavily flooded areas including No. 43 street (No. 40) in Hiep Binh Chanh ward, street 12 (No. 41) in An Khanh ward, street 24 (No.48) in Linh Dong ward, etc. District 2 is a typical area flooded due to construction which doesn't pay attention to the background during construction work. It is a fact that many residential areas were built with foundations lower than the high tide level, so they are often flooded by tides. More specifically, the flood situation is more severe for the areas near the large canals and low-lying as Thao Dien ward and a part of Binh An ward (Figure 4).

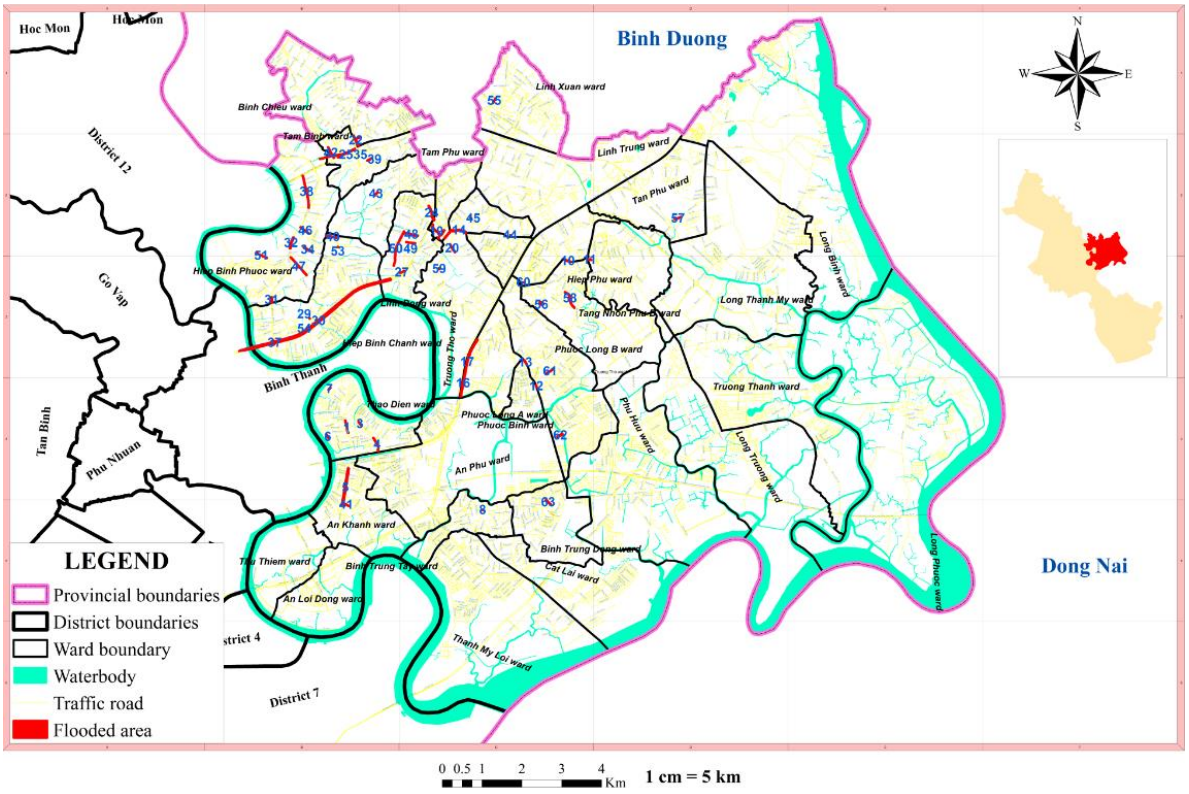


Figure 4. The location of flooded points in Thu Duc City, Ho Chi Minh City, Vietnam.

Flooding by tidal currents cause phenomena such as embankment breaches, overflow, and flooding in certain areas, such as Thu Duc District (Hiep Binh Phuoc, Hiep Binh Chanh, Ling Dong, Tam Phu, Tam Binh) and District 2 (Thao Dien, An Loi Dong, An Khanh), impacting the lives and activities of residents. This leads to environmental pollution, compromised living conditions, compromised sanitation and safety, challenges in food safety, disruptions to traffic, and difficulties in the mobility of residents.

Table 1. The current flooding situation of all flooding points.

No.	Street	Flooding range		Average flooded depth (m)	Average flooded area (m^2)	Average flooding time (minutes)	Causes of flooding		
		From	To				Tide	Precipitation	Tide & precipitation
1	Quoc Huong	Tong Huu Dinh	Alley 76	0.225	2080	80		x	
2	Quoc Huong	Tong Huu Dinh	Street 65	0.2	1550	260		x	
3	Thao Dien	Alley 95	Police of Thao Dien Ward	0.3	2700	-			x
4	Thao Dien	Hanoi highway	Xuan Thuy	0.15	2200	195			x
5	Tran Nao	Dieu Giac Temple	2nd Street	0.2	16800	38		x	
6	Nguyen Van Huong	In front of Hoang Anh Gia Lai Apartment		0.15	1500	200		x	
7	Nguyen Van Huong	Before House No. 170		0.2	1800	165		x	
8	Nguyen Duy Trinh	In front of the children's culture house in District 2		0.2	2000	40		x	
9	Luong Dinh Cua	Thu Thiem Bridge	Cat Lai ferry	0.13	3250	120			x
10	Le Van Viet	Phong Phu Communal House	Alley 201	0.15	1440	120		x	
11	La Xuan Oai	House number 34	Outlet	0.2	2000	90			x
12	Do Xuan Hop	People's Committee of Phuoc Binh Ward		0.35	1400	160		x	
13	Do Xuan Hop	Vocational colleges		0.23	1490	38		x	
14	Vo Van Ngan	Dang Van Bi	Thu Duc Market	0.2	4000	120		x	
15	Ha Noi Highway	Power pole T7C	Power pole T30C	0.2	8000	100		x	
16	Ha Noi Highway	Charging station	Rach Chiec bridge	0.2	1440	120		x	
17	Ha Noi Highway	Rach Chiec bridge	Tay Hoa	0.2	400	35		x	
18	Dang Thi Ranh	Duong Van Cam	To Ngoc Van	0.2	900	15		x	
19	Duong Van Cam	Le Van Tach	Dang Thi Ranh	0.4	1840	40		x	
20	Ho Van Tu	Kha Van Can	Police Department of Truong Tho Ward	0.175	1120	25		x	
21	Le Thi Hoa	Provincial Street 43	House number 21	0.5	3355	35		x	
22	Provincial Street 43	Binh Chieu	Highway 1	0.4	3600	25		x	
23	To Ngoc Van	Railway	Pham Van Dong	0.35	1300	60		x	
24	To Ngoc Van	Linh Dong	Linh Tay stream outlet	0.2	1200	25		x	
25	Highway 1	Thanh Binh gas station	Go Dua overpass	0.15	2000	20		x	
26	Kha Van Can	Duong Van Cam	Thu Duc Post Office	0.21	3453	10		x	
27	Kha Van Can	Gas station 7/27	Vo Uu Temple						
28	Kha Van Can	No. 1 outlet	House number 617						
29	Kha Van Can	Rach Mon area							

No.	Street	Flooding range		Average flooded depth (m)	Average flooded area (m^2)	Average flooding time (minutes)	Causes of flooding		
		From	To				Tide	Precipitation	Tide & precipitation
30	Pham Van Dong	House number 148	House number 7/29	0.3	3900	40		x	
31	National Highway 13	Ong Dau Bridge	Temple of Binh Trieu	0.146	916	71		x	
32	National Highway 13	Gia Dinh Shoe Company	Hiep Binh Street	0.2	1600	50		x	
33	Dang Van Bi								
34	Hiep Binh	1st Street	Hiep Binh Secondary School						
35	Go Dua	National Highway 1A (Binh Phuoc Overpass)	To Ngoc Van	0.2375	1767	58		x	
36	Le Van Tach	House number 3	Duong Van Cam						x
37	Binh Trieu Crosstrees	Binh Trieu Roundabout							x
38	Old Highway 13	Highway 1	Highway 13	0.2	4200	120		x	
39	Street no. 10	To Ngoc Van Street	Alley 30, street 10	30	360	60		x	
40	No. 43 street	Tam Binh	Rach Dia	0.3	1318.5	30			x
41	Street 12	Tran Nao	Residential Project of Caric Joint Stock Company	0.3	1000	90		x	
42	Alley 789	At the beginning of Song Hanh National Street 1	Thu Duc Farmers Market	0.2	1120	30		x	
43	Alley 37, Acacia Tree Street	After Chau Hung Pagoda	Ring Street 2 intersection (under construction)	0.3	880	30		x	
44	Alley 95, Vo Van Ngan Street	House number 95/26	House No. 95/36	0.5	80	30		x	
45	Alley 2, Street 17, Quarter 5	Street 17	At the end of the alley	0.3	13395900	60		x	
46	Alley 606 R13	National Street 13	Adjacent to Hong Long project	0.2	400	120			x
47	Sixth Street, Quarter 6	National Street 13	Kinh Do canal	0.2	600	150			x
48	Street 24	Street 22	Linh Dong Street	0.4	2400	120			x
49	Street 26	House No. 10	Linh Dong Street	0.4	-	120		x	
50	Street 30	Street 22	House number 47	0.5	-	120			x
51	Alley 34 - Street 36	Street 34	People's houses	0.3	240	30		x	
52	Tam Tam Xa Street	Duong Van Cam	At the end of the route	0.5	880	90			x
53	42nd Street	Tam Binh	People's houses	0.3	4050	24		x	
54	Alley 384 Pham Van Dong	Pham Van Dong	Street 26	0.5	960	30		x	
55	Alley 2 Street 13	Street No. 13	At the end of the route	0.4	750	60		x	
56	Street 8	Street No. 8	At the end of the route	0.5	700	60		x	

No.	Street	Flooding range		Average flooded depth (m)	Average flooded area (m^2)	Average flooding time (minutes)	Causes of flooding		
		From	To				Tide	Precipitation	Tide & precipitation
57	Cau Xay Street	Alley 41	Alley 15	0.3	1440	40		x	
58	Dinh Phong Phu Street	Street No.6	End of the Phong Phu Bridge	0.3	700	40		x	
59	Alley 99 Street 11	House No. 99/9A	At the end of the route	0.3	150	60		x	
60	Alley 26 Tu Xuong street	House No. 26B/2F	At the end of the route	0.4	150	45		x	
61	Duong Dinh Hoi Street	Street 359	Street 10	0.2	1600	40		x	
62	Lien Phuong Street	Tang Nhon Phu Street	Alley 2	0.3	600	45		x	
63	Street 63	Nguyen Duy Trinh Street	Project Zone Binh's population Trung Dong area 19.3ha	0.5	650	40-60		x	

3.1.1. The flooding in the period of five years from 2010-2015

Statistical results from 2010 to 2015 show that local flooding occurred on a series of streets such as Quoc Huong, Thao Dien, Nguyen Van Huong, Le Van Viet, ... (the total of 37 heavily flooded street locations), as a result, not only the traffic situation becomes congested, but also thousands of people face difficulties due to flooding. The number of times flooded from 2010 to 2015 and the average depth of flooding during the 2010-2015 period is described in Figure 5 and Figure 6, respectively.

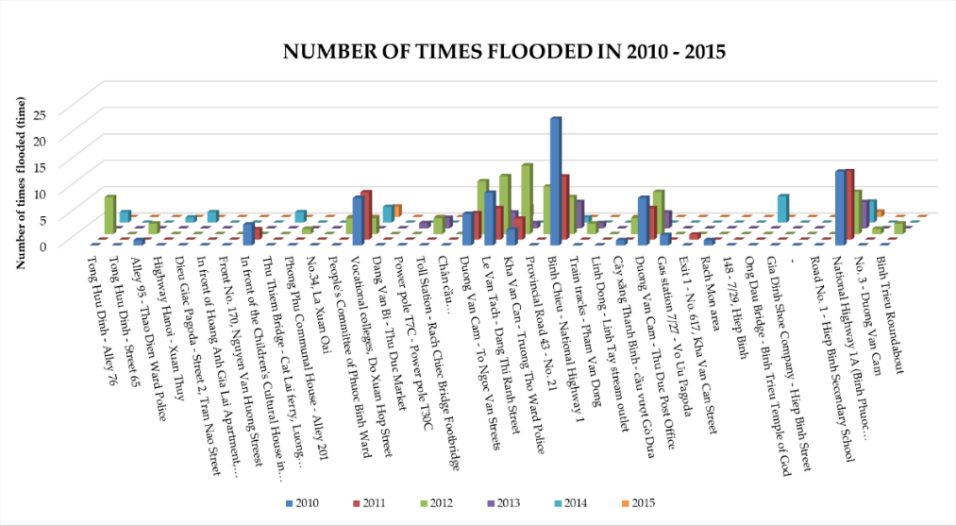


Figure 5. The number of times flooded from 2010-2015.

Particularly, Quoc Huong Street (the segment from Tong Huu Dinh Street to alley 76) was heavily flooded due to heavy precipitation from 2012 to 2014 and doubled in 2014 with average flooding of 80 minutes (Figure 5). Meanwhile, the sewer system was blocked leading to Dang Thi Ranh Street (from Duong Van Cam to To Ngoc Van Street) being flooded 10 times in 2012 (Figure 5). The segment of Duong Van Cam Street (from the intersection of Le Van Tach and Dang Thi Ranh) was regularly flooded due to heavy precipitation over the years, with the highest number of floods in 2012 (11 times) (Figure 5), at the average depth of 0.3 m flooding (Figure 6). Similarly, Ho Van Tu Street (the segment of the Kha Van Can Street to the police station of Truong Tho ward) was flooded 13 times (Figure 5) in 2012 with an average recorded depth of 0.22 m (Figure 6).

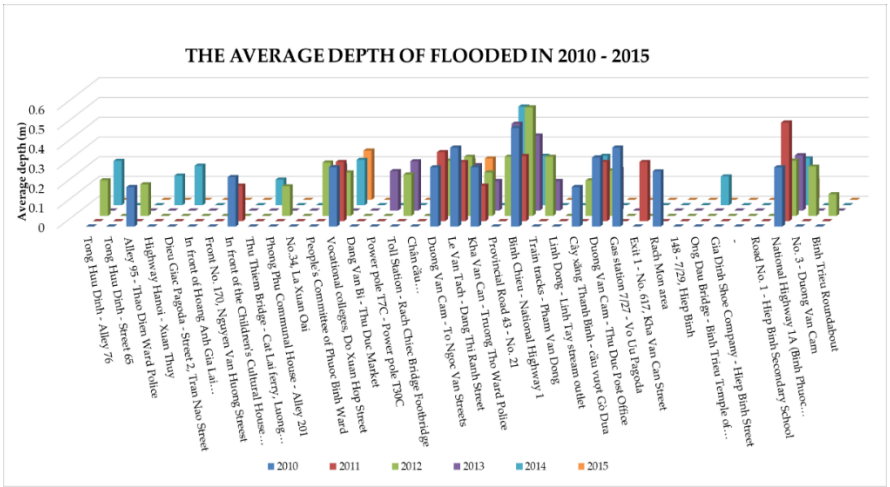


Figure 6. The average depth of flooding in the period 2010-2015.

It is noticeable that due to the deterioration of the drainage system, heavy rain has flooded some locations to an alarming level. TL43 street (the segment from Binh Chieu to National Highway 1) was an alarming flood route, with the highest frequency of flooding up to 24 times in 2010 (the average depth of 0.5 m) (Figure 5 and Figure 6). A decreasing trend of flooding depth was recorded in the street during the 2011-2015 period, but the highest frequency of flooding experienced was very high, compared to other streets. Additionally, in the segment of Go Dua street (from the Binh Phuoc overpass to the intersection of To Ngoc Van street), which is the most flooded place in the years 2010-2012, the number of flooding times was quite high, with approximate 8-14 times in 1 year (Figure 5). The statistical data shows that the average recorded depth of flooding in the street was up to 0.5 m (Figure 6).

According to the flooding due to the combination of heavy precipitation and high tide, the area in front of the Vocational College (Do Xuan Hop street) had the highest number of floods in 2010 and 2011 (9 times) (Figure 5). Heavy precipitation and high tide caused flooding in this area; in addition, the drainage system did not respond to the current situation, so flooding occurred frequently.

3.1.2. The flooding in the period of five years from 2016-2021

Compared with the flooding in the previous period, an improvement was noted in that the number of times flooding has decreased significantly, but the frequency of flooding has increased. The number of flooded times and the average depth of flooding in the streets in the period of five years from 2016 to 2021 are illustrated in Figure 7 and Figure 8.

The segment of Quoc Huong Street (Tong Huu Dinh Street - Street 65) has been flooded since 2016. The largest number of recorded floods in the whole period is 26 times (Figure 7), with the average depth of flooding of 0.23 m (Figure 8). The average flooded length was 135 m with an average flooded area of 1111 m^2. The frequency of the flooded location decreases gradually, at 5 times in 2021 (Figure 7) with a flood length of nearly 150 m and a recorded flooded area of 1400 m^2.

Although there is an operating sewer system, however, the segment of Thao Dien street (Alley 95 to Thao Dien ward police station) was still flooded sometimes up to 0.3 m in 2020 (Figure 8) with an average flood length of 270 m^2 and the average flooded area of 2700 m^2. High tide is the main cause of flooding in the area. Furthermore, it has directly affected people's lives and activities around the area. It is the cause of environmental pollution, food safety, and hygiene, obstructing traffic, and affecting the movement of people.

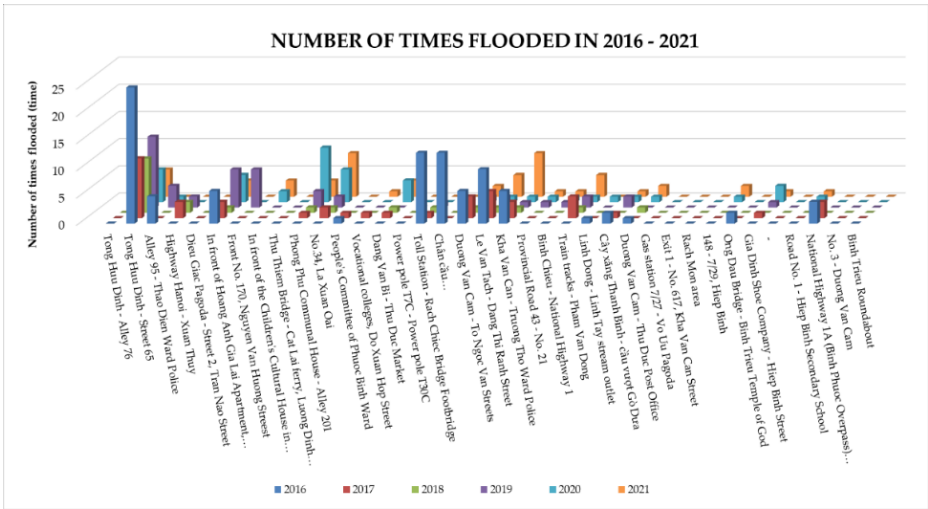


Figure 7. The number of times flooded from 2016-2021.

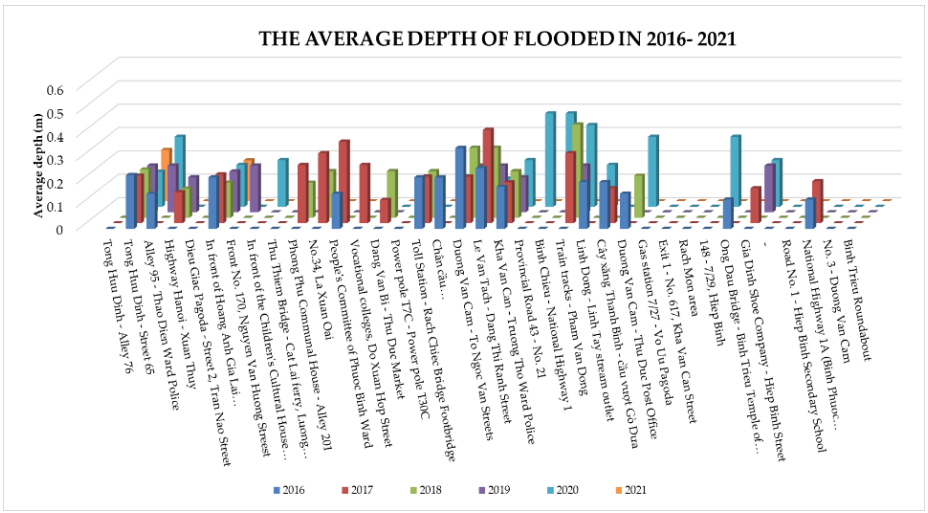


Figure 8. The average depth of flooding in the period 2016-2021.

3.2. The causes of flooding

3.2.1. Natural causes

Natural causes include topography, heavy precipitation, and tidal current. Studies on heavy precipitation and tidal current have been studied on many topics and projects. Therefore, in this study, the topography of the area is studied in more detail.

Thu Duc city has a rich and diverse topography, with an average height from 5m to 25m. There are 3 types of topography in this area including (Figure 2).

The elevation varies from 2.0 m to 30.0 m, concentrated in the north of the city, where it is adjacent to Binh Duong province. This place is not flooded by the tide but flooded by heavy precipitation, particularly Vo Van Ngan Street, To Ngoc Van Street...

The elevation varies from +1.5 m to +2.0m distributed in the south of the city such as An Khanh and Thao Dien wards... This place is flooded due to the combination of heavy precipitation and tides.

The elevation below 1.5m is distributed inwards along the Saigon and Dong Nai rivers. Typically, Hiep Binh Phuoc ward has an elevation of mainly 0-1 m and a part of 1-2 m while Hiep Binh Chanh ward has a terrain of 0-1m, accounting for about 50%. Furthermore, Thu Thiem ward has an elevation of 50% 0-1 m and 50% 1-2 m. Additionally, Long Phuoc and Thanh My Loi wards have mostly 0-1m of elevation.

3.2.2. Human causes

This study focuses on the causes related to management and planning in the region, including flood control and drainage planning

(1) Flood control planning

Flood control planning has been approved but not be completed until 2020. For planning 1547: approved by the Prime Minister in Decision 1547/2008 to propose irrigation solutions to solve the flooding situation in Ho Chi Minh City, specifically: Period to 2012: implementation of tidal control solutions, actively lowering the water level on the axis canals surrounding the right bank of the Saigon - Nha Be river. The period after 2012: implementing control measures in the confluence of Dong Nai - Saigon rivers. However, at present, there are no specific plans and guidelines for implementation. For planning 752: Master plan on the drainage of HCMC until 2020 funded by JICA, approved by the Prime Minister in 2001 in Decision No. 752/QĐ-TTg. However, this project only focuses on solutions to prevent flooding due to precipitation in the central area of Ho Chi Minh City.

(2) Drainage planning

Urbanization of the city has reduced the natural regulation of the watershed surface. When most of the land is concreted, and plasticized to build houses, factories, etc., there will be an increase in surface runoff to drainage locations. However, two prominent problems cause flooding: Due to lack of sewer or no sewer; overloaded drains, and due to urbanization and leveling of drainage sources. Many drainage projects have not been completed on schedule due to the following reasons: lack of capital, problems in compensation and clearance, relocation of underground works, problems in procedures, lack of construction contractors who can complete the project. According to the drainage planning map in District 2, District 9, and Thu Duc District by 2020, several flooded locations due to no drainage system, lack of drainage systems, and unresponsive drainage systems have not yet been planned, specifically:

In District 2: Locations 1, 2, 6, 7, and 41 (Figure 4 and Table 1) were flooded due to heavy precipitation but there is no drainage plan until 2020.

In District 9: Locations 56, 57, 59, and 62 (Figure 4 and Table 1) were also flooded due to heavy precipitation but there was no planned drainage system.

In Thu Duc District: Most of the flooding points due to precipitation are because the drainage system is degraded and there is no sewer system and up to now, there is no planned sewer system for these areas. Locations: 14, 19, 20, 27, 29, 30, 35-38, 40-44, 46-49, and 52-55 (Figure 4 and Table 1) are flooded locations due to drainage but there is no drainage system.

Thus, the slow implementation of the plans and the lack of implementation guidelines have affected the current flooding situation in the study area. Besides, the lack of synchronization in drainage planning makes this situation more serious. Flooding in District 2 and Thu Duc District are the prime examples.

The main causes identified are quite consistent with the primary causes of flooding in Ho Chi Minh City according to the research by [35].

Thu Duc City is aiming to become a green and smart city, and accordingly, the appropriate flood control strategies are as follows:

1. Implement rapid and coordinated flood prevention planning.

2. Hierarchical risk zoning: The hierarchical risk zoning is based on three criteria that constitute flood risk, including the level of the flooding event, vulnerability, and exposure. These criteria are then combined with input data, including flood assessment maps for extreme flooding events, such as the combination of heavy rainfall, tide, flood and climate change, and maps of population distribution and current infrastructure to identify areas with high, moderate, low, and no flood risk.

3. Tidal control corridors: Unlike traditional approaches using embankments or dikes, tidal control corridors help create integrated spaces for multiple infrastructure solutions to serve the goal of flood and salinity control. Tidal control corridors can narrow around areas with low foundations instead of covering an entire large area like traditional embankments. Additionally, to seal off tidal control corridors, it is necessary to construct tidal gates at canal entrances in the flood control corridors.

4. Reservoirs are the main solution for storing overflow caused by rain when the system cannot drain water out due to high river water levels.

4. Conclusions

Flooding in Thu Duc City has been happening for many years now and every year, there are some new flood spots in urbanizing areas. In general, tidal flooding only occurs in some areas, flooding due to heavy precipitation is a common form of flooding in Thu Duc City. Heavy precipitation combined with high tide caused some streets in the central area of District 2 and Thu Duc with low foundations to be frequently flooded. New flooding points appeared mainly in the Thu Duc District and a few points in the west of District 9.

Compared to the drainage planning map in District 2, District 9, and Thu Duc District by 2020, several flooded locations due to no drainage system, lack of drainage systems, and unresponsive drainage systems have not yet been planned. Especially in the Thu Duc District, almost all locations flooded caused of heavy precipitation but there was no drainage system.

The main reasons for such flooding are the slow implementation of the master plans, the lack of implementation guidance, and the lack of synchronization in the drainage planning which makes this situation more serious. However, this study has not yet classified flood risk zones, an essential premise for flood control strategies in the study area.

In addition, forecasting flooding corresponding to climate change scenarios for future planning has not been undertaken. These are limitations and represent the next research following this study.

Author Contributions: Conceptualization, N.T.P.T., T.T.K.; Methodology, N.T.P.T., T.T.K., N.K.P.; Formal analysis, N.T.P.T., T.T.K.; Investigation, N.T.P.T., T.T.K., N.K.P.; Resources, N.T.P.T., T.T.K.; Data curation, N.T.P.T.; Writing–original draft preparation, N.T.P.T., T.T.K.; Writing–review and editing, N.T.P.T., T.T.K.; Visualization, N.T.P.T., T.T.K.; Supervision, N.K.P.; Project administration, N.K.P.; Funding acquisition, N.K.P.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lian, J.; Xu, K.; Ma, C. Joint impact of rainfall and tidal level on flood risk in a coastal city with a complex river network: a case study of Fuzhou City, China. *Hydrol. Earth Syst. Sci.* **2013**, *17(2)*, 679–689.
2. Huong, H.T.L.; Pathirana, A. Urbanization and climate change impacts on future urban flooding in Can Tho city, Vietnam. *Hydrol. Earth Syst. Sci.* **2013**, *17(1)*, 379–394.
3. Chen, Y.; Zhou, H.; Zhang, H.; Du, G.; Zhou, J. Urban flood risk warning under rapid urbanization. *Environ. Res.* **2015**, *139*, 3–10.
4. Nguyen, H.D.; Fox, D.; Dang, D.K.; Pham, L.T.; Viet Du, Q.V.; Nguyen, T.H.T.; Dang, T.N.; Tran, V.T.; Vu, P.L.; Nguyen, Q.H.; et al. Predicting future urban flood risk using land change and hydraulic modeling in a river watershed in the central province of Vietnam. *Remote Sens.* **2021**, *13(2)*, 262.

5. Park, K.; Lee, M.H. The development and application of the urban flood risk assessment model for reflecting upon urban planning elements. *Water* **2019**, *11(5)*, 920.
6. Jha, A.K.; Bloch, R.; Lamond, J. Cities and flooding: a guide to integrated urban flood risk management for the 21st century. World Bank Publications, 2012.
7. Bertilsson, L.; Karin Wiklund, K.; de Moura Tebaldi, I.; Rezende, O.M.; Veról, A.P.; Miguez, M.G. Urban flood resilience–A multi-criteria index to integrate flood resilience into urban planning. *J. Hydrol.* **2019**, *573*, 970–982.
8. Chan, F.K.S.; et al. "Sponge City" in China–a breakthrough of planning and flood risk management in the urban context. *Land Use Policy* **2018**, *76*, 772–778.
9. Wang, B.; et al. Urban resilience from the lens of social media data: Responses to urban flooding in Nanjing, China. *Cities* **2020**, *106*, 102884.
10. Cao, A.; et al. Future of Asian Deltaic Megacities under sea level rise and land subsidence: current adaptation pathways for Tokyo, Jakarta, Manila, and Ho Chi Minh City. *Curr. Opin. Environ. Sustain*. **2021**, *50*, 87–97.
11. Nicholls, R.J.; et al. Ranking port cities with high exposure and vulnerability to climate extremes: exposure estimates. 2008: OECD Environment Working Papers.
12. Hanson, S.; et al. A global ranking of port cities with high exposure to climate extremes. *Clim. Change,* **2011**, *104(1)*, 89–111.
13. Ngoc, T.T.; et al. Ho Chi Minh City growing with water-related challenges. Water, Megacities Global Change. 2016.
14. Arnbjerg-Nielsen, K.; et al. Impacts of climate change on rainfall extremes and urban drainage systems: a review. *Water Sci. Technol*. **2013**, *68(1)*, 16–28.
15. Vachaud, G.; et al. Flood-related risks in Ho Chi Minh City and ways of mitigation. *J. Hydrol.* **2019**, *573*, 1021–1027.
16. Tu, T.T.; Nitivattananon, V. Adaptation to flood risks in Ho Chi Minh City, Vietnam. *Int. J. Clim. Change Strategies Manage*. **2011**, *3(1)*, 61–73.
17. Ho, L.P.; et al. Integrated urban flood risk management approach in context of uncertainties: Case study Ho Chi Minh city. *La Houille Blanche* **2014**, *6*, 26–33.
18. Binh, L.T.H.; Umamahesh, N.; Rathnam, E.V. High-resolution flood hazard mapping based on nonstationary frequency analysis: case study of Ho Chi Minh City, Vietnam. *Hydrol. Sci. J.* **2019**, *64(3)*, 318–335.
19. Dahm, R.; Diermanse, F.; Phi, H.L. On the flood and inundation management of Ho Chi Minh City, Vietnam. Proceeding of the International Conference on Flood Resilience: Experiences in Asia and Europe. 2013.
20. Storch, H.; Downes, N.K. A scenario-based approach to assess Ho Chi Minh City's urban development strategies against the impact of climate change. *Cities* **2011**, *28(6)*, 517–526.
21. Downes, N.K.; et al. Understanding Ho Chi Minh City's urban structures for urban land-use monitoring and risk-adapted land-use planning, in Sustainable Ho Chi Minh City: Climate Policies for Emerging Mega Cities. Springer. 2016, pp. 89–116.
22. Erkens, G.; et al. Sinking coastal cities. *Proc. Int. Assoc. Hydrol. Sci.* **2015**.
23. Thiep, T.H.; Truong, N.X. Inundation in Ho Chi Minh City: current situation, cause and solutions. *Int. J. Res.* **2021**, *2(7)*, 27–33.
24. Leitold, R.; et al. Flood risk reduction and climate change adaptation of manufacturing firms: Global knowledge gaps and lessons from Ho Chi Minh City. *Int. J. Disaster Risk Reduct.* **2021**, *61*, 102351.
25. Santos, M.d.C.O. Enabling water sensitive urban design principles in Ho Chi Minh City for flooding resilience. in Reframing urban resilience implementation: 11th International Forum on Urbanism Congress. 2018.

26. Xuan, P.T.H.; Nhut, N.M. Suggestion from experience of some asian countries regarding "City-inside City": A case study of Thu Duc City of Ho Chi Minh City. Proceeding of the 18th International Symposium on Management (INSYMA 2021). 2021.
27. Camenen, B.; et al. Monitoring discharge in a tidal river using water level observations: Application to the Saigon River, Vietnam. *Sci. Total Environ.* **2021**, *761*, 143195.
28. Ho, L. Impacts of climate changes and urbanisation on urban inundation in Ho Chi Minh City. Proceeding of the 11th international conference on urban drainage, Edinburgh, Scotland, UK. 2008.
29. Mark, O.; et al. Potential and limitations of 1D modelling of urban flooding. *J. Hydrol.* **2004**, *299(3-4)*, 284–299.
30. Guo, K.; Guan, M.; Yu, D. Urban surface water flood modelling–a comprehensive review of current models and future challenges. *Hydrol. Earth Syst. Sci.* **2021**, *25(5)*, 2843–2860.
31. Luo, P.; et al. Urban flood numerical simulation: Research, methods and future perspectives. *Environ. Modell. Software* **2022**, *156*, 105478.
32. Lin, L.; Wu, Z.; Liang, Q. Urban flood susceptibility analysis using a GIS-based multi-criteria analysis framework. *Nat. Hazard.* **2019**, *97*, 455–475.
33. Tayyab, M.; et al. Gis-based urban flood resilience assessment using urban flood resilience model: A case study of Peshawar city, Khyber Pakhtunkhwa, Pakistan. *Remote Sens.* **2021**, *13(10)*, 1864.
34. Huong, N.T.T. Status quo and traffic management measures dealing with flooded roads due to adverse weather for District 9 in Thu Duc City, HCMC. *Transport Commun. Sci. J.* **2022**, *73(5)*, 486–501.
35. Giam, N.M.; et al. The main causes of flooding in Ho Chi Minh City. *VN J. Hydrometeorol.* **2023**, *747*, 21–36. (In Vietnamese)

JOURNAL OF
HYDRO-METEOROLOGY

Research Article

Assessing the level of air pollution at some small-scale household waste incinerators in Hai Hau district, Nam Dinh province

Nguyen Tien Dung[1]*, Vu Duc Toan[2]*, Ngo Tra Mai[3], Nguyen Ngoc Anh[3]

[1] Institute of Environmental Science and Technology, Vietnam Cooperative Alliance; nguyentiendunghb78@gmail.com
[2] ROOM strong research, Environmental and life science research laboratory, Thuyloi University; vuductoan@tlu.edu.vn; huongntl@tlu.edu.vn
[3] Institute of Physics, Vietnam Academy of Science and Technology; ntmai@iop.vast.vn; ann@iop.vast.vn

*Corresponding author: nguyentiendunghb78@gmail.com; vuductoan@tlu.edu.vn; Tel.: +84–914953335

Received: 12 October 2023; Accepted: 27 November 2023; Published: 25 December 2023

Abstract: In this study, sixteen air samples from two small-scale household waste incinerators in Hai Hau district, Nam Dinh province, Vietnam, were collected and investigated for some pollutants such as SO_2, NO_2, CO, H_2S, and BTEX. The results showed that the concentration of SO_2, NO_2, CO, and H_2S were correspondingly in the range of 351.1 -450.0, 30300-36600, 200.7-212.0, and 78.9-100.1 µg/m^3. In addition, the BETX concentration ranged from 5.16 to 20.7 µg/m^3 for Benzen, from 21.63 to 120.66 µg/m^3 for Toluen, from 3.13 to 12.57 µg/m^3 for (m, p)-Xylene, and from 1.58 to 7.73 µg/m^3 for o-xylene. All the values of SO_2, NO_2, CO, and H_2S concentration were higher than standard concentrations in QCVN 05: 2013/BTNMT, however, the BETX concentration was lower than the standards in QCVN 06:2009/BTNMT.

Keyword: BTEX; Concentration of pollution; Small scare incinerators.

1. Introduction

Waste generation has been increasing as long as society has developed. In 2016, the total amount of waste included into factor all commercial activity and households was 2538 million tons in the European Union (EU) based on the figure [1]. To reduce the amount of waste, the EU has been guiding the European environment policy from 2014 to 2020 and vision in 2050 in which waste treatment by incineration was limited [2]. Nevertheless, regardless of these policies, a significant amount of incinerators have continued to operate [3].

The main reason why the EU has recommendations to minimize waste treatment by incinerators is due to the negative effects on human health. A number of toxins that are created during the incineration process or are found in the trash have an adverse effect on both humans and the environment [4]. The cancer incidence of residents near municipal solid waste incinerators was studied by [5], and the findings indicated a statistically significant increase in the risk of lung, colorectal, and stomach cancer. According to epidemiological data, there is evidence linking air pollution, particularly particulate matter (PM), to an increase in respiratory and cardiovascular-related issues. According to observations, between 40% and 70% of airborne fine particulate matter originates from combustion. This type of particulate matter has been linked to a decline in lung function and an increase in the prevalence of chronic obstructive pulmonary disease [6].

In the air environment, many harmful pollutants, particularly particulate matter (PM), volatile organic compounds (VOCs), heavy metals, and polychlorinated dibenzofurans and p-dioxins (PCDD/Fs), can be released by waste treatment incinerators [7]. The study [8] showed that PCDD/F concentrations were found in indoor air (0.29-0.85 pg/m^3) and indoor dust (14.93-649.70 ng/kg) samples taken from residences nearby to an East China municipal solid waste incineration (MSWI) plant. Study [9] reported that when sixty-four samples from four MSWI power plants were tested, 102 target volatile organic compounds (VOCs) were found, including o-xylene, toluene, ethylbenzene, etc. According to [10], 25 VOC species were found in the chimney plume and 29 VOC species were found in the discharge workshop. A high temperature of 1000^{o}C was used throughout the incineration process to decompose most of the organic materials in the rubbish, releasing CO, CO_2, VOCs, and other gases from the stack. Toluene was a typical byproduct of waste combustion, but benzene can be generated by the recombination of C_3H_3 and the interaction of C_3H_3 and other species during the burning process.

In Vietnam, according to the National Environmental Status Report 2019 on solid waste, there are currently 1,322 solid waste treatment facilities, including 381 incinerators [11]. Some small-sized incinerators for MSW treatment, many of these incinerators do not meet the requirements of QCVN 61-MT:2016/BTNMT [12]. Some incinerators are damaged and degraded after a period of operation. Some incinerators meet the requirements of QCVN 61-MT:2016/BTNMT, but when applied in localities, they encounter some problems such as MSW having low calorific value, high humidity, and the operating level of the companies. Factories are still weak, do not comply with technical requirements, or do not operate the exhaust gas treatment system, leading to uncontrollable waste generation. [13] designed a waste treatment system using incineration with high-temperature combustion systems (900-1100^{o}C) to ensure solid waste burns completely and does not produce gases. poison. The laboratory-scale incinerator with a capacity of 5kg/h is tested and the exhaust quality is evaluated. The results show that the exhaust gas after combustion was surveyed with absorption solutions, with a NaOH concentration of 0.5N, and the concentration of NO_2, CO, and SO_2 emissions met the requirements according to QCVN61-MT:2016/BTNMT. However, to our knowledge, studies specifically evaluating air pollution in small-scale waste incinerators have not been evaluated in Vietnam or abroad.

Hai Hau district, Nam Dinh province is a district with many attractive tourist attractions, so it has attracted many people from all over the province as well as other provinces, so the amount of waste generated daily in the area is quite large. To solve the waste problem, the District People's Committee has focused on investing in small-scale waste incinerators. According to statistics from the People's Committee of Nam Dinh province, currently throughout Hai Hau district, there are 29 domestic solid waste incinerators, including 20 Losiho technology incinerators, with a capacity of 400 kg/hour using fuel natural gas [14].

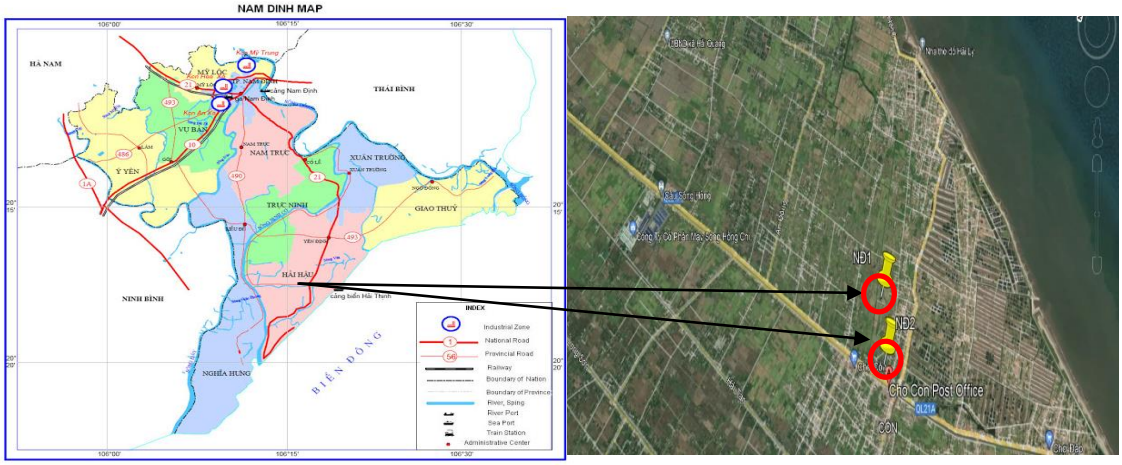


Figure 1. Location of the study area.

In this study, we focus on assessing air pollution at two small-scale waste incinerators in Hai Hau, Nam Dinh province.

2. Materials and methodology

2.1. Study area and sampling

2.1.1. Study area

The study area in Nam Dinh was selected in Con Town (NĐ1) and Hai Ly commune (NĐ2), Hai Hau district (Figure 1).

2.2. Methodology

2.2.1. Investigation and survey

Conducted investigation and survey of 29 waste incinerators in Hai Hau district, Nam Dinh province. Based on the survey results, the plementation team selected 02 incinerators with technology and waste composition representative of the research area to conduct sampling and analysis.

2.2.2. Sampling

A total of 16 samples were collected from two incinerators in Hai Hau district. Repeating sampling time is September, and November 2022 at each location of the waste incinerator. Samples were taken according to the instructions in Circular No. 10/2021/TT–BTNMT on technical regulations for environmental monitoring and management of information and data for environmental quality monitoring [15]. Basic parameters such as: SO_2, NO_2, CO, H_2S and TSP samples are taken with a Gas Sampling Pump: 1-5 LPM Model: 224- PCXR8KD SKC - USA, measuring range from 5-5000 lm/min.

BTEX group parameters: taken with Japanese Shibata HV 500R 5Y0164 gas sampling device, measuring range 100-800 L/min.

2.2.3. Study methods

SO_2: Analysis According to TCVN 5971:1995 - Ambient air - Determination of the mass concentration of sulfur dioxide - Tetrachloromercurate (TCM) pararosaniline method [16].

NO_2: Analysis According to TCVN 6137:2009 - Ambient air - Determination of mass concentration of nitrogen dioxide - Modified Griess- Saltzman method [17].

CO: Analysis According to Industry standard 52TCN 352:1989 on Carbon Oxides [18].

H_2S: Analysis According to Standard MASA-701 - Method for sampling H_2S gas [19].

TSP: Analysis According to TCVN 5067-1995: Air quality - Weighing method for determination of dust content [20].

BTEX group parameters: In compliance with NIOSH Manual of Analytical Method 1501, BTEX air sampling was performed. Charcoal sorbent tubes were used for BTEX adsorption (SKC, USA) [21]. After being collected, every sample was kept at -10°C and examined within 72 hours. Analytical reagents are all substances. After obtaining the BTEX standard, CS2, and n-hexane from Merck KgaA, Darmstadt, Germany, it was utilized without any purification procedures. The BTEX chemicals were extracted from gently shaking activated carbon with 2 mL of carbon disulfide (CS2) for 60 minutes using a vial containing 2 mL of CS2. After that, the solvent was put into a GC vial, and BTEX chemicals were measured using gas chromatography (GC/MS) with a mass spectrometer detector (Thermo Scientific Model IQS 7000). There was one capillary column utilized, the DB-WAX (30 m × 0.25 mm i.d × 0.25 µm thickness). The BTEX technique detection limits varied between 4.1 and 12.3 ng/m^3. Nearly every sample exceeded the limits of detection. For 10% of the

samples that were collected, double analyses were carried out, and the results showed that the relative standard deviation values for each BTEX compound ranged from 5% to 9%. The back section of the tubes was separately analyzed in order to control the breakthrough of the sorbent tubes regarding BTEX compounds; the results showed that none of the target compounds were significantly detected in the back section for all of the samples analyzed (15 samples chosen at random from real samples).

3. Results and discussion

3.1. Assess the current status of pollution some basic parameters around the incinerator area of SSI

To assess the current status of air pollution in the area surrounding household waste burning, the research team conducted two rounds of sampling and sample analysis on September 8, 2022 and November 9, 2022, the analysis results are presented in Tables 1.

Table 1. Results of analysis of some basic indicators in the air environment around the incinerator area.

Compound	Unit	September	November	QCVN 05:2013/BTNMT [22]
A. Nam Dinh 1				
TSP	µg/m^3	314.7 - 335.2 (322.4 ± 8.1)	312.8 – 332.4 (320.1 ± 7.7)	300
SO_2	µg/m^3	445.1 – 450.0 (447.0 ± 1.9)	441.2 – 447.1 (443.4 ± 2.3)	350
CO	µg/m^3	31000 – 36600 (33100 ± 2113.6)	30450 - 37000 (32906.2 ± 2589.1)	30000
NO_2	µg/m^3	201.5 – 212.0 (205.4 ± 4.1)	200.7 – 210.7 (204.4 ± 2.9)	200
H2S	µg/m^3	89.2 – 100.1 (93.3 ± 4.1)	78.9 – 97.4 (85.8 ± 7.3)	42
B. Nam Dinh 2				
TSP	µg/m^3	317.3 – 322.7 (319.3 ± 3.1)	315.7 – 321.5 (317.9 ± 3.3)	300
SO_2	µg/m^3	351.3 – 357.1 (353.5 ± 2.3)	351.1 – 354.7 (352.4 ± 1.4)	350
CO	µg/m^3	30380 – 30400 (30387.5 ± 7.9)	30300 – 30800 (30487.5 ± 197.6)	30000
NO_2	µg/m^3	207.1 – 210.5 (208.4 ± 1.3)	205.4 – 208.4 (206.5 ± 1.2)	200
H_2S	µg/m^3	91.4 – 98.5 (94.1 ± 2.8)	89.5 – 97.4 (92.3 ± 3.0)	42

The treatment system includes a dust collection silo, a cooling silo, an absorption silo and 04 water tanks and 01 clear lime water tank. For this treatment technology, the concentration of common gases such as CO_2, SO_2, NO_2,... after going through the treatment system is significantly reduced. Currently, Losiho incinerator is using wet gas treatment technology. The treatment system includes a dust collection silo, a cooling silo, an absorption silo and 04 water tanks and 01 clear lime water tank. For this treatment technology, the concentration of common gases such as CO_2, SO_2, NO_2,... after going through the treatment system is significantly reduced. However, currently, the operation and maintenance of the incinerator exhaust gas treatment system has not been carried out regularly, leading to low treatment efficiency, leading to quite high concentrations of pollutants after treatment.

CO was the most abundant species, followed by SO_2, TSP, NO_2 and H_2S in all samples.

In Nam Dinh 1, the concentration of CO, SO_2, TSP, NO_2 and H_2S was in the range of 30450-37000, 441.2-450.0, 351.1-357.1, 315.7-322.7 and 78.9-100.1 µg/m^3; while in Nam Dinh 2 were 30300-30800, 351.1-357.1, 315.7-322.7, 205.4-210.5 and 89.5-98.5µg/m^3, respectively.

The results show that all analytical indicators exceed the allowed standard compared to QCVN 05:2013/BTNMT, of which the largest exceeding criterion is H_2S, this indicator exceeds the allowable standard from 2.13 to 2.38 times [22].

Air samples around the incinerator area in Nam Dinh 1 are more polluted than in Nam Dinh 2. The reason may be that the composition of domestic solid waste in Nam Dinh 1 is more complex than that in Nam Dinh 2 (it contains many agricultural waste products, organic ingredients, inorganic waste, etc.).

3.2 Contamination of BTEX around the incinerator area

To assess the current status of BTEX pollution in the area surrounding household waste burning, the research team conducted two rounds of sampling and sample analysis on September 8, 2022, and November.

Table 2. Results of BTEX analysis in the air environment around the incinerator area.

Compound	Unit	September	November	QCVN 06:2009/BTNMT [23]
Nam Dinh 1				
Benzen	µg/m^3	7.62 - 12.71 (9.3 ± 2.1)	6.13 – 20.70 (11.5 ± 5.7)	22
Toluen	µg/m^3	33.92 - 53.94 (40.6 ± 6.4)	27.82 – 120.66 (62.6 ± 36.7)	500
(m,p)-Xylene	µg/m^3	6.3 – 7.66 (6.8 ± 0.5)	6.55 – 12.57 (10.7 ± 6.3)	1000
o-Xylene	µg/m^3	3.07 – 3.48 (3.1 ± 0.16)	2.00 – 4.13 (2.8 ± 0.18)	
Etyl Benzen	µg/m^3	2.95 - 4.56 (3.4 ± 0.6)	2.78 – 7.73 (4.6 ± 1.9)	-
B. Nam Dinh 2				
Benzen	µg/m^3	7.28 – 11.35 (8.8 ± 1.6)	5.16 – 16.82 (9.5 ± 4.6)	22
Toluen	µg/m^3	32.58 – 48.14 (38.4 ± 6.1)	21.63 – 95.90 (49.5 ± 29.3)	500
(m,p)-Xylene	µg/m^3	5.21 – 5.49 (7.0 ± 1.7)	3.13 – 12.54 (6.3 ± 3.6)	1000
o-Xylene	µg/m^3	2.6 – 2.74 (3.2 ± 0.5)	1.58 – 6.48 (3.3 ± 1.9)	
Etyl Benzen	µg/m^3	2.78 - 7.73 (4.6 ± 1.9)	2.45 – 6.41 (3.9 ± 1.6)	-

Analytical results of 02 sampling periods showed that Toluene concentration was the highest, followed by benzene, m,p-xylenes (p,m-X), ethylbenzene (E) and o-xylene (o-X). The concentrations of substances are summarized in Table 2.

In Nam Dinh 1, the concentration of benzene, toluene, m,p-xylene, o-xylene end Etyl Benzen was in the range of 6.13-20.7, 27.82-120.66, 6.55-12.57, 2.0-4.13 and 2.78-7.73 µg/m^3; while in Nam Dinh 2 were 5.16-16.82, 21.63-95.9, 3.13-12.54, 1.58-6.48 and 2.45-7.73µg/m^3, respectively.

Overall, the BTEX concentration in Nam Dinh 1 was considerably higher than in Nam Dinh 2. The BTEX concentration during the sampling period was highest in September (Table 2).

In Vietnam, there is currently no research on BTEX concentration pollution around household waste incinerator areas. However, we can see that the concentration of BTEX substances will depend on the furnace technology as well as the composition and properties

of the fuel. [24] showed that different combustion compositions produce different BTEX products. Survey results show that the waste that needs to be burned in the area is mainly organic (food scraps, leaves, tree branches); waste paper, glass, plastic bags, plastic bottles; concrete dishes, etc. Currently, Losiho incinerator's exhaust smoke treatment technology is only effective for common pollutants. For the BTEX group of organic compounds, it is not very effective, so most of the emissions generated by the BTEX group will be released into the local environment.

The concentration of organic substances in the BTEX group in September tends to be higher than in November. However, the concentration of the BTEX group measured in the study area is still lower than the allowed standard compared to QCVN 06:2009/BTNMT.

3.3 Ozone formation potential and isomeric ratio of BTEX

The percentage and ozone formation potential of BTEX are summarized in Table 3.

Table 3. Percentage (%), ozone formation potential (µg O_3/m^3) and isomeric ratio of BTEX.

Compound	Nam Dinh 1	Nam Dinh 2
A. Percentage of BTEX (%)		
Benzene	13.46–13.96	13.61–14.82
Toluene	62.47–69.14	64.15–66.61
Ethylbenzene	5.10–5.83	6.36–7.44
m,p-Xylene	8.4–13.13	8.16–9.76
o-Xylene	3.36–5.08	4.16–4.90
B. Ozone formation potential (µg O_3/m^3)		
Benzene	2.74–7.50 (4.5 ± 1.9) (a)	3.44–9.67 (5.8 ± 2.5)
Toluene	70.62–235.4 (132.4 ± 65.1)	97.1–344.3 (189.8 ± 97.8)
Ethylbenzene	6.35–19.0 (0.25 ± 0.11)	8.1–27.8 (15.5 ± 7.8)
m,p-Xylene	37.9–70.8 (50.2 ± 13.0)	49.2–101.6 (55 ± 21)
o-Xylene	15.10–28.6 (20.1 ± 5.3)	19.5–35.3 (25.4 ± 6.3)
Total BTEX	132.7–361.4 (218.5 ± 90.4)	177.4–518.7 (305.4 ± 134.9)
C. Isomeric ratio of BTEX		
T/B	4.64–5.04	4.33–4.97
(m,p-X)/E	1.65–2.25	1.33–1.57
(m,p-X)/B	0.6–0.9797	0.61–0.66
E/B	0.31–0.48	0.21–0.54
o-X/B	0.24–0.38	0.43–0.53

(a): min – max (mean±standard deviation)

The summary results in table 3 show that Toluene accounts for the highest proportion, ranging from 62.47-69.14%, followed by Benzene with a ratio of 13.46-14.82%, m,p-Xylene with a ratio of 37.9-101.6%, Ethylbenzene accounts for 5.10-7.44% and o-Xylene accounts for 3.36–5.08%.

Among BTEX species, Toluene was the largest contributor to ozone formation, followed by m,p-Xylene, o-Xylene, Ethylbenzene and the lowest was benzene (Table 3). In Vietnam, there is no announcement on the analysis rate of BTEX in the air around the incinerator area, however, if compared with the rate of BTEX in traffic locations, this rate is equivalent.

The average T/B ratio ranged from 4.33 to 5.04 in this study, which is higher than 0.7-1.3 in Hanoi [25], 0.63-4.71 in Texas, USA [26], 1.3-3.0 in Windsor, Canada [27] and 1.29-2.45 in Cairo [25]. The above-mentioned values were lower than values commonly found in developed countries such as 6.4-8.5 in Tokyo and 9.2-11.5 in Hong Kong. This study's (m,p-

X)/E values ranged from 1.33 to 2.25, in the same range in Hanoi [25]. They were remarkably lower than the E/B values of this study (from 0.35 to 0.41).

The isomeric ratio of BTEX help us identify waste sources, the toluene to benzene ratio (T/B) was higher than 2, indicating that the fixed incinerators were the main sources.

4. Conclusion

An investigation was conducted on the air pollution at a few small-scale residential waste incinerators in the Hai Hau district. Dust samples from the Losiho incinerator turned up sixteen samples. The analysis's findings indicate that burning household garbage has had an impact on the air near the incinerator, with concentrations of CO_2 (200.7-212 µg/m^3), SO_2 (351.1-450.0), NO_x (30300-36600), H_2S (78.9-100.1) and TSS (312-317). In addition, the BETX concentration ranged from 5.16 µg/m^3 to 20.7 µg/m^3 for Benzen; from 21.63 µg/m^3 to 120.66 µg/m^3 for Toluen; from 3.13 µg/m^3 to 12.57 µg/m^3 for (m, p)-Xylene; and from 1.58 µg/m^3 to 7.73 µg/m^3 for o-Xylene. All concentration of CO_2, SO_2, NO_x, H_2S exceed the allowed standard compared to QCVN 06:2013/BTNMT, in which the H_2S indicator exceeds the allowed standard from 2.13 to 2.38 times. Currently, there is almost no assessment of air pollution as well as assessment of health effects of people in areas surrounding incinerators in Vietnam. Therefore, there is a need for more specific research and evaluation on small-scale incinerators so that there can be measures to improve and reduce pollution.

To minimize negative impacts on the environment and health of incinerator operators, the article suggests applying a number of management measures such as equipping labor protection equipment for incinerator operators regularly and periodically maintaining the incinerator system, at the same time improving and installing additional exhaust gas treatment systems for the incinerator, etc. to ensure that emissions generated during the combustion process are treated before being released into the environment.

Author Contributions: Conceptualization, N.T.D., V.D.T.; Methodology, N.T.D, V.D.T., N.T.M., N.N.A.; Formal analysis, N.T.D; Investigation, N.T.D., N.N.A.; Resources, N.T.D., V.D.T.; Data curation, N.T.D.; Writing–original draft preparation, N.T.D., V.D.T.; Writing–review and editing, N.T.D., V.D.T.; Visualization, N.T.D., V.D.T.; Supervision, N.TD.; Project administration, N.T.D.; Funding acquisition, N.T.D.

Acknowledgement:

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Eurostat, S.E. Waste statistics. 2018.
2. EC. Living Well, Within the Limits of Our Planet. Prcoceeding of the 7th EAP–The New General Union Environment Action Programme to 2020, 2013.
3. Jose L. Domingo, Montse Marquès, Montse Mari, Marta Schuhmacher. Adverse health effects for populations living near waste incinerators with special attention to hazardous waste incinerators. A review of the scientific literature. *Environ. Res.* **2020**, *187*, 109631.
4. Cangialosi, F.; Intini, G.; Liberti, L.; Notarnicola, M.; Stellacci, P. Health risk assessment of air emissions from a municipal solid waste incineration plant–a case study. *Waste Manag.* **2008**, *28(5)*, 885–895.
5. Elliott, P.; Shaddick, G.; Kleinschmidt, I.; Jolley, D.; Walls, P.; Beresford, J.; Grund, Y.C. Cancer incidence near municipal solid waste incinerators in Great Britain. *Br. J. Cancer*. **1996**, *73(5)*, 702–710.
6. Dugas, T.R.; Lomnicki, S.; Cormier, S.A.; Dellinger, B.; Reams, M. Addressing emerging risks: scientific and regulatory challenges associated with environmentally persistent free radicals. *Int. J. Environ. Res. Public Health.* **2016**, *13*, 573.

7. Bai, Y.; Guo, W.; Wang, X.; Pan, H.; Zhao, Q.; Wang, D. Utilization of municipal solid waste incineration fly ash with red mud-carbide slag for eco-friendly geopolymer preparation. *J. Clean. Prod.* **2022**, *340*, 130820.
8. Yu, J.; Li, H.; Liu, Y.; Wang, C. PCDD/Fs in indoor environments of residential communities around a municipal solid waste incineration plant in East China: Occurrence, sources, and cancer risks. *Environ. Int.* **2023**, *174*, 107902.
9. Chengyi, S.; Zhiping, W., Yong, Y.; Minya, W.; Xianglong, J.; Guoao, L.; Jing, Y.; Liyun, Z.; Lei, N.; Yiqi, W.; Yuxi, Z.; Yang, L. Characteristic, secondary transformation and odor activity evaluation of VOCs emitted from municipal solid waste incineration power plant. *J. Environ. Manage.* **2023**, *326*, 116703.
10. Liu, S.L.; Wang, B.G.; He, J.; Tang, X.D.; Luo, W.; Wang, C. Source fingerprints of volatile organic compounds emitted from A municipal solid waste incineration power plant in Guangzhou, China. *Procedia Environ. Sci.* **2012**, *12*, 106–115.
11. Ministry of Natural Resources and Environment, National environmental status report on domestic solid waste incinerator.
12. Ministry of Natural Resources and Environment. QCVN 61: 2016/BTNMT: National technical regulation on domestic solid waste incinerator.
13. Huong, P.T.M., Tung, N.Q.; Ngan, N.H.; Van, D.T.C; Huy, H.V.; Yen, P.T.T. Study on the treatment process of air pollution and ash from the solid waste incinerator. *J. Sci. Technol.* **2020**, *56*, 120–124.
14. People's Committee of Nam Dinh province. Project on management and treatment of household waste in Nam Dinh province for the period of 2020 - 2025.
15. Circular No. 10/2021/TT–BTNMT on technical regulations for environmental monitoring and management of information and data for environmental quality monitoring.
16. TCVN 5971:1995 - Ambient air - Determination of the mass concentration of sulfur dioxide - Tetrachloromercurate (TCM) pararosaniline method.
17. TCVN 6137:2009 - Ambient air - Determination of mass concentration of nitrogen dioxide - Modified Griess- Saltzman method.
18. Industry standard 52TCN 352:1989 on Carbon Oxides.
19. Standard MASA-701 - Method for sampling H_2S gas.
20. TCVN 5067-1995: Air quality - Weighing method for determination of dust content.
21. NIOSH Manual of Analytical Methods (NMAM), Fourth Edition "Hydrocarbons, Aromatic: Method 1501, Issue 3, dated 15 March 2003".
22. Ministry of Natural Resources and Environment. QCVN 05: 2013/BTNMT: National technical regulation on ambient air quality.
23. Ministry of Natural Resources and Environment. QCVN 06:2009/BTNMT- National technical regulation on some toxic substances in ambient air - QCVN 06:2009/BTNMT, 2009.
24. Puttaswamy, N.; Natarajan, S.; Saidam, S.R.; Mukhopadhyay, K.; Sadasivam, S.; Sambandam, S.; Balakrishnan, K. Evaluation of health risks associated with exposure to volatile organic compounds from household fuel combustion in southern India. *Environ. Adv.* **2021**, *4*, 100043.
25. Truc, V.T.Q.; Oanh, N.T.K. Roadside BTEX and other gaseous air pollutants in relation to emission sources. *Atmos. Environ*. **2007**, *41*, 7685–7697.
26. Raysoni, A.U.; Stock, T.H.; Sarnat, J.A.; Chavez, M.C.; Sarnat, S.E.; Montoya, T.; Holguin, F.; Li, W. Evaluation of VOC concentrations in indoor and outdoor microenvironments at near-road schools. *Environ. Pollut.* **2017**, *231*, 681–693.
27. Miller, L.; Xu, X.; Wheeler, A.; Zhang, T.; Hamadani, M.; Ejaz, U. Evaluation of missing value methods for predicting ambient BTEX concentrations in two neighbouring cities in Southwestern Ontario Canada. *Atmos. Environ*. **2018**, *181*, 126–134.

JOURNAL OF
HYDRO-METEOROLOGY

Rerearch Article

Application of EnHEBIS tool to assess economic impact due to health effects from $PM_{2.5}$ pollution – A case study at Long An province, Vietnam

Phong Hoang Nguyen[1,2], Nhi Hoang Tuyet Nguyen[1,2], Long Ta Bui[1,2*]

[1] Laboratory for Environmental Modelling, Faculty of Environment and Natural Resources, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Viet Nam

[2] Vietnam National University Ho Chi Minh City (VNU–HCM), Linh Trung Ward, Thu Duc District, Ho Chi Minh City, Viet Nam

* Corresponding author: longbt62@hcmut.edu.vn; Tel.: +84–918017376

Received: 5 October 2023; Accepted: 29 November 2023; Published: 25 December 2023

Abstract: Assessment of the effects of short-term $PM_{2.5}$ exposure on human health is one of the problems that need to be addressed within the framework of sustainable development. The goal of this study is to quantify the health-economic impacts of $PM_{2.5}$ pollution for a specific province - Long An province in Vietnam using the EnHEBIS (Environment, Public Health, and Economic Benefit Management Support Integrated System) software package. The study outcomes showed that significant health and economic effects could occur in areas with high $PM_{2.5}$ concentrations and dense population concentrations. Prominent results presented that acute exposure to $PM_{2.5}$ pollution from May to December 2018 caused 265 (95% CI: -12; 422) premature deaths, which approximately 60% of all early deaths due to all-caused respiratory diseases (RDs) with 155 (95% CI: 23; 170) cases. Corresponding to the health impacts, the economic values of Long An province also suffered a loss of around 1.15% of the total value of the gross regional product (GRDP) with about 1,270 (95% CI: -57; 2,021) billion VND, equivalent to roughly 170 million USD, along with a significant decline in the working time for the adult group. Although uncertainties have remained in this study, these results have shown the extent of economic, health, and environmental damage when $PM_{2.5}$ pollution occurs. The highlights are the basis for proposing measures to control and improve the local ambient air quality.

Keywords: Air pollution; $PM_{2.5}$; Premature mortality (PMOR); Health assessments; Economic impact.

1. Introduction

Air pollution issues have caused diverse impacts on human health, ecosystems, tourism, and climate [1–2], which is identified as a global health priority in the sustainable development goals (SDGs) related to health (Goal 3), cities (Goal 11) and energy (Goal 7) [3]. In particular, PM_{10} and $PM_{2.5}$ are among the most widely monitored and studied components of ambient air pollution [4]. Outdoor particulate matter (PM) is considered the most serious air pollutant in urban areas since it causes negative health effects and is also a cause of heart vascular diseases, respiratory irritation, as well as pulmonary dysfunction [5]. Especially, $PM_{2.5}$ is likely to easily penetrate deep into the lungs, reaching the alveoli and causing greater health risks than PM_{10} because the higher surface area per unit mass

increases the absorption capacity and condensation of toxic substances [6]. Moreover, continuous acute (short-term) exposure or long-term (chronic) exposure to $PM_{2.5}$ can increase the risk of PMOR [7–8]. The cause-and-effect relationship between $PM_{2.5}$ pollution and the deterioration of public health from acute respiratory and cardiovascular diseases to related chronic diseases has been fully documented in the medical literature [9–10].

As $PM_{2.5}$ is an extremely important factor causing ambient air quality deterioration, it has become a notable research issue in the scientific community [11–12]. Many studies have been carried out to understand the physical, chemical, and optical properties of $PM_{2.5}$, and also to determine the driving factors affecting the spatio-temporal allocation of $PM_{2.5}$ [13–14]. The health risk assessment process includes three main components, such as hazard assessment based on air pollution concentrations, individual vulnerability based on dose-response functions, and exposure probability [15]. Both available epidemiological and toxicological studies have shown that the smaller the particle size, the greater the risk of human health effects occurring [16]. Furthermore, long-term exposure to $PM_{2.5}$ pollution is able to lead to increased risk by 6% of PMOR [17]. Based on the study results by [18–19] have shown that the number of premature deaths globally tends to grow sharply from 1.4 million cases in 2010 and is expected to increase to 3.6 million cases by 2050.

Besides, these health issues have caused a considerable economic burden due to growing medical spending, increased number of lost workdays, and decline in labor supply [2]. In 2017, there were about 1.24 million premature deaths in China attributed to air pollution [20], which also resulted in economic losses of 101.39 billion USD, accounting for 0.91% of the gross domestic product (GDP) [21]. According to available statistics, severe smog events occurred in the first quarter of 2013 in China, affecting approximately 13.5% of the land area and around 800 million people [22]. It was estimated that without applying a pollution control policy in China, the PM pollution problems were likely to cause a loss of roughly 2% of the GDP value and 25.2 billion USD in health costs in the year 2030 [23].

Based on the overview of previous study outcomes, it is emphasized the need to assess not only the health effects caused by $PM_{2.5}$ exposure but also the economic losses associated with these $PM_{2.5}$-related negative health impacts. Currently, although the complex socio-economic effects of outdoor $PM_{2.5}$ exposure are worthy of consideration. Nevertheless, most of the studies are focused on large urban areas in Vietnam, such as Hanoi and Ho Chi Minh City (HCMC), whilst Long An province is a satellite locality of HCMC, rarely paid attention to. Also, there is extremely little related information and data in Long An province. Thus, we chose Long An province as the study area to evaluate distributions of $PM_{2.5}$ pollution, one of the most concerned environmental issues, using a coupled WRF/CMAQ model in the period of May to December 2018. At the same time, a module of the EnHEBIS software for estimating human health-economic impacts was applied to better understand the health impact of outdoor $PM_{2.5}$ exposure in Long An province. How do acute human health effects associated with short-term $PM_{2.5}$ exposure impact local provincial macroeconomics, as well as what are the levels of economic losses caused by $PM_{2.5}$ pollution? It is considered as an advanced tool to enable planners and decision-makers to effectively quantify population exposure to outdoor $PM_{2.5}$ pollution issues.

Thus, the overall approach to conducting this study is the combination of the Weather Research and Forecast (WRF) model, the Community Multiscale Air Quality Modeling System (CMAQ) model, and a module of human health impact estimations and economic loss rapid quantification (Module 3) of the EnHEBIS software. EnHEBIS is the main product of the Vietnam National University Ho Chi Minh City (VNU-HCM) level project of type B with code B2023-20-23. The software allows quantifying benefits of reducing

outdoor $PM_{2.5}$ pollution, public health, and economic value achieved from controlling, minimizing precursor emissions, and optimizing emissions control costs. The EnHEBIS software consists of four integrated modules, including (1) Module 1 to calculate and visualize the improvement of surface air quality under changes of controlled emission scenarios, (2) Module 2 to quantify the cost of implemented control measures to reduce precursor emissions, (3) Module 3 to estimate economic and public health benefits or losses, and (4) Module 4 to analyze and evaluate the effectiveness of the achieved benefit-cost relationship. Compared to some other models, such as AirQ+ or BenMAP-CE that merely allow the development of a single computing scenario discretely with many complex parameters which are almost difficult to access in Vietnam, EnHEBIS is likely to the setup and analysis of many different scenarios with relevant input parameters selected appropriately for Vietnam's unique conditions, especially for HCMC and its surrounding areas. Also, the EnHEBIS software has been built into the WebGIS platform and be accessed through authorized accounts without having to be installed on a computer, which is considered a main advantage. Therefore, the EnHEBIS software not only allows calculations but is also used for managing estimating databases simultaneously in a wide range of different regions.

2. Method

The research method is shown in Figure 1, which is a coupled WRF/CMAQ model system, and the quantification models of health impacts and economic losses (Module 3) of the EnHEBIS package software. The coupled WRF/CMAQ models generated data on hourly and daily mean $PM_{2.5}$ levels from May to August 2018 in the study area. Meanwhile, Module 3 of the EnHEBIS software has been built to quantify human health effects related to short- and long-term outdoor $PM_{2.5}$ exposure. PMOR and morbidity-related impacts are converted into total annual health expenditure and occupational injuries per capita due to $PM_{2.5}$ exposure. The study scope is limited to estimating and evaluating short-term health effects. A brief description is to introduce the coupled WRF/CMAQ modeling system since details of all setup processes and model inputs have been presented by [24]. The CMAQ (version 5.2.1) and WRF (version 3.8) models were applied for $PM_{2.5}$ level estimations. These models have a spatial resolution of 3.0 km × 3.0 km and a modeling setup similar to

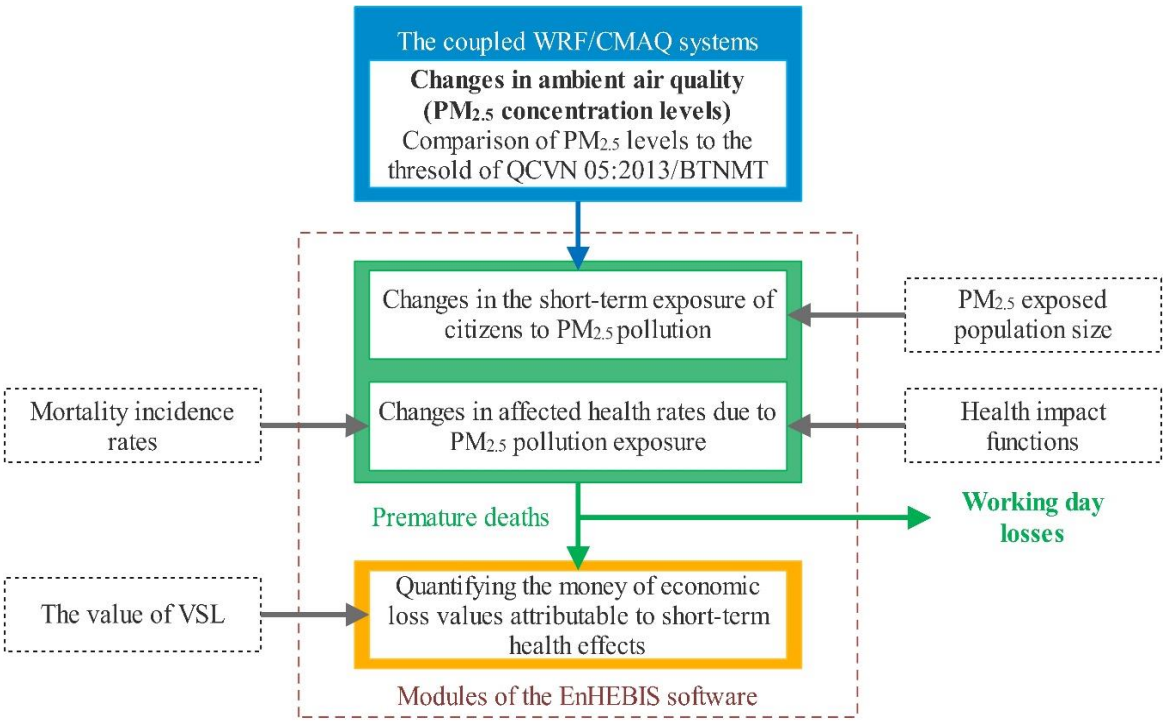


Figure 1. The diagram of research framework steps.

our previously available study [24]; in particular, WRF created meteorological conditions as input factors to CMAQ. Computing simulations in CMAQ used CB05, AER06, and the Regional Acid Deposition Model (RADM) as the gas phase mechanism, the aerosol module, and the aqueous phase chemistry module, respectively. In addition, CMAQ's boundary conditions are applied from the global model of GEOS-Chem [25] with a spatial resolution of 2.0° × 2.5°.

2.1. Study area

Long An belongs to both the Mekong Delta Region (MDR) and the Southern Key Economic Region, with geographical coordinates from 105°30'30" to 106°47'02" E and 10°23'40" to 11°02'00" N, far about 45 km from the center of Ho Chi Minh City along National Highway 1 (Figure 2). The entire study area has a natural land area of about 4,500 km^2, of which agricultural land accounts for 74% and more than 50% of the land area of Dong Thap Muoi region, including provinces of Long An, Tien Giang, and Dong Thap. Long An province's population was 1.678 million people in 2018 (Figure 2), of which about 1.406 million people live in rural areas, and 272 thousand inhabitants live in urban areas [26]. In general, the geographical location of Long An province is quite favorable for connecting MDR and the Southeast region of Vietnam. The province's economy is increasingly developing, affirming its role with an economic growth rate of 10.4% in 2018, including the sectors of agriculture, forestry and fishery rising by 4.9%, the industrial production and construction sector increasing by 15.4%, and the trade and service sector growing by 6.7% [26]. Furthermore, the province's economic structure has shifted towards industrial development, and the industry sector accounts for 47.5%. Also, Long An province ranked second among 13 localities in MDR and third in the country in terms of Provincial Competitiveness Index (PCI) in the year 2018 [26].

Acute health effects due to short-term $PM_{2.5}$ exposure are estimated across Long An province; hence, the exposed population size includes the entire population of one city and thirteen districts belonging to Long An province in 2018, as shown in Figure 2. This dataset is referenced in the Long An Statistical Yearbook for 2018, considered for all age and gender groups [26]. By 2018, the province's population contributed to around 9% of the total population of MDR, of which the districts of Duc Hoa, Can Giuoc, Can Duoc, and Ben Luc frequently have

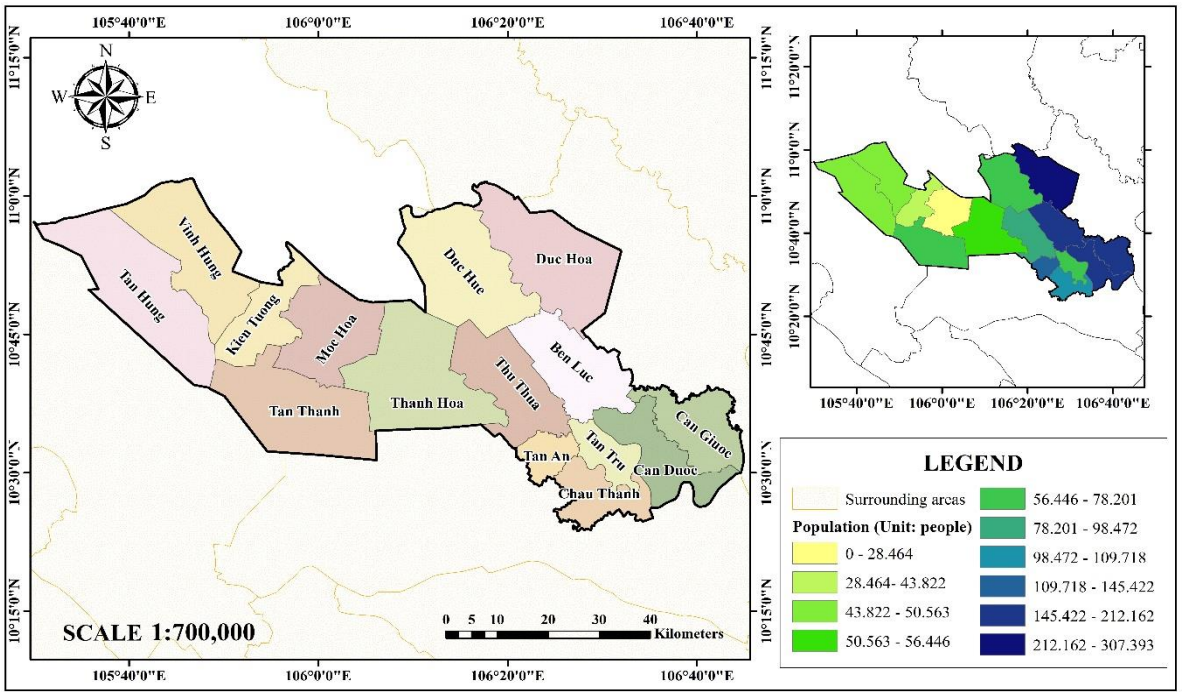


Figure 2. Geographic location (left) and district-level population size (right) of Long An province in the year 2018.

a larger population size than other districts. In particular, the population of these districts was 307,393 people (contributing to 18.31%), 212,162 people (accounting for 12.62%), 187,481 people (accounting for 11.17%), and 180,041 people (contributing to 10.72%), respectively. Nevertheless, the population density in Tan An City and the districts of Can Giuoc, Can Duoc, and Duc Hoa is the highest in Long An province, respectively being 1,779.2 people/km^2, 986,3 people/km^2, 850,3 people/km^2, and 723,1 people/km^2 [26–27].

2.2. The EnHEBIS module estimating health effects and economic losses

2.2.1. Model of health impact estimations

Short-term $PM_{2.5}$ exposure increase over time has resulted in potential health issues known as health damage, which is classified into morbidity and PMOR [23]. Table 1 lists all exposure–response functions (ERFs) (also known as health damage functions) that were used. Most studies have shown that the relative risk (RR) value for various types of health damage has a linear correlation with the pollutant concentration level [7, 28–29]. However, a few recent studies have suggested a non-linear relationship, especially when exposed to high pollution levels [30–31]. As shown in equation (1), we merely considered ERFs presenting a linear relationship with $PM_{2.5}$ exposure. There are three PMOR-related health endpoints are considered in this study, including PMOR due to all-caused diseases of respiratory (RDs), cardiovascular (CVDs), and other circulatory system (CSDs), respectively. It is assumed that ERFs are constant, specifically for the group (i) of PMOR-related health endpoints applied according to the studies of [32–34] and the group (ii) of occupational injuries due to illness based on the study by [23].

$$RR_{p,r,y,m,e,g}(C) = \begin{cases} 1, & C_{p,r,y} \le C_{0p} \\ 1 + ERF_{m,e,g}(C_{p,r,y} - C0_p), \text{linear function}, & C_{p,r,y} > C_{0p} \end{cases} \quad (1)$$

where $RR(C)$ is the RR value of the health endpoint type at concentration level C (unit: case/person/year or day/person/year), C is the concentration level of $PM_{2.5}$, C_0 is the threshold level of $PM_{2.5}$ that is likely to cause health effects (10 μg/m^3 for $PM_{2.5}$ according to available epidemiological studies applied according to [32–34]) and ERF is the health damage function or the concentration-response function. Moreover, the suffixes p, r, y, m, e, and g represent air pollutant ($PM_{2.5}$), area, year of calculation, health endpoint groups (premature death or illness-attributed occupational injuries), the specific type of health endpoints, and the value range (including medium, lower, and upper).

The total number of health-affected cases by each health endpoint is calculated by multiplying RR with the $PM_{2.5}$ exposed population and given incidence rates in equation (2) as follows:

$$EP_i = \begin{cases} POP_{r,y,m}(RR_{p,r,y,m,e,g}(C) - 1), \text{for linear morbidity function} \\ POP_{r,y,m} I_{r,\text{"all cause"}}(RR_{p,r,y,m,e,g}(C) - 1), \text{for linear mortality function} \end{cases} \quad (2)$$

where POP is the size of the exposed population (unit: person) including children (from 0 to 14 years old), adults (from 15 to 64 years old), and the elderly (≥ 65 years old) and $I_{r,\text{"all cause"}}$ is the baseline rate of annual mean natural deaths for all health endpoints.

Based on the equation (2) above and a method similar to the Global Burden of Disease 2019 of Risk Factors Collaborators [35], the total cases of premature deaths attributable to short-term $PM_{2.5}$ exposure are determined as shown in equation (3).

$$EP_i = \sum_a \sum_d \left(POP_i \times AgeP_{i,a} \times MB_{i,a,d} \times \frac{RR_{a,d} - 1}{RR_{a,d}} \right) \quad (3)$$

where POP_i is the population size exposed to $PM_{2.5}$ pollution in study area i, $AgeP_{i,a}$ is the proportion of population in age group a in study area i, $MB_{i,a,d}$ is the baseline mortality rate of

disease type d for citizens in age group a in study area i, and $RR_{a,d}$ is the relative risk (RR) value of disease type d in the population of age group a.

Table 1. Summary of ERFs with 95% Confidence Interval (CI**).**

Type of damage	Unit	ERFs		
		Mean	Lower	Up
Early deaths due to RDs [a]		0.00090	0.00023	0.00156
Early deaths due to CVDs [b]		0.00017	-0.00002	0.00035
Early deaths due to CSDs [c]		0.00040	-0.00016	0.00096
Work loss days [d]	day/(person.µg.m.year)	0.02070	0.01760	0.02380

(*) Note: [a] referenced from [34]; [b] referenced from [35]; [c] referenced from [36]; [d] referenced from [23]

2.2.2. Model of estimations of annual per capita work loss rate

The region's total annual work loss day (WLD) is the sum of the number of days lost from work due to illness, and the cumulative number of days lost from work due to chronic mortality in adults (ages 15 to 65) is presented in equation (4). Based on the study results by [23], it is assumed that about 4% of the total number of chronic deaths in the total calculated early deaths are because of $PM_{2.5}$ exposure in the adult group. The annual per capita work loss rate (WLR) may be determined by dividing WLD by the annual number of working days and the number of workers in the population, shown in equation (5). Simultaneously, WLR is also applied to estimate the actual labor force after subtracting the number of lost works, shown in equation (6).

$$\mathrm{WLD}_{p,r,y,g} = \sum_{m} \mathrm{EP}_{p,r,y,m,"wld",g} + \sum_{e,y'\le y} \mathrm{EP}_{p,r,y',"mt",e,g} \times 4\% \times \mathrm{DPY} \quad (4)$$

$$\mathrm{WLR}_{p,r,y,g} = \frac{\mathrm{WLD}_{p,r,y,g}}{\mathrm{DPY}_{p,r,y,"adult"}} \quad (5)$$

$$\mathrm{LAB}_{p,r,y,g} = \mathrm{LAB0}_{r,y}(1 - \mathrm{WLR}_{p,r,y,g}) \quad (6)$$

where WLD is the total number of work loss days per year (unit: days/year), WLR is the per capita annual work loss rate, LAB is the number of the labor force after considering lost works, LAB0 is the number of the labor force under normal circumstances (not affected), DPY is the annual average number of working days per capita (commonly the mean level being 5 $days.week^{-1}$ × 52 $weeks.year^{-1}$ = 260 $days.year^{-1}$), the corresponding subscripts "wld" and "mt" of e and m are the cases of "days lost from work" and "chronic mortality" respectively, as well as the subscript "adult" represents for adults.

2.2.3. Rapid assessment model of PMOR-associated economic losses

Techniques of the economic loss determination attributable to health effects due to short-term $PM_{2.5}$ exposure have been developed quite early. Most of the approaches to assess and quantify into the money $PM_{2.5}$ pollution-related human health effects in many nations have been built on the basis of applying survey, assessment, and estimation methods developed by the World Health Organization and the World Bank and currently widely used in China, the United States (the U.S.), Canada, and Australia [36]. In the cost-benefit analysis method of environmental programs of the U.S. and the European Union, PMOR is considered a special value called the "Value of Statistical Life" (VSL). This is the total value that inhabitants have to pay for reducing their risks of death, which means growing life.

VSL is a function that increases with income level in the study area, which may be estimated according to equation (7) based on the method of benefit value transfer. Meanwhile, the quantification of the VSL value for Vietnam is inherited from research [24]. From there, changes in the total number of PMOR cases can be quantified into money using equations (7) and (8) [37–38] as follows:

$$VSL = VSL_0 \times \left(\frac{I}{I_0} \right)^{\alpha} \quad (7)$$

$$\Delta E = VSL \times \Delta H \quad (8)$$

where VSL is the value of statistical life of the province in the estimating year (unit: VND or USD), VSL_0 is the VSL value in the base year (unit: VND or USD), I is the annual per capita average income of the province in a given year (unit: VND or USD), I_0 is the local annually per capita income in the base year (unit: VND or USD), α is the coefficient adjusted according to income for calculating VSL value, and ΔE is the total economic losses associated with early deaths because of $PM_{2.5}$ pollution (unit: VND or USD).

3. Results and Discussion

3.1. Results of spatio-temporal PM2.5 allocation changes

Figures 3 and 4 showed that in the period from May to December 2018, the 24-hour mean $PM_{2.5}$ level ranged from 13.8-67.8 μg/m^3, in which $PM_{2.5}$ levels in some areas exceeded the threshold (with 50 μg/m^3) of the national ambient air quality standard

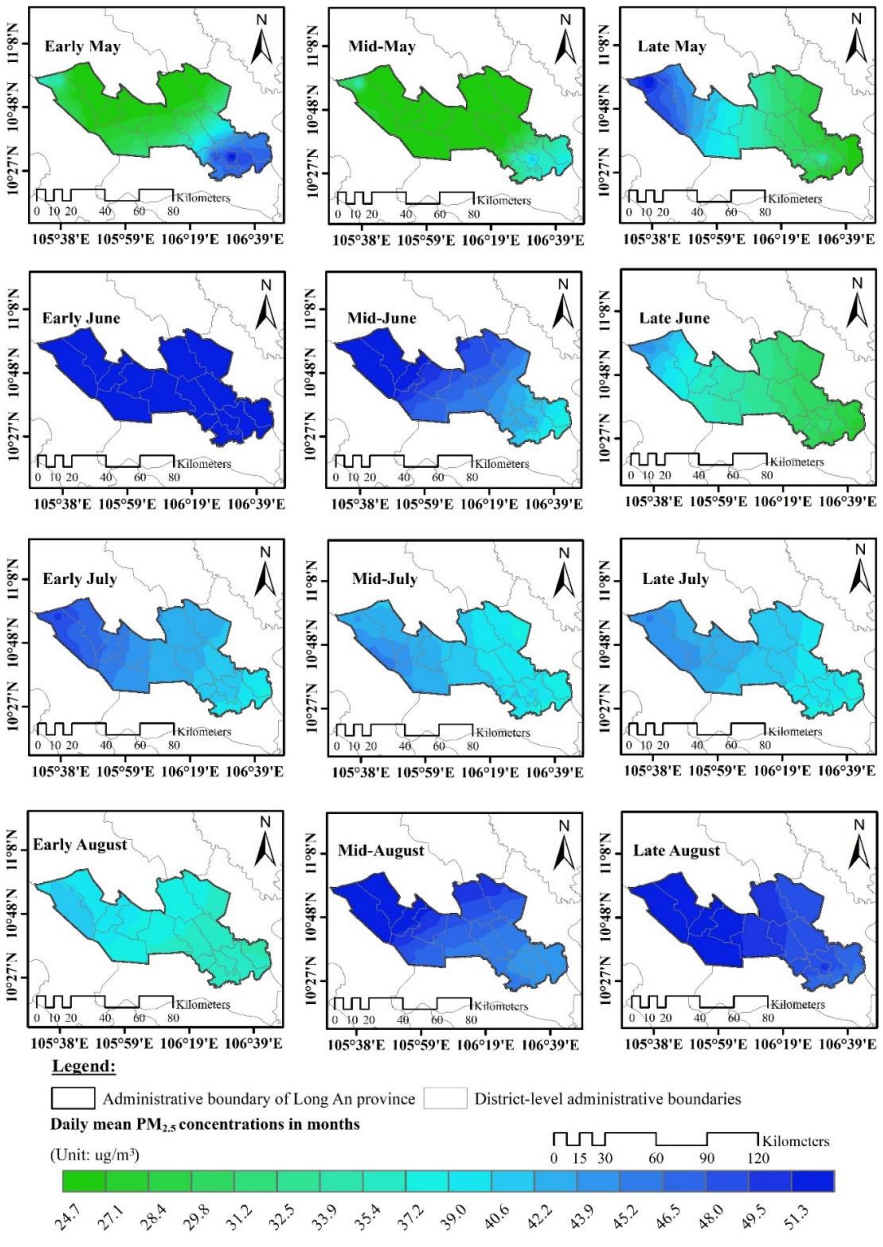


Figure 3. Spatio-temporal distributions of 24-hour average PM2.5 levels (early, mid-, and late of the months) in Long An province from May to August 2018.

(NAAQS) (or QCVN 05:2013/BTNMT) up to 1.36 times. Between May and August 2018, high $PM_{2.5}$ allocations almost occurred in the west and northwest of Long An province, consisting of Tan Hung, Vinh Hung, Kien Tuong, Tan Thanh, and Moc Hoa Districts. In contrast, from September to December 2018, high $PM_{2.5}$ distributions were observed commonly in the south and southeast of Long An province, such as Duc Hoa, Chau Thanh, and Can Duoc Districts. These are localities with large populations and strongly developed industrial activities. In general, $PM_{2.5}$ concentrations tended to be high at the month beginning, then gradually declined from the 11th to the 20th monthly, and increased gradually towards the end of May, June, and July 2018. The daily mean $PM_{2.5}$ variations in May, June, and July 2018 are from 21.6-53.5 µg/m^3, 27.8-67.8 µg/m^3, and 37.4-49.9 µg/m^3, respectively. Meanwhile, PM2.5 levels in August and October 2018 grew steadily from the beginning to the end of these months, varying from 34.7-58.7 µg/m3 and 14.1-46.5 µg/m^3, respectively. In contrast, $PM_{2.5}$ concentrations tended to reduce steadily from the beginning to the end of September, November, and December 2018, which were $PM_{2.5}$ level fluctuations of 13.3-33.3 µg/m^3, 18.3-65.5 µg/m^3, and 37.6-64.8 µg/m^3, respectively.

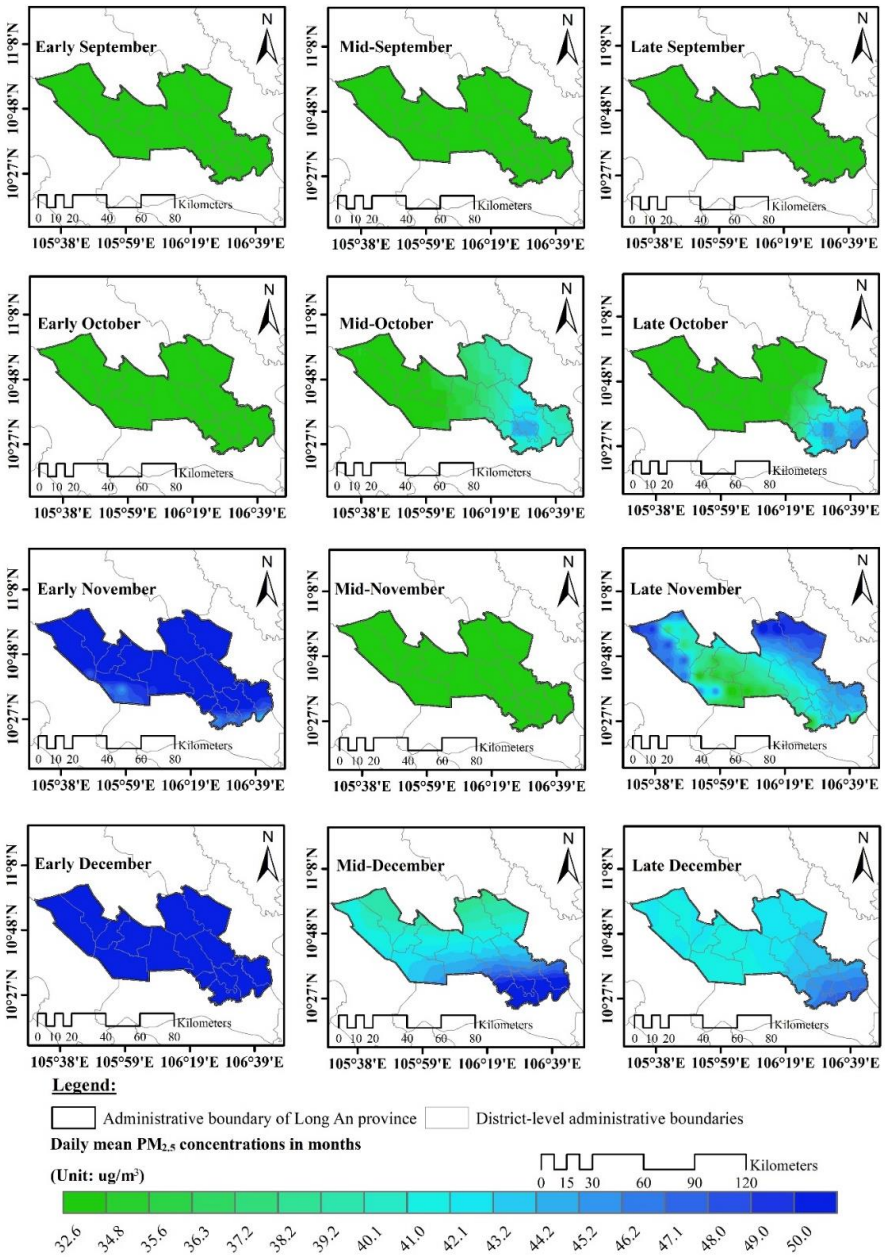


Figure 4. Spatio-temporal allocations of daily mean $PM_{2.5}$ levels (early, mid-, and late of the months) in Long An province from September to December 2018.

3.2. Results of the effect assessment on public health attributable to $PM_{2.5}$ exposure

Table 2 and Figure 5 reported that the total health-affected cases in Long An province from May to December 2018 were about 265 (95% CI: -12; 422) PMOR cases. In particular, there were about 148 (95% CI: -7; 236) premature deaths between May and August 2018, including around 86 (95% CI: 13; 95) cases related to RDs (accounting for 58.1%), 19 (95% CI: -2; 38) cases due to CVDs (contributing to 12.8%), and 43 (95% CI: -18; 103) cases attributable to CSDs (contributing to 29.1%). Meanwhile, the PMOR cases from September to December 2018 tended to decline roughly 1.3 times, having about 117 (95% CI: -5; 186) PMOR cases. Specifically, there were respectively around 68 (95% CI: 10; 74) cases, 15 (95% CI: -2; 30) cases, and 34 (95% CI: -14; 81) cases, which related to RDs, CVDs, and CSDs, respectively. In this period, the proportion of PMOR cases due to RDs was the highest (58.2%), whilst the percentage of premature deaths associated with CVDs was the lowest (12.7%). Generally, the types of RDs-associated health endpoints are the most frequent health issues caused by short-term $PM_{2.5}$ exposure in the study area, followed by CSDs and CVDs. In particular, the total early deaths because of RDs, CVDs, and CSDs in Long An province during the evaluated period were 155 (95% CI: 23; 170) cases (accounting for 58.5%), 33 (95% CI: -4; 68) cases (accounting for 12.5%), and 77 (95% CI: -31; 184) cases (contributing to 29.1%), respectively.

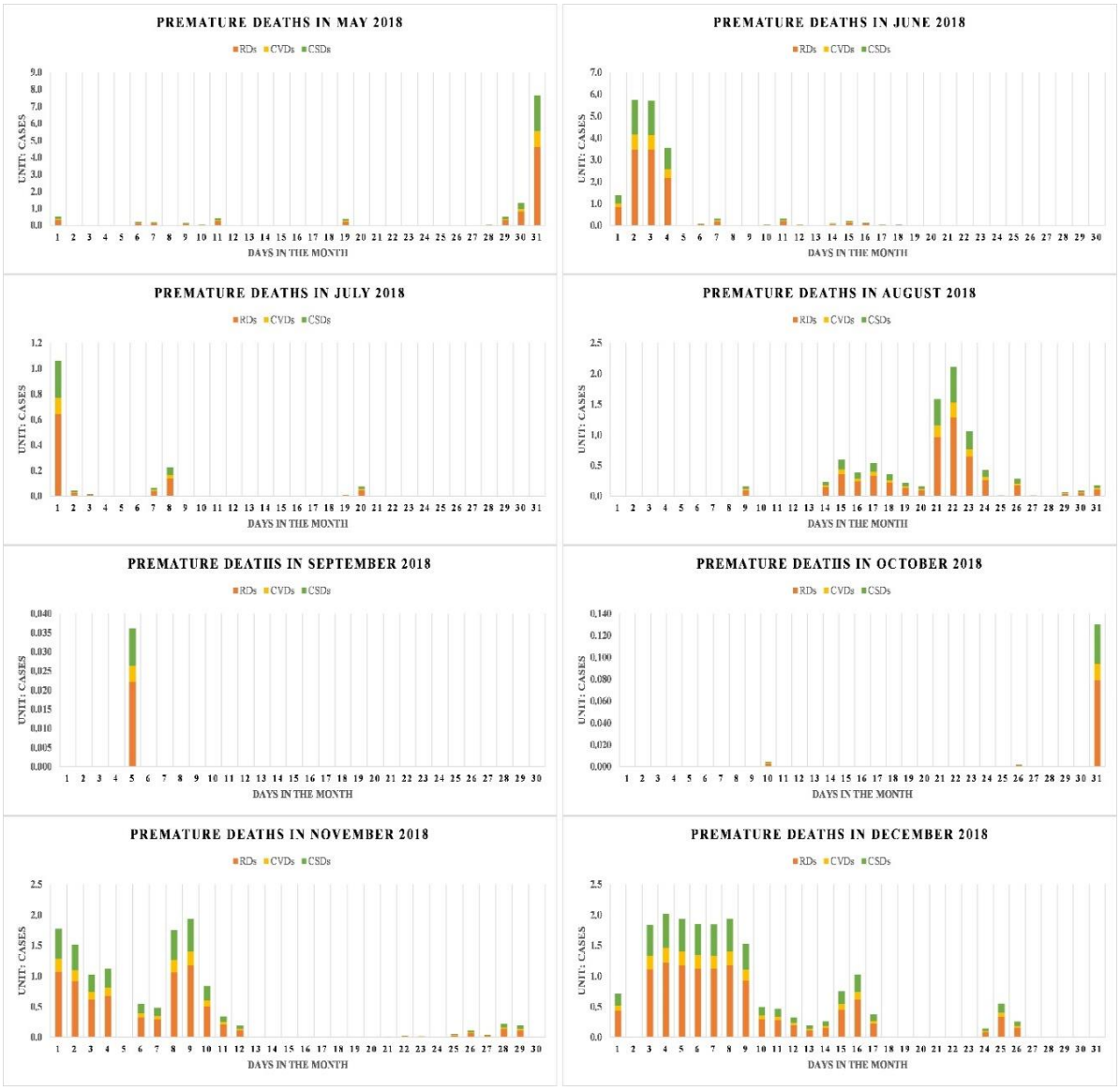


Figure 5. Changes in daily PMOR cases because of RDs, CVDs, and CSDs attributable to acute $PM_{2.5}$ exposure in Long An province from May to December 2018.

By each month, the total PMOR cases declined from May to October 2018, then rose sharply until December 2018. September and October 2018 had the lowest number of early deaths with 68 (95% CI: -3; 109) cases, whilst June and December 2018 had the highest number of PMOR cases with 69 (95% CI: -3; 110) cases.

Moreover, it could be found that the higher cases of PMOR occurred commonly from the early to the middle of June, July, and December 2018. In contrast, they almost happened from the middle to the end of May and August 2018, whilst the total PMOR cases were observed majorly at the beginning and the end of November 2018. Specifically, the highest number of PMOR cases occurring by each assessed month was approximately 8 (95% CI: 0; 11) cases on 31 May 2018, 6 (95% CI: 0; 8) cases on 2 June 2018, 1 (95% CI: 0; 2) cases on 1 July 2018, 2 (95% CI: 0; 3) cases on 22 August 2023, 2 (95% CI: 0; 3) cases on 9 November 2023, and 2 (95% CI: 0; 3) cases on 4 December 2018.

Table 2. Results of the estimations of PMOR cases in each month and each kind of health endpoint because of acute $PM_{2.5}$ exposure (with 95% CI).

Month	$EP_{PM2.5,\ RDs}$[1]	$EP_{PM2.5,\ CVDs}$[2]	$EP_{PM2.5,\ CSDs}$[3]
May 2018	24.8 (95% CI: 3.8; 27.5)	5.4 (95% CI: -0.6; 11.0)	12.6 (95% CI: -5.1; 29.7)
Jun. 2018	39.6 (95% CI: 6.0; 44.0)	8.6 (95% CI: -1.0; 17.5)	20.0 (95% CI: -8.1; 47.4)
Jul. 2018	3.3 (95% CI: 0.5; 3.6)	0.7 (95% CI: -0.1; 1.4)	1.6 (95% CI: -0.7; 3.9)
Aug. 2018	18.5 (95% CI: 2.7; 20.4)	3.9 (95% CI: -0.5; 8.0)	9.2 (95% CI: -3.7; 21.9)
Sept. 2018	0.08 (95% CI: 0.01; 0.09)	0.0 (95% CI: 0.0; 0.0)	0.0 (95% CI: 0.0; 0.1)
Oct. 2018	0.30 (95% CI: 0.04; 0.33)	0.1 (95% CI: 0.0; 0.1)	0.1 (95% CI: -0.1; 0.4)
Nov. 2018	27.5 (95% CI: 4.1; 30.0)	5.8 (95% CI: -0.7; 12.0)	13.7 (95% CI: -5.5; 32.7)
Dec. 2018	40.5 (95% CI: 6.0; 44.0)	8.6 (95% CI: -1.0; 17.7)	20.2 (95% CI: -8.1; 48.1)
Total	**154.6 (95% CI: 23.2; 169.8)**	**33.1 (95% CI: -3.9; 67.9)**	**77.5 (95% CI: -31.3; 184.2)**

(*) Notes: [1] PMOR cases due to RDs; [2] PMOR cases due to CVDs; [3] PMOR cases due to CSDs

3.3. Results of the impact assessment associated with labor time decline

Based on datasets from the Long An Statistical Yearbook in 2018 [26], the labor force (> 15 years old) in this study area obtained 1,006.7 thousand people, including male workers with 544.6 thousand people (contributing to 54.1%) and female workers with 462.1 thousand people (accounting for 45.9%). Furthermore, the labor force in rural areas was significantly higher than that in urban areas; specifically, there were 156.3 thousand laborers in urban areas (accounting for 15.5%) and 850.4 thousand laborers in rural areas (accounting for 84.5%). Simultaneously, by age group division, there were 126.3 thousand laborers aged 15 to 24 years old, 604.0 thousand laborers aged 25 to 49 years old, and 276.4 thousand people over 50 years old in Long An province.

High outdoor $PM_{2.5}$ concentration exposure often results in declining working days because of additional morbidity and PMOR. Citizens who live in environmental conditions with better ambient air quality are likely to work more due to their living quality (or even their children) being less able to get sick. In particular, the cases occurring in the adult group (< 65 years old) have significantly reduced their working time. The outcomes presented that the average number of days lost from work during the period of May to December 2018 was approximately 16.5 days/person. It contributed around 6.35%, compared to the annual average number of working days per capita of 260 days/year. At the same time, the highest number of days lost from work occurred respectively in June and December 2018, with about 5.2 days/person (accounting for 2.0%) and 8.3 days/person (accounting for 3.2%).

3.4. Results of the estimations of PMOR-related economic losses

The combinations of PMOR-related health impact results due to acute $PM_{2.5}$ exposure based on the specific types of health endpoints and the theories of Modue 3 in the EnHEBIS software had been applied to rapidly quantify the economic loss values ($HE_{PM2.5,\ mortality}$) in Long An province between May and December 2018. The determination of economic losses was based on Vietnam's VSL in 2018 of 4,789 million VND (equivalent to about 642 million USD, 2018 US$). The value of VSL measures inhabitants' willingness to pay for better and safer living and working conditions, as well as being used to assess the benefits of policies and implementing solutions from agencies and units in both private and public sectors of the national authorities [36, 39]. Thus, the total economic loss values caused by $PM_{2.5}$ pollution was estimated at 1,270 (95% CI: -57; 2,021) billion VND, which was equivalent to approximately 170 (95% CI: -7; 271) million USD during the study period. Comparisons to the gross regional product (GRDP) in Long An province in 2018 of 110,077 trillion VND referenced from [40], the total PMOR-related economic loss values contributed to around 1.15% of the province's GRDP value in 2018.

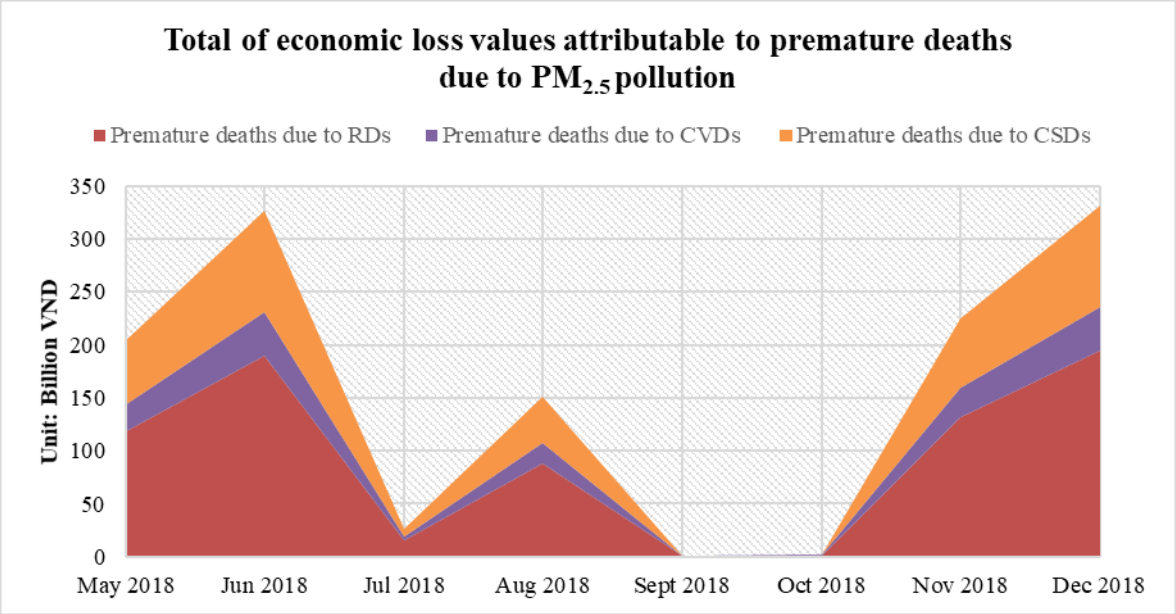


Figure 6. Results of economic loss values in each month due to $PM_{2.5}$-related short-term health impacts between May and December 2018 in this study area.

As shown in Figure 6, the total economic loss values (from May to August 2018) in Long An province attributable to PMOR cases were caused by daily $PM_{2.5}$ levels exceeded the threshold of NAAQS (QCVN 05:2013/BTNMT) with the damage of around 709 (95% CI: -32; 1,132) billion VND, equivalent to 95 (95% CI: -4; 152) million USD. In detail, the total RDs-related PMOR cases resulted in the total economic losses during this period of 413 (95% CI: 62; 457) billion VND, equivalent to approximately 55 (95% CI: 8; 61) million USD (accounting for 13.3%). In comparison to RDs, the total PMOR cases due to CVDs and CSDs caused the corresponding loss values around 89 (95% CI: -10; 182) billion VND (average of about 12 million USD) accounting for 21.5% and 208 (95% CI: -84; 493) billion VND (average of roughly 28 million USD) contributing to 6.8% of the total loss values, respectively.

From September to December 2018, the total economic loss values caused by acute $PM_{2.5}$ exposure-related early deaths due to RDs, CVDs, and CSDs was roughly 561 (95% CI: -25; 889) billion VND, equivalent to 75 (95% CI: -3; 119) million USD in Long An province (Figure 6). The economic losses were about 1.3 times lower than in the period of May to August 2018. Specifically, by each kind of health endpoint, the total loss values were 328 (95% CI: 49; 356) billion VND, equivalent to approximately 44 (95% CI: 7; 48) million USD

(accounting for 58.5%) caused by RDs. Meanwhile, there were approximately 70 (95% CI: -8; 143) billion VND (average of about 9 million USD) and 163 (95% CI: -66; 389) billion VND (average of around 22 million USD), respectively, in a total of economic losses caused by CVDs (contributing to 12.5%) and CSDs (contributing to 29.1%).

Besides, by each month, the total loss values were in the order as follows: September < October < July < August < May < November < June < December. In particular, the loss value was the highest in December 2018 with 349 (95% CI: -15; 509) billion VND (average about 47 million USD), whilst the lowest loss value occurred in September 2018 with 0.7 (95% CI: 0.0; 1.0) billion VND (average about 0.1 million USD). Meanwhile, the total economic losses in May, November, and June 2018 were respectively 204 (95% CI: -9; 326) billion VND (average around 27 million USD), 237 (95% CI: -10; 346) billion VND (average about 32 million USD), and 327 (95% CI: -15; 522) billion VND (average about 44 million USD). Generally, the loss values in October, July, and August 2018 were much lower than in May, November, and June 2018. There were 3 (95% CI: -0.1; 4) billion VND (average roughly 0.3 million USD), 27 (95% CI: -1; 43) billion VND (average around 4.0 million USD), and 151 (95% CI: -7; 241) billion VND (average about 20 million USD), respectively in the total values of economic losses in October, July, and August 2018.

3.5. Discussion

On the basis of the approach taken in this study and the estimation results of acute health impacts, as well as the economic loss values achieved, it could be found that the $PM_{2.5}$ concentration level in each district plays an important role in the estimations and also has a linear relationship to these damage values. However, the exposed population size and the population density of each locality in Long An province are the most crucial factors that have a decisive influence on the values of the estimated damage. This is completely similar and supports observations from previously available studies, such as [24, 41–42]. Compared to a similar study by [43] that evaluated PMOR due to acute $PM_{2.5}$ exposure in HCMC, the total number of PMOR cases was about 1.12 times lower. It could be seen that the study by [43] estimated for all months in 2018; moreover, the population size of HCMC in 2018 (about 8.832 million people) was also 5.26 times higher than that of the study area. Though the total economic losses caused by short-term exposure to $PM_{2.5}$ in HCMC (around 410 million USD) are significantly higher than this study results (roughly 170 million USD), they only account for 0.25% of HCMC's total GRDP compared to 1.15% of Long An's total GRDP in 2018. This is obviously because the GRDP value of HCMC is remarkably higher than that of Long An province. Nevertheless, the study by [43] noted the health effects of morbidity types, such as hospitalizations and emergency visits, and also estimated several specific types of RDs and CVDs, including acute lower respiratory infections (LRIs) due to pneumonia and acute bronchiolitis in children, chronic obstructive pulmonary disease (COPD), community-acquired pneumonia (CAP), heart failure (HF), myocardial infarction (MI)), and stroke. Simultaneously, the total corresponding economic losses are not only estimated for PMOR but are determined by the costs of illness treatment, loss of income and declined productivity labor in adults. Similarly, the total economic loss value of this study is also 2.32 times lower than the results of [44], which evaluated PMOR related to CVDs occurring in HCMC in 2018. Compared with a similar study by [45] in Guangzhou, China in 2010, the total PMOR cases in our study was 1.46 times higher. Meanwhile, compared to the results by [24], the total loss values in Long An province from May to December 2018 (around 170 million USD) were much higher than that in the period of January to April 2018 (roughly 8.3 million USD). However, the estimated VSL value for Long An province in 2018 was used to quantify economic losses because of PMOR in our study instead of using Vietnam's VSL in 2018 as in the study by [24] since this is likely to reduce errors in estimating total economic losses.

4. Conclusion

The study outcomes provided evidence that $PM_{2.5}$ concentration levels in this study area during the period of May to December 2018 exceeded the threshold of NAAQS (QCVN 05:2013/ BTNMT) and caused negative health impact issues, which contributed to rose PMOR and created a considerable economic burden in Long An province. Specific highlights are as follows:

Acute health effects associated with short-term $PM_{2.5}$ exposure commonly caused around 155 (95% CI: 23; 170) PMOR cases due to RDs, 33 (95% CI: -4; 68) PMOR cases due to CVDs, and 77 (95% CI: -31; 184) PMOR cases related to CSDs in Long An province in 2018.

Exposure to high $PM_{2.5}$ levels in ambient air resulted in the decline of working hours in the adult group by an average of about 16.5 days/person between May and December 2018.

The total economic losses caused by $PM_{2.5}$ pollution in this period of 2018 were quantified at 1,270 (95% CI: -57; 2,021) billion VND, equivalent to about 170 (95% CI: -7; 271) million USD (2018 US$). These PMOR-related economic losses attributable to short-term outdoor $PM_{2.5}$ exposure accounted for about 1.15% of the total GRDP of the province in 2018.

Although the highlights have remarkably contributed in terms of scientific basis to clarify the acute effects of $PM_{2.5}$ causing PMOR cases, as well as corresponding economic losses in Long An province in 2018, this study has not yet completely looked at the cases of hospitalizations, emergency visits, or even long-term $PM_{2.5}$ exposure issues (chronic). Since chronic $PM_{2.5}$ exposure is likely to result in much greater potential health risks than short-term $PM_{2.5}$ exposure. At the same time, more specific health endpoint types of all-cause RDs and CVDs have not been noted in this study; thus, it is crucial to extend and update the latest ERFs for health effect assessment in further studies. Furthermore, future studies need to continue evaluating the distributions of $PM_{2.5}$ concentrations throughout Long An province for the following years in order to make specific comments on spatio-temporal fluctuation trends. They will be a solid basis for policymakers to develop strategies and measures to control and mitigate $PM_{2.5}$ precursor emissions in Long An province in Vietnam. In particular, these solutions should focus on the major sources of the province, typically industry, construction, residential, transportation, and agriculture activities.

Author contribution statement: Developing research ideas, drawing up a draft writing plan, editing the manuscript: L.T.B.; Process data, run coupled WRF/CMAQ models, run health-economic impact simulations: P.H.N., N.H.T.N.; GIS processing, manuscript writing: P.H.N.

Acknowledgments: This research was funded by Vietnam National University Ho Chi Minh city (VNU-HCM), grant No: B2023-20-23. The authors would like to thank the Ho Chi Minh City University of Technology for the support of time and facilities from the Ho Chi Minh City University of Technology (HCMUT), VNU-HCM for this study.

Competing interest statement: The authors declare no conflict of interest.

References

1. Manisalidis, I.; Stavropoulou, E.; Stavropoulos, A.; Bezirtzoglou, E. Environmental and Health Impacts of Air Pollution: A Review. *Front. Public Heal.* **2020**, *8*, 1–13.
2. Kjellstrom, T.; Holmer, I.; Lemke, B. Workplace heat stress, health and productivity - an increasing challenge for low and middle-income countries during climate change. *Glob. Health Action* **2009**, *2(1)*, 2047.
3. Sachs, J.D.; Schmidt-Traub, G.; Mazzucato, M.; Messner, D.; Nakicenovic, N.; Rockström, J. Six Transformations to achieve the Sustainable Development Goals. *Nat. Sustain.* **2019**, *2(9)*, 805–814.
4. Nam, D.T. et al. Desgning and manufacturing ambient air quality gravity sample collection equipment. *Environ. Mage.* **2023**, *I*, 48–53.
5. Delfino Ralph, J.; Constantinos, S.; Shaista, M. Potential Role of Ultrafine Particles in

Associations between Airborne Particle Mass and Cardiovascular Health. *Environ. Health Perspect.* **2005**, *113(8)*, 934–946.
6. WHO. Burden of disease attributable to outdoor air pollution. 1211 Geneva 27, Switzerland, 2011.
7. Hoek, G. et al. Long-term air pollution exposure and cardio-respiratory mortality: A review. *Environ. Heal. A Glob. Access Sci. Source* **2013**, *12(1)*, 43.
8. Shang, Y. et al. Systematic review of Chinese studies of short-term exposure to air pollution and daily mortality. *Environ. Int.* **2013**, *54*, 100–111.
9. Perone, G. Assessing the impact of long-term exposure to nine outdoor air pollutants on COVID-19 spatial spread and related mortality in 107 Italian provinces. *Sci. Rep.* **2022**, *12(1)*, 1–24.
10. Hayes, R.B. et al. $PM_{2.5}$ air pollution and cause-specific cardiovascular disease mortality. *Int. J. Epidemiol.* **2019**, *49(1)*, 25–35.
11. Li, X.; Ma, Y.; Wang, Y.; Liu, N.; Hong, Y. Temporal and spatial analyses of particulate matter (PM_{10} and $PM_{2.5}$) and its relationship with meteorological parameters over an urban city in northeast China. *Atmos. Res.* **2017**, *198*, 185–193.
12. Munir, S. et al. Modeling particulate matter concentrations in Makkah, applying a statistical modeling approach. *Aerosol Air Qual. Res.* **2013**, *13(3)*, 901–910.
13. Hallquist, M. et al. The formation, properties and impact of secondary organic aerosol: Current and emerging issues. *Atmos. Chem. Phys.* **2009**, *9(14)*, 5155–5236.
14. Hien, T.T.; Chi, N.D.T.; Nguyen, N.T.; Vinh, L.X.; Takenaka, N.; Huy, D.H. Current Status of Fine Particulate Matter ($PM_{2.5}$) in Vietnam's Most Populous City, Ho Chi Minh City. *Aerosol Air Qual. Res.* **2019**, *19(10)*, 2239–2251.
15. Scherer, D.; Fehrenbach, U.; Lakes, T.; Lauf, S.; Meier, F.; Schuster, C. Quantification of heat-stress related mortality hazard, vulnerability and risk in Berlin, Germany. *J. Geogr. Soc. Berlin* **2014**, *144(3-4)*, 238–259.
16. Samet, J.; Wassel, R.; Holmes, K.J.; Abt, E.; Bakshi, K. Research Priorities for Airborne Particulate Matter in the United States. *Environ. Sci. Technol.* **2005**, *39(14)*, 299A-304A.
17. Pope, C.A.; Dockery, D.W. Health effects of fine particulate air pollution: Lines that connect. *J. Air Waste Manag. Assoc.* **2006**, *56(6)*, 709–742.
18. OECD. OECD Environmental Outlook to 2050, 2012.
19. Lim, S.S. et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **2012**, *380(9859)*, 2224–2260.
20. Yin, P. et al. The effect of air pollution on deaths, disease burden, and life expectancy across China and its provinces, 1990–2017: an analysis for the Global Burden of Disease Study 2017. *Lancet Planet. Heal.* **2020**, *4(9)*, e386–e398.
21. Maji, K.J.; Ye, W.F.; Arora, M.; Shiva Nagendra, S.M. $PM_{2.5}$-related health and economic loss assessment for 338 Chinese cities. *Environ. Int.* **2018**, *121*, 392–403.
22. Huang, R.J. et al. High secondary aerosol contribution to particulate pollution during haze events in China. *Nature* **2014**, *514(7521)*, 218–222.
23. Xie, Y.; Dai, H.; Dong, H.; Hanaoka, T.; Masui, T. Economic Impacts from $PM_{2.5}$ Pollution-Related Health Effects in China: A Provincial-Level Analysis. *Environ. Sci. Technol.* **2016**, *50(9)*, 4836–4843.
24. Bui, L.T.; Lai, H.T.N.; Nguyen, P.H. Benefits of Short-term Premature Mortality Reduction Attributed to $PM_{2.5}$ Pollution: A Case Study in Long an Province, Vietnam. *Arch. Environ. Contam. Toxicol.* **2023**, *85*, 245–262.
25. Bey, I. et al. Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *J. Geophys. Res. Atmos.* **2001**, *106(D19)*, 23073–23095.
26. Long An Provincial Statistics Office. Statistical Yearbook of Long An Province 2020, Tan An City, 2021.

27. GSO. Statistical Yearbook of Vietnam 2018, Ha Noi Capital: The Statistical Publishing House, 2019.
28. Chen, R. et al. Association of particulate air pollution with daily mortality: The China air pollution and health effects study. *Am. J. Epidemiol.* **2012**, *175(11)*, 1173–1181.
29. Pope III, C.A. et al. Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution. *JAMA* **2002**, *287(9)*, 1132–1141.
30. Apte, J.S.; Marshall, J.D.; Cohen, A.J.; Brauer, M. Addressing Global Mortality from Ambient $PM_{2.5}$. *Environ. Sci. Technol.* **2015**, *49(13)*, 8057–8066.
31. Burnett, R.T. et al. An Integrated Risk Function for Estimating the Global Burden of Disease Attributable to Ambient Fine Particulate Matter Exposure. *Environ. Health Perspect.* **2014**, *122(4)*, 397–403.
32. Cai, J. et al. Association between $PM_{2.5}$ exposure and all-cause, non-accidental, accidental, different respiratory diseases, sex and age mortality in Shenzhen, China. *Int. J. Environ. Res. Public Health* **2019**, *16(3)*, 401.
33. Orellano, P.; Reynoso, J.; Quaranta, N.; Bardach, A.; Ciapponi, A. Short-term exposure to particulate matter (PM_{10} and $PM_{2.5}$), nitrogen dioxide (NO_2), and ozone (O_3) and all-cause and cause-specific mortality: Systematic review and meta-analysis. *Environ. Int.* **2019**, *142*, 105876.
34. Qu, Y. et al. Short-term effects of fine particulate matter on non-accidental and circulatory diseases mortality: A time series study among the elder in Changchun. *PLoS One* **2018**, *13(12)*, 1–12.
35. Murray, C.J.L. et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **2020**, *396(10258)*, 1223–1249.
36. Chinh, N.T. Evaluate economic losses due to pollution and environmental degradation. Hanoi capital: National Political Publishing House, 2013.
37. Zhou, Z. et al. The health benefits and economic effects of cooperative $PM_{2.5}$ control: A cost-effectiveness game model. *J. Clean. Prod.* **2019**, *228*, 1572–1585.
38. Zhao, N. et al. Field-based measurements of natural gas burning in domestic wall-mounted gas stove and estimates of climate, health and economic benefits in rural Baoding and Langfang regions of Northern China. *Atmos. Environ.* **2020**, *229*, 117454.
39. Thinh, H.B. Social forecasts about the Covid-19 pandemic. *Vietnam Social Sciences Journal* **2022**, *3*, 3–12.
40. HCMC Statistical Office. Part II: Actual Situation of Economic Growth of Key Economic Region of South Vietnam in the Period of 2010-2018, in *Ho Chi Minh City Economy and the Southern Key Economic Region*, Ho Chi Minh City: Ho Chi Minh City Statistical Office, **2019**, 19–30.
41. Bui, L.T.; Nguyen, P.H.; My Nguyen, D.C. Linking air quality, health, and economic effect models for use in air pollution epidemiology studies with uncertain factors. *Atmos. Pollut. Res.* **2021**, *12(7)*, 101118.
42. Nhung, N.T.T. et al. Mortality benefits of reduction fine particulate matter in Vietnam, 2019. *Front. Public Heal.* **2022**, *10*.
43. Bui, L.T.; Nguyen, P.H. Evaluation of the annual economic costs associated with $PM_{2.5}$-based health damage – a case study in Ho Chi Minh City, Vietnam. *Air Qual. Atmos. Heal.* **2022**, *16*, 415–435.
44. Dang, T.N. et al. Mortality and economic burden of $PM_{2.5}$ on cardiovascular disease in Ho Chi Minh City in 2018. *Vietnam J. Prev. Med.* **2021**, *31(6)*, 9–18.
45. Ding, D. et al. Evaluation of health benefit using BenMAP-CE with an integrated scheme of model and monitor data during Guangzhou Asian Games. *J. Environ. Sci. (China)* **2016**, *42*, 9–18.

Table of content