MODELING AND PROPAGATION OF NOISE OF
A WIND TURBINE PROJECT IN THAILAND

NATTAPAT CHAROENTANGPRASERT

A Thesis Submitted in Partial Fulfillment of the Requirements for
the Degree of Master of Engineering in Industrial Systems
and Environmental Engineering
Suranaree University of Technology
Academic Year 2022

แบบจำลองและการกระจายตัวของเสียงที่เกิดจาก
โครงการกังหันลมในประเทศไทย

นายณัฐภัทร เจริญตั้งประเสริฐ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมระบบอุตสาหกรรมและสิ่งแวดล้อม
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2565

MODELING AND PROPAGATION OF NOISE OF
A WIND TURBINE PROJECT IN THAILAND

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Master's Degree

Thesis Examining Committee

....

(Assoc. Prof. Boonchai Wichitsathian)

Chairperson

....

(Assoc. Prof. Netnapid Tantemsapya)

Member (Thesis Advisor)

....

(Dr. Chatpet Yossapol)

Member

....

(Prof. Dr. Watinee Thavorntam)

Member

....

(Assoc. Prof. Dr. Chatchai Jothiyangkoon)

Vice Rector for Academic Affairs and

Quality Assurance

....

(Assoc. Prof. Pornsiri Jongkol)

Dean of Institute of Engineering

ณัฐภัทร เจริญตั้งประเสริฐ: แบบจำลองและการกระจายตัวของเสียงที่เกิดจากโครงการ
กังหันลมในประเทศไทย
อาจารย์ที่ปรึกษา: รองศาสตราจารย์ ดร.เนตรนภิส ตันเต็มทรัพย์, 111 หน้า

คำสำคัญ: เสียงรบกวนจากกังหันลม/ ระบบสารสนเทศภูมิศาสตร์/ แผนที่เสียงรบกวน/ แบบจำลอง
การทำนายเสียงรบกวน/ SPSS MODELER

การศึกษานี้มีวัตุประสงค์เพื่อวิเคราะห์การกระจายตัวของเสียงจากกังหันลมในพื้นที่ชุมชน
ชนบทในภาคตะวันออกเฉียงเหนือของประเทศไทย โดยการสร้างแผนที่เสียงและตัวแบบทำนายเสียง
แผนที่เสียงถูกสร้างขึ้นโดยใช้ระบบสารสนเทศภูมิศาสตร์จากซอฟต์แวร์ ArcGIS 10.5 และใช้วิธีการ
ประมาณค่าช่วงแบบ kriging โดยวิเคราะห์ภูมิสถิติ (geostatistical analyst) ระดับเสียงที่ใช้เพื่อ
สร้างแผนที่เสียงถูกเก็บข้อมูลประจำปีตั้งแต่ พ.ศ. 2561 ถึง พ.ศ.2564 จำนวน 40 ตำแหน่ง โดยรอบ
ขอบเขตฟาร์มกังหันลมเป็นระยะ 800 เมตร นอกจากนี้ ระดับเสียงได้รับการเก็บรวบรวมจากสถานี
ตรวจวัดระดับเสียงอัตโนมัติเพื่อวิเคราะห์ความแตกต่างระหว่างระดับเสียงในช่วงเวลากลางวันและ
กลางคืนโดยใช้วิธีการวิเคราะห์เชิงสถิติแบบทดสอบทีและเปรียบเทียบระดับเสียงกับมาตรฐานเสียง
รบกวน แบบจำลองทำนายเสียงรบกวนพัฒนาโดยใช้ซอฟต์แวร์ IBM SPSS Modeler ระดับเสียงและ
ข้อมูลสภาพอากาศได้ถูกวัดทุก 5 นาทีเป็นเวลา 3 วัน จากขอบเขตฟาร์มกังหันลมเป็นระยะ 400
เมตรในทิศเหนือตะวันออก มีการพัฒนาจำลองจากอัลกอริทึม 5 รูปแบบ คือ แบบจำลองCHAID
แบบจำลอง CART แบบเชิงเส้น (Linear) ระบบประสาทเทียม (Neural network) และแบบจำลอง
แบบการรวมโมเดล (Ensemble) ตัวแปรที่ใช้เป็นชุดข้อมูลในการพัฒนาแบบจำลอง คือ ระยะทาง
เวลา ความเร็วลม ทิศทางลม อุณหภูมิ ความชื้น และความดัน อัลกอริทึมถูกเปรียบเทียบเพื่อหา
อัลกอริทึมที่เหมาะสมที่สุดโดยการประเมินผล

โดยผลลัพธ์ที่จากแผนที่เสียงแสดงให้เห็นว่าระดับเสียงในพื้นที่ที่อยู่อาศัยอยู่ระหว่าง 45 ถึง
60 เดซิเบล(เอ) ในพื้นที่ฟาร์มกังหันลมอยู่ระหว่าง 45 ถึง 70 เดซิเบล(เอ) และในพื้นที่เกษตรกรรมอยู่
ระหว่าง 40 ถึง 55 เดซิเบล(เอ) ผลการตรวจวัดที่ได้จากสถานีตรวจวัดระดับเสียงอัตโนมัติแสดงให้
เห็นว่าระดับเสียงในหมู่บ้านทางตะวันออกเฉียงใต้อยู่ระหว่าง 41.5 ถึง 87.6 เดซิเบล(เอ) และใน
หมู่บ้านทางตะวันออกเฉียงเหนืออยู่ระหว่าง 29.9 ถึง 81.8 เดซิเบล(เอ) มีความแตกต่างระหว่างระดับ
เสียงในช่วงเวลากลางวันและกลางคืนอย่างมีนัยสำคัญ ผลการเปรียบเทียบระดับเสียงที่วัดได้กับ
มาตรฐานเสียงรบกวนพบว่าเขตชุมชนที่ใกล้ถนนได้รับการสัมผัสระดับเสียงในช่วงกลางคืนเกินกว่าค่า
แนะนำจากองค์การอนามัยโลก (World Health Organization) ผลลัพธ์ที่ได้จากแบบทำนายเสียง
รบกวนพบว่าระดับเสียงจะดังขึ้นเมื่อเข้าใกล้กังหันลมและพบว่าระดับเสียงที่เกิดขึ้นได้รับอิทธิพลจาก
กิจกรรมของมนุษย์โดยเฉพาะในช่วงเช้าและเย็นซึ่งคาดว่าเกิดจากการจราจร ผลจากการเปรียบเทียบ
อัลกอริทึมพบว่า แบบจำลอง Ensemble ถูกพิสูจน์ว่าเป็นเทคนิคที่เหมาะสมที่สุดจาก
การเปรียบเทียบด้วย R^2 MAE และ RMSE จากการตรวจสอบแบบจำลอง Ensemble

แบบ Cross validation พบว่า MAE และ RMSE มีค่าแตกต่างกันที่ร้อยละ 5.89 และ 10.08 แสดงว่าแบบจำลอง Ensemble ไม่เกิดการ Overfitting

        ผลการศึกษาแสดงให้เห็นว่าการจราจรเป็นแหล่งกำเนิดเสียงหลักซึ่งอาจเป็นปัจจัยเสี่ยงต่อสุขภาพของชาวบ้านในชุมชนใกล้เคียงซึ่งมีความจำเป็นในการปรับปรุงการควบคุมเสียงเพื่อลดและควบคุมระดับเสียงในอยู่ในมาตรฐานโดยผู้ที่มีส่วนเกี่ยวข้องสามารถใช้ข้อมูลนี้เพื่อระบุพื้นที่ที่จำเป็นต้องควบคุมและลดระดับมลพิษเสียง

สาขาวิชาวิศวกรรมสิ่งแวดล้อม

ปีการศึกษา 2565

ลายมือชื่อนักศึกษา _____

ลายมือชื่ออาจารย์ที่ปรึกษา _____

NATTAPAT CHAROENTANGPRASERT: MODELING AND PROPAGATION OF NOISE OF A WIND TURBINE PROJECT IN THAILAND
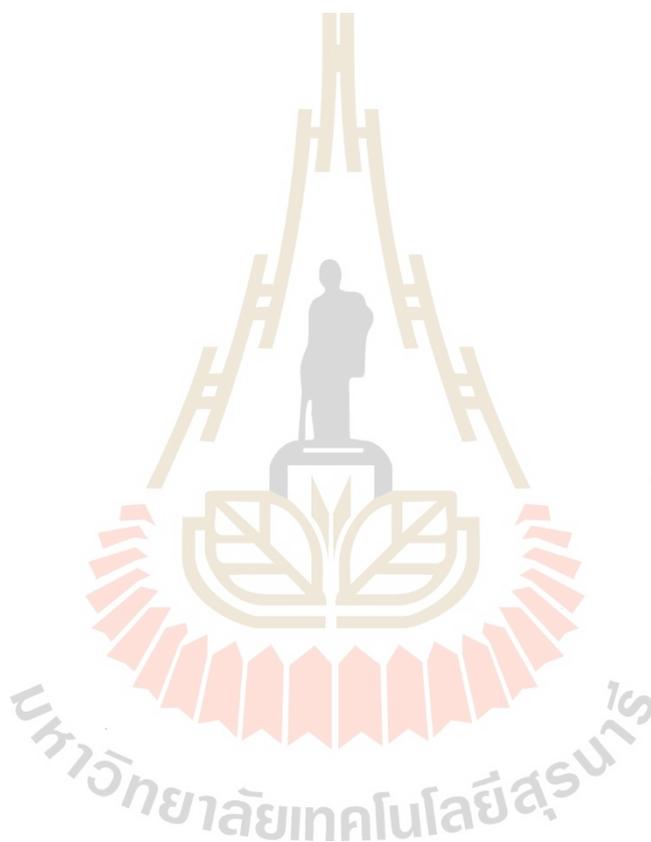THESIS ADVISOR: ASSOC. PROF. NETNAPID TANTEMSAPYA, 111 PP.

This study analyzes wind turbine noise propagation in a rural area of northeast Thailand by developing noise maps and a noise prediction model. Noise maps were generated using geographic information systems (GIS) from ArcGIS 10.5 software, performed by the kriging interpolation method on geostatistical analyst. Noise levels were measured annually from 2018 to 2021 at 40 locations, extending 800 meters from the wind farm boundary to create the noise maps. Additionally, noise levels were collected from automated sound monitoring stations to differentiate between daytime and nighttime periods using independent sample t-tests and were compared with noise standards. The noise prediction models were developed using IBM SPSS Modeler software. Noise levels and ambient meteorological conditions were measured at 5-minute intervals for three days, extending 400 meters from the wind farm boundary in the northeast direction. Four individual models (CHAID, CART, Linear, and Neural network) and their ensemble were developed and compared. The models' inputs included distance, time, wind speed, wind direction, temperature, humidity, and pressure, with the output being the equivalent sound level.

The results from the noise maps showed that noise levels in residential areas ranged from 45 to 60 dB(A), in the wind farm area from 45 to 70 dB(A), and the agriculture area from 40 to 55 dB(A). The results from the monitoring stations indicated that noise levels in the southeast village ranged from 41.51 to 87.56 dB(A) and in the northwest community from 29.90 to 81.82 dB(A). Daytime noise levels were found to be significantly higher than nighttime noise levels. Comparing the measured noise levels with standards revealed that roadside communities were exposed to unacceptable noise levels at night according to the World Health Organization (WHO) recommendations, with traffic identified as a significant source of noise annoyance. The results from the noise prediction model showed that field measurements indicated that sound levels were higher closer to the wind turbines, particularly in the morning and evening, which indicates the influence of human activity.

The Ensemble model was identified as the most suitable technique based on the evaluation of R^2, MAE, and RMSE, providing RMSE (10.08%) and MAE (5.89%) during cross-validation for training and testing.

The analytical results revealed that traffic was the primary noise source, potentially posing health risks to villagers, emphasizing the need for mitigation measures to improve noise control. Decision-makers can utilize this information to identify areas requiring mitigation measures and minimize the nuisance of noise pollution.

School of <u>Environmental Engineering</u>   Student's Signature _____

Academic Year <u>2022</u>   Advisor's Signature _N. Tantemsapya_

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

Page

LIST OF TABLES

LIST OF FIGURES

TABLE OF FIGURES (Continued)

Figure   **Page**

LIST OF ABBREVIATIONS

ANN	=	Artificial Neural Network
CART	=	Classification and Regression Trees
CHAID	=	Chi-Squared Automatic Interaction Detection
EPA	=	Environmental Protection Agency
GIS	=	Geographic Information System
IDW	=	Inverse Distance Weighting
ISO	=	International Organization for Standardization
L_{Aeq}	=	Equivalent Continuous Sound Level
L_{10}	=	Sound pressure level exceeded 10% of the time
L_{90}	=	Sound pressure level exceeded 90% of the time
L_{np}	=	Noise pollution level
MAE	=	Mean Absolute Error
R^2	=	Coefficient of Determination
RMSE	=	Root Mean Square Error
SVM	=	Support Vector Machine
TNI	=	Traffic Noise Index
WHO	=	World Health Organization
WTB	=	Wind Turbine
WTN	=	Wind Turbine Noise

Chapter I

INTRODUCTION

1.1 Statement of Problem

Thailand's energy consumption has been increasing in recent years. The electricity demand rises by an average of 4.8% annually from 2011 to 2015 (Tunpaiboon, 2021). Because of this problem, the Thailand government developed an alternative energy development plan in 2015, "AEDP2015" (Department of Alternative Energy Development Efficiency, 2015), and updated in 2018 "AEDP 2018-2037" (Department of Alternative Energy Development Efficiency, 2020). AEDP 2018-2037 was developed and focused on promoting energy production within the full potential of domestic renewable energy resources, aiming to increase Thailand's renewable energy production by 30% in 2037. Wind energy is one of the renewable energies promoted by AEDP 2018-2037, where the production capacity target was set to 3.0 GW. Since starting AEDP, the production capacity has increased from 224.5 MW in 2014 to 1,027 MW in 2019. In the lower northeastern region of Thailand, the potential area for the wind farm is located at the elevated edge on the western side of the Korat plateau, which is in Nakhon Ratchasima and Chaiyaphum Province. Wind speed in the area ranges from 3 to 8 m/s, generating electricity at rate of 50 to 100 W/m^2 (Chancham et al., 2014). Huai Bong Sub-District, Dankhuntod District, Nakhonratchasima Province, is a potential area for wind farms, with three wind farms in the area.

The wind turbine is a device used to generate wind energy that converts kinetic energy from the wind into electricity. While wind turbines generate clean energy, annoyance, and health effects caused by wind turbine noise have drawn much attention from the public. There are two types of noise sources generated by wind turbines; 1) aerodynamic noise from the trailing edge of blades, inflow-turbulence noise, and airfoil self-noise, and 2) mechanical noise from generators and gearboxes (Kondili & Kaldellis, 2012, p. 515). Fyhri and Aasvang (2010) reported significant relationships between noise annoyance and sleeping problems, potentially affecting human health such as dizziness, anxiety, and depression. Michaud et al. (2018) reported that visual and auditory annoyance such as noise, blinking lights, shadow flicker, visual impacts, and vibrations was increased significantly with increasing wind turbine noise levels.

One of the wind farms in Huai Bong Sub-District locates close to the community (proximately 500 meters). The vicinity community complained about the effect of wind turbine noise such as annoyance, sleep disturbance, dizziness, and headache from this wind farm. The wind farm had installed two real-time online sound level monitoring stations at 500 m southeast and east of the boundary. However, sound level monitoring is not covering all vicinity community areas. Noise mapping is a technique used to create a visual representation of sound levels across a specific geographic area. It provides an effective means of assessing noise and understanding its distribution in areas where sensitive land use is a concern. This modern approach to evaluating noise levels facilitates the planning and implementation of strategies to mitigate the detrimental effects of noise pollution (Oyedepo et al., 2019; Pandya, 2003). To investigate the impact of noise from the wind farm over a large area, noise maps were utilized. These noise maps were generated using interpolation techniques such as IDW, kriging, and spline. ArcGIS Desktop 10.5 software was employed to develop noise maps.

Additionally, the noise prediction model is one option for investigating sound levels using machine learning. Machine learning is a powerful tool that uses algorithms to enable systems to learn patterns from data to make predictions (Madhavan, 2019). It has been widely used in applications of environmental pollution such as air pollution (Athanasiadis et al., 2003), water pollution (Bellinger et al., 2017), and noise pollution (Adulaimi et al., 2021; Singh et al., 2021). However, most of the studies in Machine Learning focus on noise pollution related to traffic, while there is insufficient research on wind turbines noise pollution. Data mining helps find patterns and predict noise. The systematic measurement of the sound level, wind speed, wind direction, temperature, moisture, and air pressure to an accumulation of extensive data in time series form. Building the models allows for investigating noise processes over various factors that prepare forecasts for noise levels. In this study, the noise prediction model was performed using the IBM SPSS Modeler 14.1 as a data mining and analytics software application. Using software to model statistics of various variables related to sound level with different algorithms. It was used to model the prediction as a neural net, linear regression, KNN algorithm, SVM, C&RT, and CHAID models, and they ranked each candidate model and scored to find the best analysis.

As mentioned above, there are many studies about the potential impacts of wind turbine noise on the community. However, the wind turbine sound level exposure model prediction of this site has yet to be done. Thus, this research aims to study the potential effect of wind farm noise on the community at 800 m. radius from the boundary of the wind farm. The results can be utilized to predict the sound level

of wind turbines and the propagation of wind turbine noise, providing valuable guidance for future research. Moreover, the findings from this study will contribute to the development of future noise regulations for wind turbines in Thailand.

1.2   Objective

The objectives of this study are;

1. To study the propagation of wind turbine noise by generating a noise map.

2. To investigate sound level with a noise prediction model.

1.3   Study area

The wind farm is in Huai Bong Sub-District, Dankhuntod District, Nakhon Ratchasima Province in Thailand. The study area is 800 m. radius from the boundary of the wind farm, as shown in Figure 1.1.

1.4   Scope and limitations

The scope and limitations of this study are;

1. The study confines to investigating the noise of wind turbines at 800 m. radius from the boundary of the wind farm.

2. The field measurement collects a sound level, measurement location, wind speed, wind direction, temperature, and humidity.

3. The noise prediction model was generated with several simplifying assumptions, including that wind turbines are aerial sources of noise and reflections are ignored. The model is representative of flat or constantly sloping terrain and does not consider the effects of terrain features such as hills, trees, and buildings that can influence sound propagation.

1.5   Expected Outcome

The results can be explained a wind turbine noise propagation and predict wind turbine noise level. In addition, it can be a guide for further study.
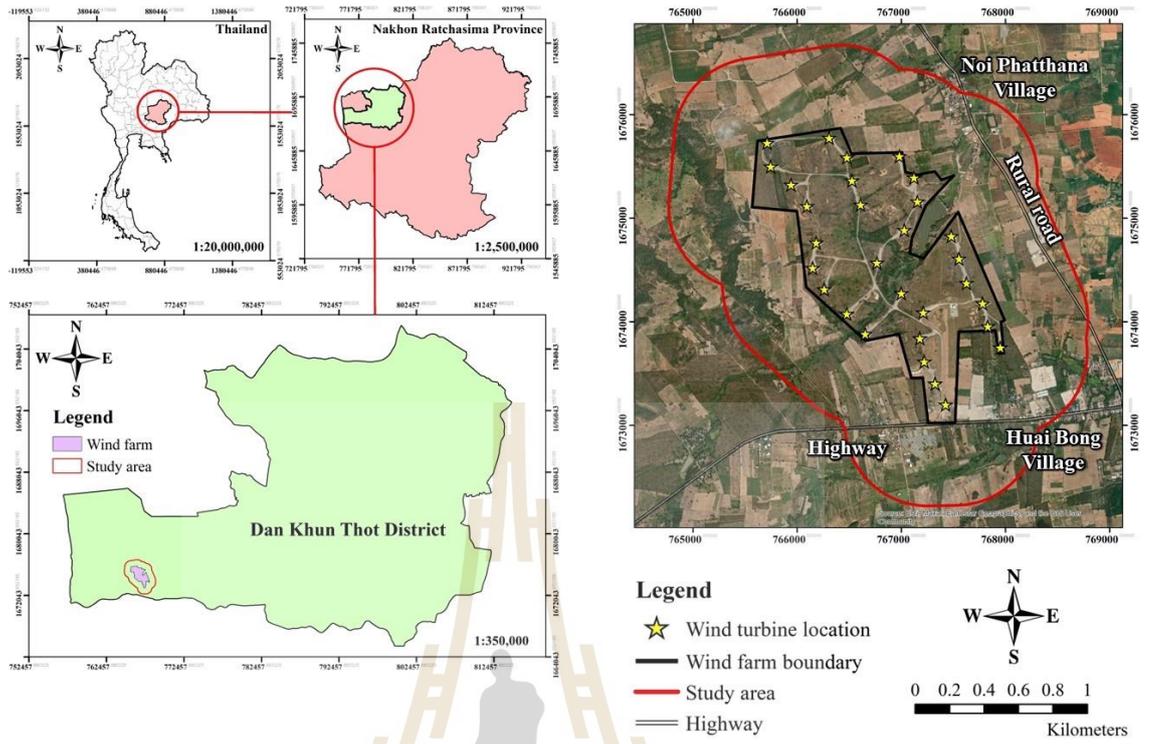
Figure 1.1 Study area

Chapter II

LITERATURE REVIEWS

2.1 Sound

2.1.1 Sound definition

Sound is a phenomenon produced by any object that is vibrating and transmitted through the medium as a pressure wave (sound wave). The number of vibrations or cycles per second is called Hertz (Hz). The range of sound frequencies that humans can be heard is approximately 20 to 20,000 Hz (Fahy & Thompson, 2015)

2.1.2 Basic properties of sound wave characteristics

A sound wave is a mechanical wave propagating through a medium; sound waves are characterized by amplitude, frequency, wavelength, and velocity (Hansen, 2001).

1) Amplitude (P_M) is the measure of the maximum displacement of particles in the medium from their resting position as the wave passes through. It changes in a single period of the sound wave. It determines the pitch of the sound. The amplitude is expressed in Pascal (Pa).

2) Frequency (f) is the number of oscillations or cycles of the wave in a single period of the sound wave. Pressure variation cycles per unit of time or cycles per second; the frequency is expressed in Hertz (Hz).

3) Wavelength (λ) is the distance of the pressure wave traveled during one cycle on the wave that is in phase, and the wavelength is expressed in the unit of length such as meter (m) or nanometer (nm).

4) Velocity (c) is the speed of sound propagation. The sound velocity in air is 343 m/s at 20°C and 1 atm, and the velocity is expressed in the unit of length per unit of time, such as meter/second (m/s).

2.1.3 Decibel scale and sound pressure level (SPL)

The sound level depends on the specific measure of sound, such as intensity, pressure, and power. It can be expressed as sound pressure level (SPL), which is the pressure of the sound wave, or as sound intensity level (SIL), which is the power of the sound wave per unit area, or as sound energy level (SEL), which is the total energy of the sound wave. These sound-level formulas result in decibels (dB), a logarithmic unit that references value (Long, 2014)

$$SPL = 20 \log_{10} (P/P_0) \qquad \text{(eq.1)}$$

When   P   =   Sound pressure is measured (in Pa)

           P_0 =   Reference sound pressure (2×10^{-5} Pa) for air.

                         Sound intensity level (SIL) is calculated by measuri2ng the sound intensity of a sound wave.

$$SIL = 10 \log_{10} (I/I_0) \qquad \text{(eq.2)}$$

When   I   =   Sound intensity is measured (in Pa)

           I_0 =   Reference sound intensity (1 pW/m²)

                         Sound energy level (SEL) is calculated by measuring the sound power of a sound wave.

$$SEL = 10 \log_{10} (E/E_0) \qquad \text{(eq.3)}$$

When   E   =   Sound energy being measured (in Pa)

           E_0 =   Reference sound energy (10^{-12} J)

The decibel readings are based on an exponential scale of sound pressure levels with a reference sound pressure. A 10 dB(A) increase in sound means 10 time-intensity greater. The sound level of common sounds rating in units of decibels is shown in Table 2.1.

Table 2.1 Decibel rating of common sounds

Sound pressure level (dB(A))		Sound description
188		Apollo lift-off, close
150		Jet engine, 10 ft away
140	Pain threshold	
130		Warning siren
125		Chain saw
120	Discomfort threshold	Thunder
115	Max under federal law	
110		Very loud music
105		Loud motorcycle or lawn mower
100	Very loud	Pneumatic air-hammer
90		The cockpit of light planes, heavy truck
85		Average street traffic
80		Lathe, milling machine, loud singing
75		Vacuum cleaner, dishwasher
70		Average radio, noisy restaurant
65	Annoying	

Table 2.1 Decibel rating of common sounds (Continued)

Sound pressure level (dB(A))		Sound description
60		Normal conversation, air conditioner
50		Light traffic, the average office
40		Library, quiet office
30		Quiet room in the home, audible whisper
20		Electric clock, the faint whisper
10	Barely detectable	Rustle of leaves
0	Hearing threshold	

Note: From Field and Long (2018).

2.1.4 Frequency weighting

The combination of different frequencies contributes to the overall sound. The weighting networks are used to evaluate frequency-weighting scales of the overall sound level in a sound measuring system (Hansen, 2001). The frequency-weighting scales are specified in IEC 60651, an international standard that sets out the methods for measuring sound pressure levels by frequency response curves for each weighting scale. The standard defines the A, B, and C frequency-weighting scales, and the sound level corrections for several weighing scales are shown in Figure 2.1.

The A-weighting: It is the most commonly used weighting widely for measuring environmental noise levels. The A-weighting applies a filter like a response of the human ear, which reduces the contribution of lower and higher frequency sounds that the average person cannot hear.

The B-weighting: It is designed to capture the effects of low-frequency sounds on structures and is used in building acoustics to measure the impact of noise on buildings.

The C-weighting: It is designed to capture the effects of high-frequency sounds on speech intelligibility and is used in audio engineering to measure the frequency response of audio equipment.

2.1.5 Sound measurement

The principle of sound measurement in this study follows the guidance note on noise assessment of wind turbine operations at EPA Licensed Sites (NG3). NG3 is a guidance note published by the environmental protection agency (EPA). It focuses on developing a standardized noise impact assessment methodology and assesses the impact of wind energy proposals on noise-sensitive locations (McAleer & McKenzie, 2011).

Measurement equipment: The basic equipment for continuous sound measurement is the sound level meter (SLM). The sound level meter is a handheld instrument with a self-contained kit and a precision microphone. The microphone responds to changes in air pressure from sound waves. The sound level meter standards class is specified by tolerance and accuracy and has Class 1 and Class 2.

1. the tolerance limits of Class 1 at the 1,000 Hz are +/- 1.9 dB(A)
2. the tolerance limits of Class 2 at the 1,000 Hz are +/- 2.2 dB(A)

Measurement positions: The measurement positions should be taken at the nearest noise- sensitive location affected by wind turbine noise. The measurement should avoid noise reflection impact by monitoring positioned at least 3.5 m away from the reflecting surface and high 1.2 to 1.5 m from the typical surface.

Measurement periods: The measurement periods should be a minimum of 50 ten-minute intervals for one week in a wind direction downwind from the turbine to the noise-sensitive location.
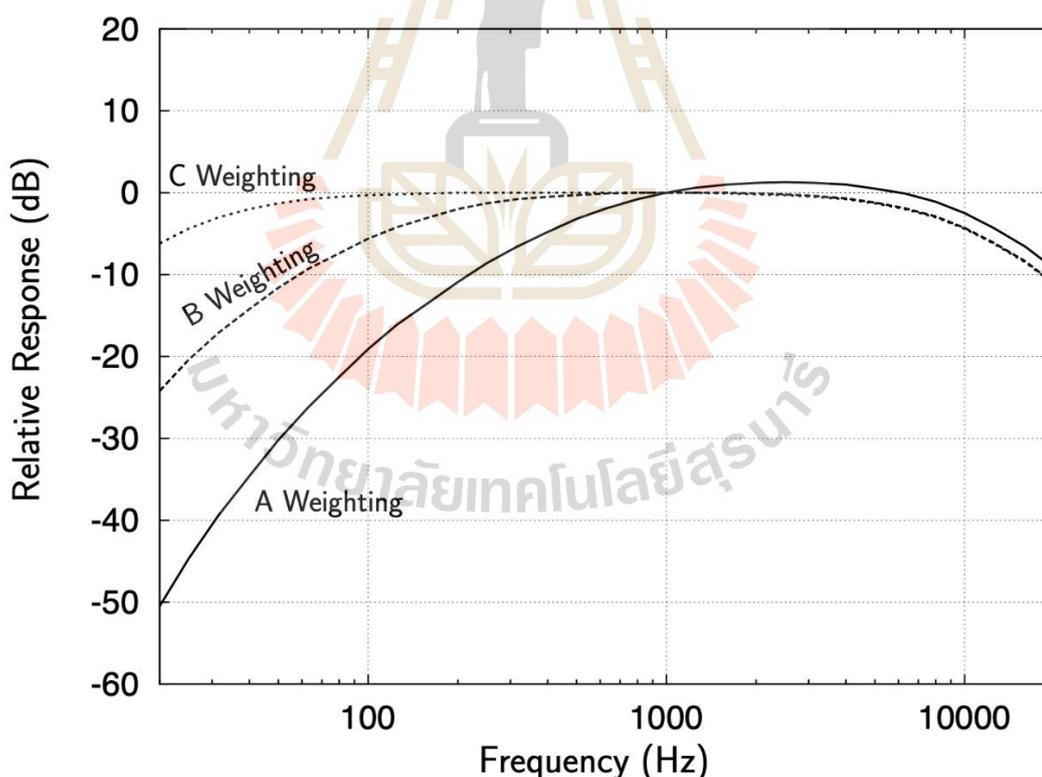
Figure 2.1 Frequency-weighting scales
from University of Alberta Faculty of Engineering (2020).

2.2   Noise

2.2.1   Noise definition

Noise is a sound defined as unwanted, annoying, unpleasant loud such as a distant train, whistle, or a neighbor's barking dog in the middle of the night. It can cause health problems, such as sleep disturbance, poorer work and school performance, hearing impairment, etc. (World Health Organization, 2010).

2.2.2   Noise categories

Noise can be considered a type of complex sound. Complex sounds are characterized by having multiple frequencies and varying amplitudes over time. They can be categorized into three main groups based on their sources: anthrophony, geophony, and biophony (Servick, 2014). Anthrophony refers to sounds produced by human activity. Geophony refers to sounds produced by non-living elements of the environment, particularly geophysical processes. Biophony refers to sounds produced by wildlife and other living organisms.

2.2.3   Common types of environmental noise

The common types of environmental noise encompass a wide range of sources that can contribute to noise pollution. World Health Organization (2022) defines environmental noise as unwanted sounds or a set of sounds that causes annoyance or has adverse health effects. Here are some examples of the sources that generate environmental noise:

Transport: This category includes noise generated by various modes of transportation, such as road vehicles, trains, airplanes, and ships.

Industrial activities: Industrial operations, including factories, manufacturing plants, and machinery, can generate significant noise.

Construction sites: Construction activities involving heavy machinery, equipment, and tools can create high noise levels.

Public works and services: Noise can arise from public works and services such as road repairs, maintenance activities, and utility services.

Cultural, sporting, and leisure activities: These activities can involve noise sources such as music from clubs, concerts, and festivals.

Neighborhood: Noise in residential areas can come from various sources. Outdoor sources may include heat pumps, motorized gardening equipment, and construction activities in nearby areas.

2.2.4   Effect of noise on human health

Environmental noise can have a range of detrimental effects on human health, affecting physical and psychosocial well-being (Bechtel & Churchman, 2003).

The physical effects of noise pollution include hearing damage, as continuous exposure to loud noise can lead to hearing loss or impairment. Sleep disturbances are also common, as noise during nighttime disrupts sleep patterns and can result in sleep deprivation, fatigue, and impaired cognitive function during the day. Exposure to excessive noise triggers a stress response in the body, elevating heart rate, blood pressure, and stress hormone levels, which can contribute to chronic health issues. High background noise levels can also interfere with effective communication, causing increased stress and frustration in social interactions. The psychosocial effects of noise pollution can impact mental and emotional well-being. Continuous noise exposure can cause annoyance and irritation, decreasing overall satisfaction with the environment and quality of life. Living in noisy environments can contribute to chronic psychological stress, resulting in reduced mental well-being. Noise distractions can impair concentration, productivity, and performance in tasks that require focus and attention. Furthermore, noise pollution can disrupt social activities and community interactions, diminishing the quality of social interactions and community cohesion (World Health Organization, 2022).

2.2.5 Noise Standards and Regulations

At present, there are no common international noise standards or regulations. The World Health Organization (WHO) has developed Environmental noise guidelines with recommended noise levels for protecting human health from environmental noise from various sources. Thailand and other countries have established standards for noise pollution from different activities for the daytime and nighttime, as shown in Table 2.2.

Table 2.2 Noise level standards of some given countries

Country/Organization	Industrial		Commercial		Residential	
	Day	Night	Day	Night	Day	Night
Australia[1]	55	55	55	45	45	35
India[1]	75	70	65	55	55	45
Japan[1]	60	50	60	50	50	40
Thailand[2]	80	80	-	-	70	70
United States of America[1]	70	60	60	50	55	45
World Health Organization[3]	65	65	55	55	53	45

Note: [1]   Chauhan and Pande (2010)

[2]   Notification of the National Environment Committee Issue 15 BE 2540, (1997), The Standard of Generic Sound Level

[3]   Hurtley (2009)

2.3   Wind Turbine Noise

2.3.1   Anatomy of a wind turbine

The wind turbine consists of four main elements: nacelle, rotor, tower, and footing (NSW Wind Energy Handbook, 2002).

Rotor: The rotor consists of a hub and blades with a shaft connecting them to the gearbox and generator. Wind turbines have three aerodynamically designed blades made of materials such as carbon fiber or fiberglass. These blades are optimized to maximize energy generation while minimizing noise. In fixed-speed turbines, the blade angle adjusts automatically to maintain a constant rotation speed, while variable-speed turbines rotate faster with increasing wind speeds. The rotor's primary function is to capture the wind's energy and convert it into mechanical energy for electricity generation. Blades may exceed 30 meters in length, giving a rotor diameter of 60 to 80 meters.

Nacelle: The nacelle is a large housing structure at the top of the tower. The gearbox and generator contain houses that convert the wind's kinetic energy into mechanical energy. The nacelle is typically designed to be aerodynamic and is responsible for connecting the rotor and the tower.

Tower: The tower is the tall structure that supports the entire wind turbine. It provides the necessary height to capture the stronger, more consistent wind speeds available at higher altitudes. Their height varies with the generator's size and the blades' length, and the large generators may have towers as high as 100 meters.

Footing: The footing of a wind turbine is a large concrete slab buried underground, typically with a diameter of 7-12 meters or more and a depth of 1-2 meters.

2.3.2   Source of wind turbine noise

When the wind flows past a wind turbine, the blades take the kinetic energy from the wind and rotate. The rotation speed of the wind turbine depends on wind speed and the specific design of the turbine. As the blades move through the air, this movement primarily produces noise. The sources of noise emitted from wind turbines operation can be divided into two categories; 1) Mechanical noise and 2) Aerodynamic noise (Kondili & Kaldellis, 2012)

1) Mechanical noise: Mechanical noise is noise that originates from the structure of the wind turbine and is emitted from its surfaces. It is caused by various mechanical components within the turbine, such as the gearbox, generator, yaw drives, cooling fans, and auxiliary equipment. This noise is generated by their relative motion, mechanical rotation, and the dynamic response between these components.

2) Aerodynamic noise: Aerodynamic noise is the noise produced due to the airflow interacting with the turbine blades. It occurs when the wind passes over the turbine's rotating blades, generating noise as a byproduct of this aerodynamic interaction.
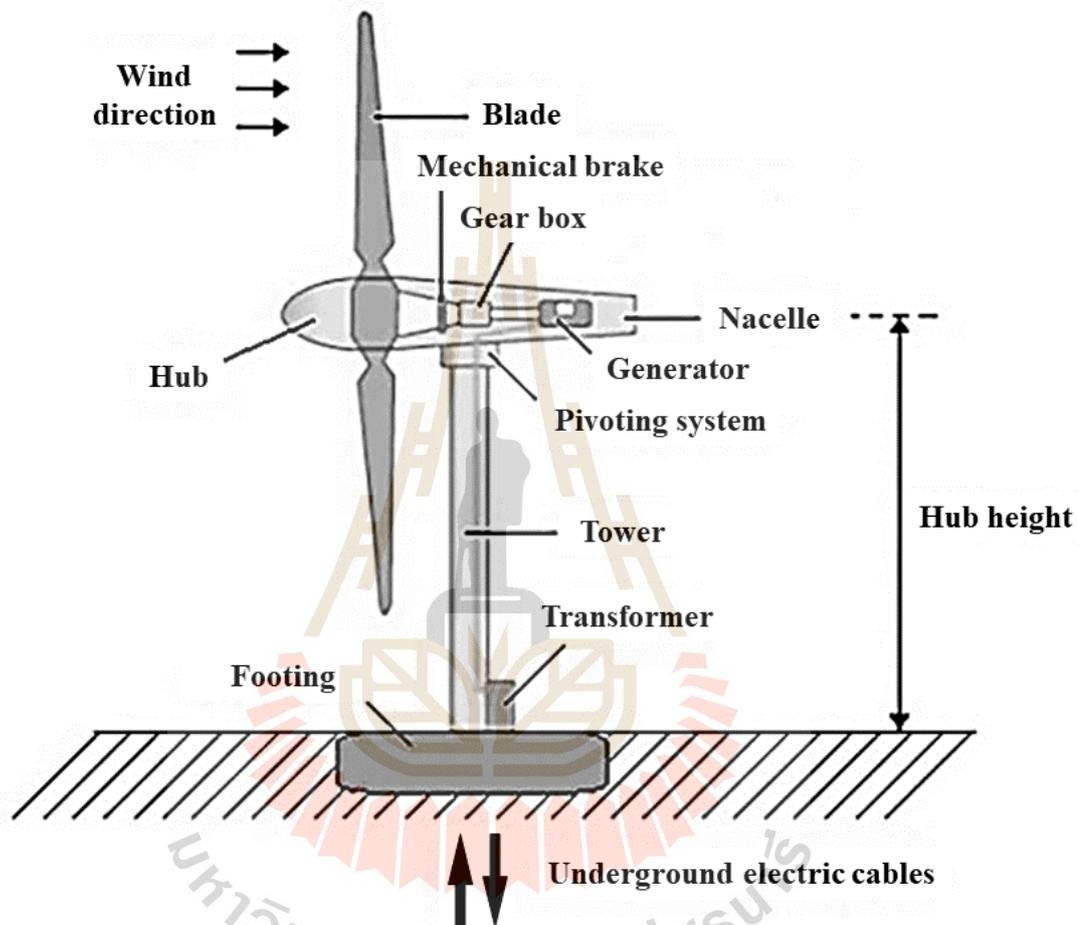
Figure 2.2 The main components of a wind turbine

from Năstase (2017)

2.3.3 Wind turbine noise characteristics

Wind turbine noise can be generated from four types of noise (Tonin, 2012). The types of noise are 1) tonal, 2) broadband, 3) low frequency, and 4) impulsive.

1) Tonal: Tonal is a discrete frequency noise generated by wind turbine components such as meshing gears. The tonal interaction with a rotor blade surface or unsteady flows is non-aerodynamic instability. It is often associated with the blade passing frequency, which depends on the blades' number and rotational speed. Tonal noise can manifest as a steady, like a hum or a whine.

2) Broadband: Broadband is a continuous distribution noise with frequencies greater than 100 Hz. It is an interaction of wind turbine blades with atmospheric turbulence. The aerodynamic interaction between the blades and the air generally causes broadband noise from wind turbines.

3) Low frequency: Low frequency is a noise with frequencies ranging from 20 to 100 Hz. It is associated with downwind rotors turbines. It is often described as a deep rumbling or vibrating sound.

4) Impulsive: Impulsive is short acoustic impulses or thumping sounds that vary in amplitude with time. These noise events can be caused by factors such as blade interactions with turbulent air, changes in wind conditions, or mechanical impacts within the turbine.

2.4   Noise Propagation

Wind turbine noise propagation refers to how the sound generated by wind turbines spreads and travels through the surrounding environment.

2.4.1   Factors affecting wind turbine noise propagation.

As wind turbines operate, they emit noise that can travel varying distances and be influenced by factors such as distance, atmospheric conditions, terrain, and the presence of barriers.

Distance: The distance between the wind turbine and the receiver affects the intensity of noise propagation. As sound waves travel further away from the source, leading to a decrease in noise levels. This phenomenon is known as sound attenuation. The inverse square law states that sound intensity decreases by approximately 6 dB for each doubling distance from the sound source (Gray PhD, 2000). The wind turbine generates a noise level that becomes equal to the background noise level when the wind speed is approximately 12 m/s and when the distance exceeds 100 m from the receiver (Katinas et al., 2016).

Wind Speed: Wind speed plays a significant role in wind turbine noise propagation. Higher wind speeds can increase the aerodynamic interactions between the wind and the turbine blades. This interaction can increase turbulence and higher rotor speeds, generating greater noise. The equivalent continuous sound level is highly correlated with the average rotor speed of a wind turbine (Sugimoto et al., 2008).

Wind Direction: The direction from which the wind is blowing can influence the path and dispersion of wind turbine noise. Sound waves tend to travel downwind more efficiently, following the wind flow. Wind direction affects the noise levels experienced in different directions. The average sound level in cross-wind

directions is lower than in upwind and downwind directions; the noise level is predicted within 1–2 dB in different wind directions (Oerlemans & Schepers, 2009).

Temperature: Temperature inversions, where a layer of warm air is trapped above cooler air near the ground, can increase noise propagation for sources near the ground. The increasing speed of sound in warmer air within the inversion layer allows sound waves to propagate more efficiently, potentially amplifying noise audibility (Zhou et al., 2013).

Terrain and noise barriers: Terrain features, such as hills, valleys, or vegetation, can cause sound waves to be reflected, diffracted, or absorbed. Natural or man-made barriers can significantly impact noise propagation from wind turbines. These barriers can block, deflect, or partially absorb sound waves. Typical terrain and noise barriers tend to absorb energy from incident acoustic waves and reflective properties of the surface (Attenborough, 2002).

Air absorption: When sound travels through the air, it gets absorbed due to two main reasons: molecular relaxation and air viscosity. Molecular relaxation is the transition of a molecule going from an excited energy level to a lower excited level. High-frequency sounds are absorbed more than low-frequency sounds because their waves are shorter. The absorption occurs because of the friction between air particles as the sound wave moves through the air. The absorption depends on the temperature and humidity of the atmosphere (Pantazopoulou, 2010).

Ground surface conditions: When the sound hits the ground, the acoustic energy loss depends on the reflection coefficient of the surface. Sound waves lose some of their energy through reflection on hard surfaces, resulting in attenuation. (Pantazopoulou, 2010).

2.4.2 Noise propagation calculation

Noise propagation calculation involves analyzing the spread of sound waves and how sound pressure changes with distance in a medium, typically air. This includes considering the sound intensity at any distance from the source and applying the inverse square law calculation for sound, where sound pressure decreases proportionally to the square of the distance.

Sound intensity at any distance from the source

Sound intensity (I) represents sound power per unit area. As the distance from the source increases, the sound intensity decreases since the sound power spreads out over a larger surface area. The sound intensity at any distance from the source can be calculated using the equation 4.

$$I = \frac{P}{A} = \frac{P}{4\pi r^2}$$ (eq.4)

When   P   =   Power of the source (in W)

         A   =   the surface area through which the sound waves pass. (in m^2)

Inverse square law calculation for sound

The Inverse square law for sound attenuation describes how sound intensity diminishes with distance. It is inversely proportional to the square of the distance from its source. This law demonstrates that sound intensity decreases significantly as the distance from the source increases. The sound intensity can be calculated using the equation 5.

$$I_2 = I_1 \left(\frac{D_1}{D_2} \right)^2$$ (eq.5)

When I_2 =   Sound intensity at a desired distance (in W/m^2).

         I_1 =   Sound intensity at a reference distance (in W/m^2).

         D_2 =   Desired distance (in m).

         D_1 =   Reference distance (in m).

Sound propagation is emitted from a source in all directions, spreading out in a spherical manner. As the sound waves travel, the sound pressure level decreases with increasing distance according to the inverse square law. The sound pressure level can be calculated using the equation 6.

$$L_p = L_w - 10Log_{10}(2\pi r^2)$$ (eq.6)

When L_p =   Sound pressure level at a particular distance from the source (in dB(A)).

         L_w =   Sound pressure level of the source (in dB(A)).

         r   =   the distance from the sound source (in m).

The ISO 9613-2 "Acoustics – Attenuation of sound during propagation outdoors – Part2: General method of calculation" (ISO, 1996) is standard that provides guidelines for calculating the attenuation, or reduction, of sound as it propagates outdoors. It can be applied to different sound sources and covers the major mechanics of sound attenuation (Blanchard & Samanta, 2019).

L_{fT}(DW) is the equivalent continuous downwind octave band sound pressure level at a receiver location and is calculated for each point source in downwind conditions based on equation 7 (ISO, 1996).

$$L_{ft}(DW) \quad = \quad Lw + D_C - A \qquad\qquad (eq.7)$$

When $L_{ft}(DW)$ = the equivalent of continuous downwind sound pressure levels at receiver locations (in dB(A)).

Lw = the sound levels produced from the source (in dB(A)).

D_C = the directivity correction, the index of the sound levels propagates into solid angles (in dB(A)).

A = attenuation during propagation from the source to the receiver (in dB(A)).

The attenuation term (A) in equation 7, which accounts for the decrease in sound intensity over distance, is determined by equation 8 (International Organization for Standardization, 1996).

$$A \quad = \quad A_{div} + A_{atm} + A_{gr} + A_{bar} + A_{misc} \qquad (eq.8)$$

When A_{div} = the attenuation due to geometrical divergence

A_{atm} = the attenuation due to atmospheric absorption

A_{gr} = the attenuation due to the ground effect

A_{bar} = the attenuation due to a barrier

A_{misc} = the attenuation due to miscellaneous other effects (noise propagating through buildings).

$L_{AT}(DW)$ is the equivalent continuous A-weighted downwind sound pressure level that can be calculated by summing the contributing time mean square sound pressures calculated according to equations 7 and 8 for each point sound source, as specified by equation 9 (International Organization for Standardization, 1996).

$$L_{AT}(DW) \quad = \quad 10log\left\{\sum_{(i=1)}^{n}\left[\sum_{(j=1)}^{8}10^{0,1\left[L_{fT}(ij)+A_f(j)\right]}\right]\right\} \qquad (eq.9)$$

When n = the number of contributions i (sources and paths)

j = the index indicating the eight-standard octave-band mid-band frequencies from 63 Hz to 8kHz.

A_f = the denotes the standard A-weighting.

2.4.3 Regulations of noise from Wind turbine source

The international regulations for noise from wind turbines are not very uniform across different countries, although there are some general similarities among

many countries. The noise levels produced by each wind turbine generator tend to increase as the wind speed at the site increases. Additionally, the background noise often rises under such conditions, which can overlap with the noise generated by the wind turbine. The regulations of noise from wind turbine sources across different countries were obtained from a study conducted by Licitra and Fredianelli (2013).

United Kingdom, The ETSU-R-97 standard sets noise limits for wind turbines based on a combination of fixed and derived limits. The fixed limit During the daytime, the noise limit can vary within the range of 35 to 40 dB, and during nighttime, it is set at a minimum of 43 dB L_{A90}, while the derived limit considering the background noise levels and adds 5 dB.

France, The Decret 2006-1099 of 2006 provides guidance and regulations regarding the measurement and management of environmental noise. The existing guidance specifies that any new noise generated by wind turbines should not exceed the existing noise level by more than 5 dB during the day and 3 dB at night.

Germany, the noise limits are based on different areas. In industrial areas, the noise limit is 65 dB(A) during daytime and 50 dB(A) during nighttime. In residential areas, the noise limit is 50 dB(A) during daytime and 35 dB(A) during night.

Netherlands, the Dutch regulation published in 2001 provides guidelines for environmental management, including noise limits dependent on wind speed. At a wind speed of 12 m/s, the noise limit is 50 dB(A) during the daytime, 45 dB(A) during evenings and 40 dB(A) during nighttime.

New Zealand, the wind turbine noise should not exceed the background sound level by more than 5 decibels or a level of 40 dB L_{A90} (10 minutes). However, for locations classified as sensitive areas, the noise limit is further reduced to 35 dB L_{A90} (10 minutes) to ensure minimal disturbance.

South Australia and New South Wales (Australia), The predicted equivalent noise level, evaluated at all relevant receivers for each integer wind speed from cut-in to rated power of the wind turbines, should not exceed 35 dB(A) or exceed the background noise by more than 5 dB(A).

Denmark, the noise limits for wind turbines are determined based on wind speed and land use. In outdoor living areas located no more than 15 meters from open countryside, the noise limit is 44 dB(A) at a wind speed of 8 m/s and 42 dB(A) at a wind speed of 6 m/s. In noise-sensitive land use areas, the noise limit is lower, set at 39 dB(A) at a wind speed of 8 m/s and 37 dB(A) at a wind speed of 6 m/s.

Canada, the noise limits for wind turbines are determined based on wind speed and land use. In urban areas, the noise limit ranges from 45 to 51 dB(A),

while in rural areas, it ranges from 40 to 51 dB(A). These limits depend on wind speeds within the range of 6 to 10 m/s at a height of 10 meters.

Thailand, the "Announcement of the Energy Regulatory Commission Regarding the Determination of Distance for Wind Power Generation Projects and Installed Capacity for Wind Power Generation Operators" provides regulations on sound impact that the maximum allowable noise level must not exceed 10 dB(A). This requirement is specified in the announcement of the National Environmental Committee, Version 29 (B.E. 2550), and aligns with the IEC 61400-11 standard. The measurement is taken from the land area of the nearest residential dwelling or house within the community.

2.5   GIS noise mapping

GIS-based mapping has expanded in popularity in recent years, with applications in nearly every field and increased geographic data availability. It has been widely and successfully used in environmental impact studies to assess the impact of spatial phenomena such as soil pollution, air pollution, and noise on the environment. Noise mapping has been applied in several sites using GIS, such as urban planning, public health (Moteallemi et al., 2017; Oyedepo et al., 2019; Tsai et al., 2009) transportation planning (Forouhid et al., 2023), and industrial noise control (Bozkurt, 2021).

GIS software is a powerful tool to create maps that visualize information about the noise characteristics and their variations in the surrounding environment (Alam, 2011). Interpolation is the most important technique for noise mapping, and it can be used to develop contours of noise levels (Yilmaz & Hocanli, 2006). The interpolation technique takes into consideration the acoustic behavior of the topographical region. Noise contour maps can be created to show the variations in environmental noise at different times of the day in urban areas. GIS can be used to create noise contour maps that help identify areas with high noise intensity and traffic noise. It also highlights the zones most affected by noise pollution. To achieve accurate noise mapping, a clear methodology can be followed, which involves the following steps:

2.5.1   Global posting system data collection

Various methods can be used to collect precise GPS data, depending on factors such as the survey objectives, required accuracy, available equipment, and logistical considerations. Common GPS survey methods include Continuous, Static, Rapid Static, and Kinematic survey techniques. The noise data collected in the field can be integrated into GIS and displayed on a map of the urban area. The distance

between data points on the map may vary based on the level of human habitation in different regions. Each data point contains information like coordinates, location, date and time of data collection, main noise sources, noise indicators, maximum and minimum recorded noise levels, and average noise level.

2.5.2 Spatial database development

A spatial database is a collection of information organized in table form. The tables in the database are structured based on the sensitivity of the collected survey data. The spatial database is built from four types of spatial data: GPS noise locations, noise level readings, noise sources, and noise impacts. GPS noise locations can be used to identify the geographical points where noise levels were recorded. Each location is assigned a unique identifier that serves as a reference to connect the entire database. Noise level readings are described in decibels (dB). Noise sources provide information about the major sources of noise, while noise impacts study the effects of noise pollution on human health and behavior.

2.5.3 Spatial modeling

Spatial modeling can be defined as the number of grids or polygons that are aggregated to a particular form of an area. This modeling technique can be linked to GIS for data input and display. There are two main types of spatial modeling techniques: vector and raster. These techniques are applied within GIS tools to determine the spatial distribution of noise pollution.

2.5.4 Interpolation methods used in noise mapping

Interpolation methods are commonly used to estimate noise levels at unsampled locations based on measured data from monitoring stations or other sources. Interpolation helps create continuous noise maps that provide a spatial representation of noise levels throughout an area of interest. Several interpolation methods are utilized, including IDW, kriging, Gaussian Process Regression, Spline, and Radial Basis Functions. Among these methods, two popular and commonly utilized approaches for noise mapping are Inverse Distance Weighting (IDW) and Kriging, including the following:

Inverse Distance Weighting (IDW): IDW is a widely used interpolation method in noise mapping. It assigns weights to nearby measured points based on their distance to the target location (Wu & Hung, 2016). The weights are inversely proportional to the distances, meaning closer points have more influence on the interpolated value. the size of the search neighborhood is directly related to the distribution of reference points in the region and the distances between these points (Figure 2.3). IDW assumes a smooth spatial variation and is simple to implement, but

it does not consider the spatial correlation between points.
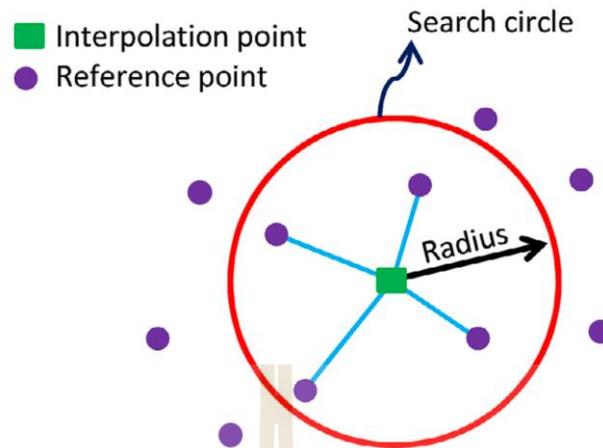
Figure 2.3 How IDW works

Adapted from Harman et al. (2016)

The power parameter is used to control the influence of nearby data points on the interpolated values. It determines the rate at which the weights assigned to neighboring points decrease with distance. The power parameter allows you to adjust the influence of the data points, emphasizing either closer points or giving more weight to points farther away. The value at the interpolation point from the IDW method can be calculated using equation 10.

$$N_0 = \frac{\sum_{i=1}^{n} N_i P_i}{\sum_{i=1}^{n} P_i}$$ (eq.10)

When N_0 = the value at interpolation point

N_i = the value at reference point

P_i = the power parameter determines the weight of the value at the reference point

n = the number of measurement points

Inverse Distance Weighting has several advantages, including its simplicity, ease of understanding, and efficiency. However, it is sensitive to outliers and lacks an indication of error (Longley, 2005).

Kriging: Kriging is a geostatistical interpolation method that considers both spatial correlation and spatial trend in the data (Wu & Hung, 2016). Kriging is a geostatistical method similar to IDW, but unlike IDW where weights are determined based only on the inverse of distances, kriging considers both proximity and spatial correlation when assigning weights to data points for estimation (Harman et al., 2016). It estimates the values at unsampled locations by considering the neighboring point values and their spatial relationships. Kriging is an interpolator that can be exact or

smoothed depending on the measurement error model. It is very flexible and allows for the investigation of spatial auto correlation and cross-correlation graphs.

Kriging utilizes statistical models that provide various output surfaces, including predictions, prediction standard errors, probability, and quantiles. Using kriging requires making careful decisions due to its flexibility. Kriging assumes that the data is from a consistent random process, and some methods assume the data follows a normal distribution. Kriging includes several variants such as ordinary kriging, simple kriging, universal kriging, and cokriging. The value at the interpolation point from the Kriging method can be calculated using equation 11.

$$\hat{Z}(x_0) = \sum_{i=1}^{n} w_i(x_0)Z(x_i) \qquad (eq.11)$$

When $\hat{Z}(x_0)$ = the value at interpolation point

$Z(x_i)$ = the value at reference point

$w_i(x_0)$ = the power parameter determines the weight

of the value at the reference point

n = the number of measurement points

Ordinary Kriging (OK): Ordinary kriging is one of the most widely used variants of kriging. It assumes that the mean value is unknown and estimates it from the data. OK provides optimal estimates by minimizing the estimation error variance. It is suitable for cases where the mean value varies spatially.

Simple Kriging (SK): Simple kriging assumes a known constant mean value. Unlike ordinary kriging, it does not estimate the mean from the data. SK is appropriate when the mean is known and constant across the study area. It is less commonly used compared to ordinary kriging.

Universal Kriging (UK): Universal kriging expands on ordinary kriging by incorporating additional covariates or trend variables that influence the spatial variation. It allows for modeling systematic trends or spatially varying means in the data. By including these covariates, UK can capture more complex spatial patterns and provide improved estimates.

Co-kriging: Co-kriging, also known as multivariate kriging, is used when multiple variables are correlated and available for analysis. It extends the principles of kriging to estimate one variable based on the values of other related variables. Co-kriging takes advantage of the spatial relationship between variables to improve the estimation of each variable of interest.

Inverse Distance Weighting (IDW) and Kriging are commonly used spatial interpolation techniques that can be evaluated based on a comprehensive analysis of various referenced studies. The advantages and disadvantages of Inverse Distance

Weighting and Kriging are based on the various referenced studies in term of interpolation study.

Advantages of Inverse Distance Weighting:

Schloeder et al. (2001) concluded that IDW performs similarly to kriging and is generally more accurate than spline interpolation.

Lu and Wong (2008) developed a new form of IDW that estimates data values at unsampled locations based on spatial patterns found in their neighborhood, potentially enhancing its performance.

Disadvantages of Inverse Distance Weighting:

Kravchenko (2003) reported that the accuracy of IDW interpolation performance is significantly affected by the presence of spatial structure. And the variograms have a significant potential to enhance kriging performance more than they enhance IDW performance.

Harman et al. (2016) reported that IDW produces better results with a smaller search circle radius and a homogeneous distribution of data.

Advantages of Kriging:

Schloeder et al. (2001) concluded that IDW performs similarly to kriging and is generally more accurate than spline interpolation.

Bishop and McBratney (2001) found that kriging process can enhance its performance by including secondary data, such as color aerial photos.

Kravchenko (2003) reported that variograms have a significant potential to enhance kriging performance, resulting in better results compared to IDW.

Harman et al. (2016) reported that the choice of variogram model in the Kriging Method significantly impacted the results. The Cubic variogram model consistently outperformed the Exponential, Linear, and Quadratic models across all grid resolutions.

Disadvantages of Kriging:

Mueller et al. (2001) reported that the performance of kriging depends heavily on the existence of spatial structure and sampling density, although there is little overall difference in performance between IDW and kriging.

Bekele et al. (2003) concluded that while kriging generally performs better than IDW, a regression-based autocorrelated error model offers greater flexibility for interpolation.

2.5.5 Geostatistical Analyst

The Geostatistical Analyst is a geostatistical tool that is integrated with GIS modeling environments. It enables GIS professionals to accurately measure

the statistical error of predicted surfaces. The Geostatistical Analyst involves three key steps (Johnston et al., 2001) :

1. Exploratory spatial data analysis
2. Spatial structural analysis
3. Surface prediction and assessment of results

1) Exploratory spatial data analysis

Exploratory spatial data analysis is utilized to explore the distribution of data, search for outliers and trends, and examine spatial autocorrelation. The tools used in exploratory spatial data analysis include Histogram, Voronoi Map, Trend Analysis, and Semivariogram/Covariance Cloud.

1.1) Histogram tool

The histogram is a tool that provides a one-variable description of your data. It displays the frequency distribution for the dataset of interest and calculates summary statistics.

Normal distribution examination: Examining the normal distribution of data in geostatistical analysis is important to ensure that the data meets the required assumptions for certain methods. This involves assessing measures such as mean, median, skewness, and kurtosis. If the data deviates from a normal distribution, applying transformations to bring it closer to a normal distribution can improve the accuracy of the analysis.

Outlier detection: Detecting outliers is important in geostatistical analysis as they can negatively impact prediction surfaces and semivariogram modeling. The histogram tool helps identify potential outliers located in the tails of the distribution. Further investigation is needed for isolated extreme values surrounded by significantly different values. Correcting or removing incorrect outliers due to data entry errors is essential for more accurate analysis.

1.2) Voronoi map tool

The voronoi map tool help identify and analyze local outliers, which deviate from the surrounding points within the normal range of the dataset. These maps are created by establishing shared polygon borders around the sample points, ensuring that any location within a polygon is closer to its corresponding sample point than to any other point.

1.3) Semivariogram/Covariance Cloud

The semivariogram/covariance cloud is a valuable tool for analyzing spatial autocorrelation within a dataset. It visually represents the empirical

semivariogram for pairs of locations based on their distance. This analysis helps in evaluating the characteristics of spatial autocorrelation, examining spatial correlation patterns, exploring examining directional influences, and identifying outliers.

1.4) Trend analysis tool

The trend analysis tool offers a valuable approach for detecting global trends in the dataset. It visualizes the data in a three-dimensional representation, with sample point locations displayed on the x-y plane and the values represented by the height of sticks in the z-dimension. By drawing a best-fit line (polynomial) through the projected points, trends in specific directions can be modeled. A flat line indicates the absence of a trend. If a trend is identified, it must be eliminated to ensure data stationarity, a requirement for employing kriging as an interpolation method.

2) Spatial structural analysis

In the geostatistical analysis using the geostatistical wizard, various parameters such as the input data layer, attribute field, kriging type, data transformation, and trend type were selected. Semivariogram models were then developed for each combination by determining the appropriate lag size, fitting a spherical semivariogram model, and calculating parameter values. Directional influences were taken into account by utilizing the directional search tool to develop an anisotropic semivariogram model.

3) Surface prediction and assessment of results

Geostatistical techniques employ statistics to create surfaces that incorporate the statistical properties of measured data. These techniques, including various kriging methods such as ordinary, simple, universal, probability, indicator, and disjunctive kriging, along with cokriging, not only generate prediction surfaces but also provide error or uncertainty surfaces, allowing for the assessment of prediction accuracy. The Geostatistical Analyst offers tools to determine appropriate parameters for the analysis. In the kriging process, the spatial structure of the data is quantified through variography, where a spatial-dependence model is fitted to the data. To make predictions for unknown locations, kriging utilizes the fitted model, the spatial data configuration, and the values of nearby sample points. These methods can produce prediction and error surfaces.

The assessment of results is essential for obtaining unbiased and accurate predictions of parameter values, as well as valid prediction standard errors.

In a model that produces unbiased predictions, the mean of the prediction errors should be close to zero. Evaluating assessing prediction accuracy involves considering the root-mean-square standardized prediction error, which should be close to 1, and the average standard error, which should be minimized or close to zero.

2.6   Noise prediction model with machine learning

Machine learning is a branch of artificial intelligence (AI) that uses data and algorithms to improve accuracy (Madhavan, 2019). It is a powerful tool that uses algorithms to enable systems to learn patterns from data to make predictions. It incorporates analysis and forecast using statistical models, machine learning, and mathematical algorithms, such as neural networks or decision trees. These platforms enable researchers to apply advanced algorithms and statistical techniques to predict data, such as MATLAB, R, Python, SAS, IBM SPSS Modeler, Microsoft Azure ML, and Apache Spark ML.

Machine learning is an effective technique for predicting and evaluating environmental pollution. In the field of air pollution, data mining, and machine learning algorithms are being increasingly utilized to analyze large datasets and identify patterns and correlations such as air pollutants, air concentrations epidemiology, air conditions, and health outcomes (Athanasiadis et al., 2003; Bellinger et al., 2017). In the field of water pollution, machine learning algorithms have been applied to assess and predict water quality in various water environments, that have been applied to evaluate the water quality in different water environments, such as surface water, groundwater, drinking water, sewage, and seawater (Bellinger et al., 2017). Furthermore, machine learning approaches have also been employed to analyze noise pollution levels. Kumar et al. (2014) proved the Artificial Neural Network (ANN) approach as a powerful technique for traffic noise modeling by replacing linear regression analysis with advanced modeling techniques such as ANN. Not only has ANN been used to predict traffic noise, but decision trees, random forests, generalized linear models, and artificial neural networks are also used to predict traffic noise (Adulaimi et al., 2021; Singh et al., 2021).

2.6.1   Algorithms of machine learning

Algorithms are computational models designed to make predictions or forecasts based on existing data. These algorithms utilize various mathematical and statistical techniques to analyze patterns and relationships within the data enabling them to predict outcomes from unseen data points. There are two main types of prediction algorithms: classification and numeric algorithms.

Classification algorithms: Classification algorithms are machine learning algorithms that categorize or assign labels to data points based on their features. The classification algorithm is used to build a model that can accurately predict the category of new data points, such as decision trees, Support Vector Machines (SVM), k-nearest neighbors (k-NN), CHAID, and random forests.

Numeric algorithms: Numeric algorithms or regression algorithms are a type of machine learning algorithm used to predict continuous numeric values based on input features. These algorithms analyze the relationships between the input and target variables to create a model that can estimate the new numeric value, such as linear regression, decision tree regression, support vector regression, and neural networks.

These descriptions provide a more detailed understanding of the mechanisms employed by each algorithm and how they operate and make predictions based on the given data. The information is sourced from Ambika (2020), Syed Muzamil and Dharmendra Singh (2019), and The International Business Machines Corporation (2021b).

Decision Trees: Decision trees are hierarchical structures where each node represents a feature or attribute, and each branch represents a decision based on that attribute. It is utilized for classification and regression tasks. The tree is constructed by recursively partitioning the data based on feature values to minimize or maximize information gain at each step. A decision tree starts with a root node, which does not have any incoming branches.
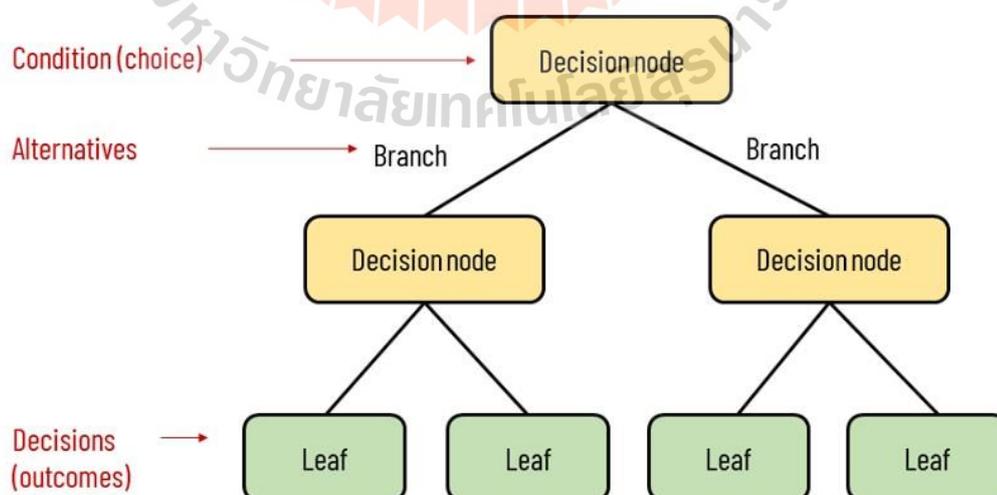
Figure 2.4 Elements of decision tree diagram
From Kosarenko (2021)

Support Vector Machines (SVM): SVM constructs hyperplanes or decision boundaries in high-dimensional space to separate data points from different classes. Its objective is to maximize the margin between the hyperplane and the nearest data points, known as support vectors. By identifying this hyperplane, SVM can successfully generalize to unseen data points, providing reliable predictions. SVMs are versatile and applicable to both classification and regression tasks. They are commonly encountered in various fields, such as bioinformatics, image recognition, and text analysis.
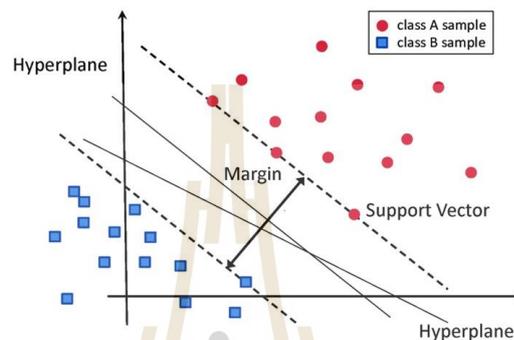
Figure 2.5 Classification of data by support vector machine (SVM)

Adapted from García-Gonzalo et al. (2016)

K-nearest neighbors (k-NN): k-NN classifies data points by considering the majority vote of their k nearest neighbors in the feature space. The distance metric is employed to determine the proximity between data points. Class labels are assigned based on the title that appears most frequently among the k nearest neighbors of a data point, a technique commonly referred to as "majority voting" and widely used in literature. It is worth noting that k-NN can be applied to both classification and regression tasks. The main difference between classification and regression is that classification is employed for predicting discrete values or categories, while regression is used for estimating continuous values.
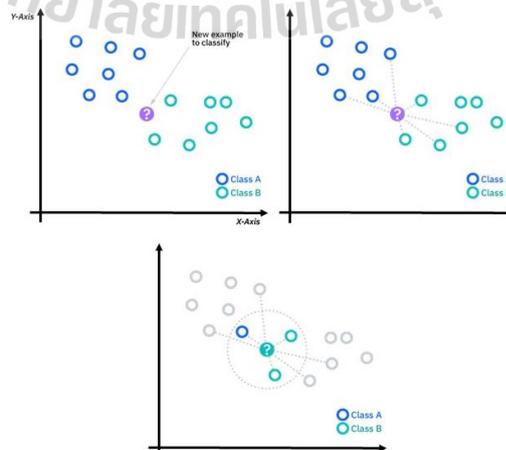
Figure 2.6 K-nearest neighbors diagram

Adapted from The International Business Machines Corporation (2018)

Chi-squared Automatic Interaction Detection (CHAID): CHAID is a decision tree-based algorithm that uses the chi-squared test for statistical significance to determine the best attribute for splitting the data at each step. It is particularly suitable for categorical or nominal target variables. CHAID can produce nonbinary trees, allowing splits with more than two branches, resulting in broader trees compared to binary growing methods. This algorithm is compatible with various input types and accommodates case weights and frequency variables.

Random Forest: Random Forest is an ensemble learning technique that combines the predictions of multiple individual decision trees to enhance accuracy and robustness. By creating a random subset from the training data and constructing decision trees based on different features and splitting criteria, Random Forest mitigates overfitting and improves generalization. The final prediction is obtained by aggregating the individual tree predictions through majority voting or averaging. It is utilized for classification and regression tasks.
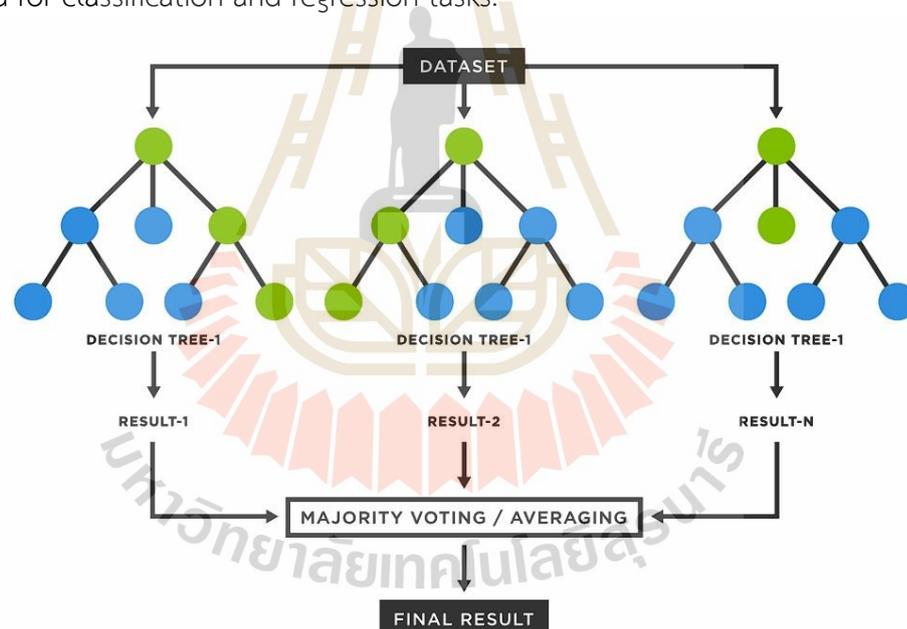
Figure 2.7 Random Forest algorithm diagram

Adapted from Sharma (2020)

Linear Regression: Linear regression establishes a linear relationship between independent variables (features) and a dependent variable (target). It aims to find the best-fit line that minimizes the sum of squared differences between the observed and predicted values. The algorithm works by fitting a linear equation to the training data, minimizing the sum of squared differences between the observed and predicted values. Linear Regression finds the best-fitting line for a relationship between the variables.
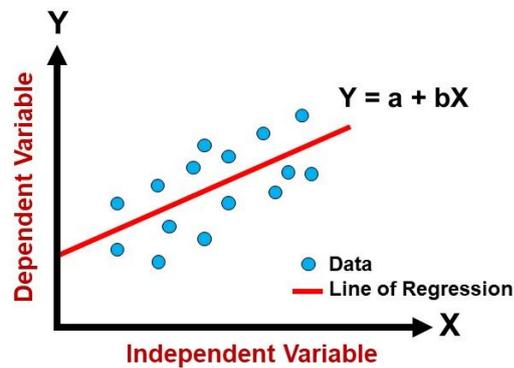
Figure 2.8 Linear regression algorithm

Decision Tree Regression: Decision tree regression is similar to decision trees for classification but is used for predicting continuous numeric values. The predicted value for a new data point is the average or weighted average of the target variable within the leaf node. Each internal node in the tree represents a feature or attribute, while the leaf nodes provide the predicted numerical value. The splitting process is based on metrics such as mean squared error or mean absolute error. This method aims to minimize the overall prediction error.
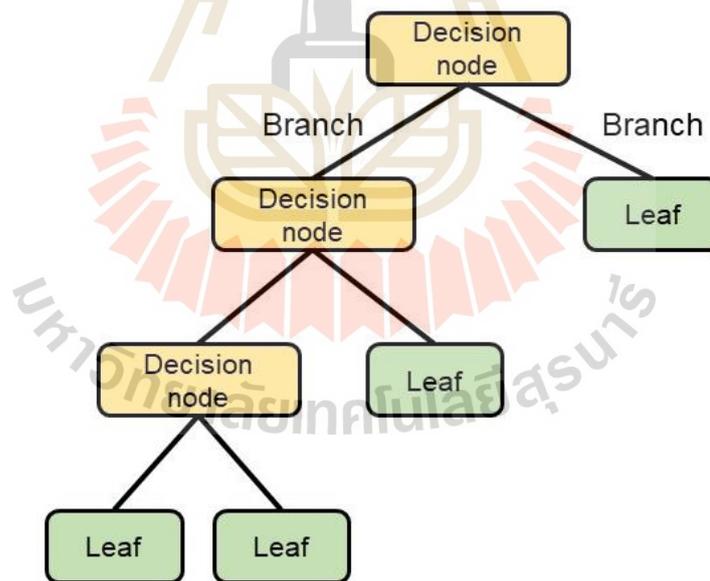
Figure 2.9 Decision tree regression algorithm diagram

Support Vector Regression: Support Vector Regression (SVR) is an extension of SVM for regression problems. It uses support vectors and hyperplanes to perform regression and estimate continuous values. Both SVM and SVR utilize the concept of support vectors, which are the data points that are closest to the decision boundary or hyperplane. SVR aims to find the hyperplane that minimizes the error between the predicted and actual continuous values.

Neural Networks: Neural networks, also known as Artificial Neural Networks (ANNs), consist of interconnected layers of artificial neurons called nodes or units, inspired by the structure and functioning of the human brain. They process data through weighted connections, apply activation functions, and have the ability to learn and recognize complex patterns and relationships in the data. Neural networks can handle both classification and regression tasks and are particularly effective in learning from large and high-dimensional datasets.
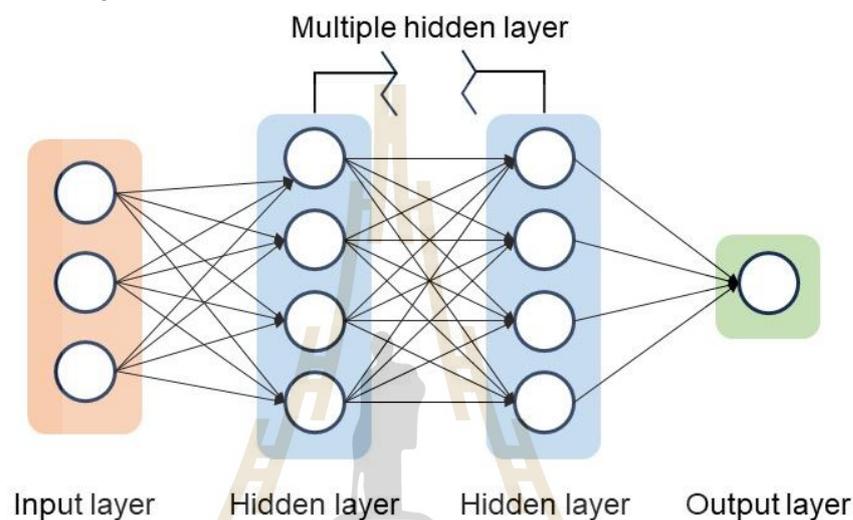
Figure 2.10 Neural networks algorithm diagram

Ensemble methods: Ensemble methods combine multiple models to make predictions. Random Forests, for example, create an ensemble of decision trees, where each tree is trained on a subset of the data. The final prediction is determined by aggregating the predictions of individual trees. Ensemble methods improve prediction accuracy and generalization by leveraging the diverse perspectives and collective wisdom of multiple models.
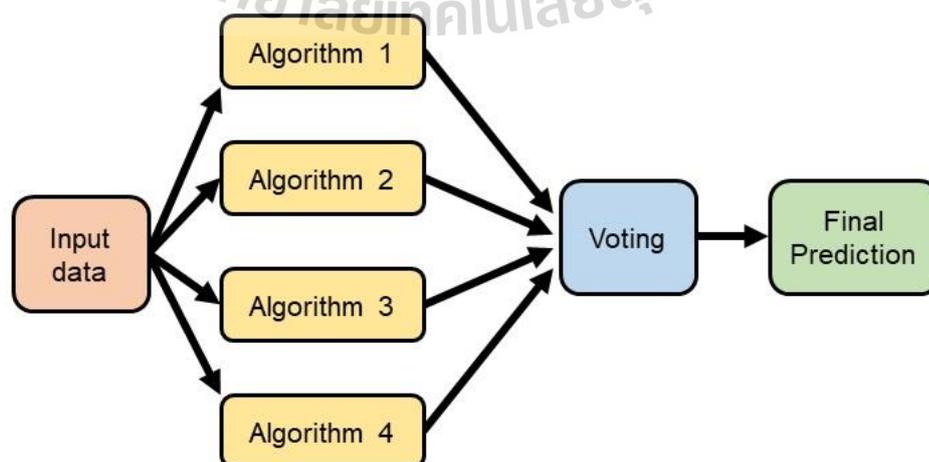
Figure 2.11 Ensemble methods diagram

2.6.2 Machine learning performance evaluation

Machine learning performance evaluation is a crucial aspect of developing and assessing the effectiveness of predictive models. It involves quantifying the accuracy and reliability of the model's predictions by comparing them to the actual values. Performance evaluation assists in studying model selection, improvement, and optimization. Various metrics are used for evaluation, such as R-squared (R^2), mean absolute error (MAE), Mean Squared Error (MSE), and mean squared error (RMSE) (Chicco et al., 2021).

R-squared (R^2): R^2 or the coefficient of determination, measures the proportion of the variance in the dependent variable that can be explained by the independent variables in a regression model. It ranges between 0 and 1, with a higher value indicating a better fit of the model to the data.

Mean Absolute Error (MAE): MAE measures the average absolute difference between the predicted and actual values. It provides a measure of the average magnitude of errors without considering their direction. Smaller MAE values indicate better prediction accuracy.

Mean Squared Error (MSE): MSE measures the average squared difference between the predicted and actual values. It squares the errors, penalizing larger errors more heavily. MSE is widely used as an objective function in regression models. Like MAE, smaller MSE values indicate better prediction accuracy.

Root Mean Squared Error (RMSE): RMSE is the square root of the MSE. It measures the standard deviation of the residuals or errors in a regression model. RMSE is often used to evaluate the accuracy of prediction models, with smaller values indicating better performance.

$$R^2 = 1 - \frac{\sum_i (x_i - y_i)^2}{\sum_i (x_i - \overline{y_i})^2} \qquad \text{(eq.12)}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i| \qquad \text{(eq.13)}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2 \qquad \text{(eq.14)}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \qquad \text{(eq.15)}$$

When x_i = The measured values

y_i = The predicted values

$\overline{y_i}$ = The mean values

2.6.3   Comparison of data mining tools

Data mining is an advanced data analysis technique that involves the process of discovering hidden patterns and relationships within large datasets. This process combines artificial intelligence (AI), machine learning (ML), and statistical analysis to identify data trends and make predictions based on those trends. With numerous tools available in the market, it is essential to compare their features, strengths, and limitations to determine the most suitable tool for specific data mining tasks. In the study conducted by Chou et al. (2018), four popular data mining tools were explored: RapidMiner Studio, Microsoft Azure Machine Learning Studio, WEKA, and IBM SPSS Modeler.

The data mining tool described by Wolff (2022) provides a detailed description of their data mining tool, functionalities, capabilities, and applications.

RapidMiner Studio: It is a free and open-source data science platform that is based on a Java application. It is designed to provide multiple tools for data analysis tasks and features hundreds of algorithms for data preparation, machine learning, deep learning, text mining, and predictive analytics.

Advantages: - Free and open-source platform.
                - User-friendly visual interface.
                - Extensive library of algorithms for various data analysis tasks.
                - Support for machine learning, deep learning, text mining, and predictive analytics.
                - Active community support.
Disadvantages: - Limited scalability for big data processing.
                - Requires some level of programming knowledge for advanced customization.

Microsoft Azure Machine Learning Studio: It is a cloud-based platform that allows users to build, deploy, and manage machine learning models. It features a drag-and-drop interface and offers built-in algorithms, and automated machine learning capabilities. The platform enables users to quickly create and deploy predictive models as analytics solutions.

Advantages: - Cloud-based platform with scalable infrastructure.
                - Integration with other Azure services.
                - Automated machine learning capabilities.
                - Collaboration and deployment features.
                - Seamless integration with Microsoft ecosystem.

Disadvantages: - Reliance on Azure services.

                        - Potential scalability and performance limitations.

                        - Limited customization options compared to other tools.

Weka: It is a free and open-source machine learning software with a large collection of machine learning algorithms that is based on a Java application. It widely used for educational purposes and provides a large collection of algorithms and techniques for data analysis.

Advantages: - Open-source tool with a large collection of algorithms.

                        - User-friendly graphical interface.

                        - Suitable for educational purposes.

                        - Extensive data preprocessing capabilities.

                        - Active community support.

Disadvantages: - Limited scalability for large datasets.

                        - Less suitable for big data analytics.

                        - Limited integration options with other tools or platforms.

IBM SPSS Modeler: It is a visual data science and machine learning solution designed for data mining and predictive analytics. It is a user-friendly data mining tool that supports various modeling techniques. It features a visual interface for ease of use and offers capabilities for data preparation, transformation, and integration with other SPSS products.

Advantages: - User-friendly visual interface.

                        - Broad range of modeling techniques.

                        - Support for data preparation and transformation.

                        - Integration with other SPSS products.

                        - Strong documentation and support.

Disadvantages: - Proprietary software with licensing costs.

                        - Limited customization options compared to some other tools.

                        - Less flexible for advanced users.

In the comparison of RapidMiner Studio, Microsoft Azure Machine Learning Studio, WEKA, and IBM SPSS Modeler. IBM SPSS Modeler was the most effective platform for the baseline analysis, outperforming other AI techniques and producing the best performance among the models evaluated (Chou et al., 2018).

IBM SPSS Modeler is a multipurpose software that suites on data mining and include numerous methods, It supports deep learning tasks, various data structures, and time series analysis (Bruxella et al., 2014).

This study utilizes IBM SPSS Modeler through a 30-day free trial subscription. The trial period includes access to trial support provided through the Stack Overflow forum, ensuring that users can seek assistance and guidance during their exploration of the software. With the comprehensive capabilities of SPSS Modeler, researchers can effectively prepare, blend, explore, and model their data without the need for programming expertise.

2.7   Literature review of noise prediction and noise mapping

Noise pollution is a significant environmental concern that requires thorough study and analysis. The noise propagation prediction can be achieved by applying mathematical formulas such as ISO 9613-2 and various modeling software tools like Nord2000, CONCAWE, IMMI Software, and SoundPLAN. These tools provide valuable insights into how noise travels and its potential impact on different areas. Furthermore, noise pollution evaluation can be effectively conducted by utilizing Geographic Information Systems (GIS) and developing noise maps with software applications such as ArcGIS, CadnaA®, SoundPLAN, and Openwind®. These software tools enable researchers and policymakers to visualize and analyze noise pollution patterns, assisting in identifying areas that require mitigation measures.

In conducting this study, a comprehensive literature review of noise mapping was undertaken, focusing on selecting relevant research that aligns with the objectives and scope of the present investigation. The literature review of noise prediction and mapping is shown in **Table 2.3**.

Table 2.3 Literature review of noise prediction and noise mapping

Title	Study purpose	Method	Finding	References
Literature review of noise prediction and noise mapping				
Analytical procedure for constructing noise contours	To present an analytical procedure for developing a noise contour map of the workplace.	Develops mathematical formulas to estimate combined noise levels at predetermined locations and generate a noise contour map.	1. The proposed procedure allows engineers to construct noise contour maps quickly. 2. Estimated noise levels provide the assessment of noise impact and determine areas where hearing protection devices are required.	Nanthavanij et al. (1999)
Road traffic noise: GIS tools for noise mapping and a case study for Skåne region	To address traffic noise pollution and develop a noise calculator software package for creating noise maps.	The noise calculator is based on the noise model described in the Nordic prediction method for road traffic noise and simulates a road traffic noise map using the ArcMap GIS package.	1. 5.65% of the population is affected by noise levels exceeding limit. 2. Nordic Prediction accuracy is based on predictions and mathematical models created by real noise levels in different situations.	Farcaş and Sivertunb (2010)
Traffic noise prediction with Nord2000-an update	To provide accurate predictions of road traffic noise levels over complex terrain and under various weather conditions.	Predict third-octave band levels of road traffic noise and calculate population exposure from yearly noise levels using Nord2000.	1. Nord2000 provides accurate predictions. The standard deviation was 1 dB up to 400 m, and above 400 m, it up to 2 dB for flat ground. 2. The effectiveness of noise barriers appears slightly lower when positioned downwind and perpendicular to the road.	Kragh (2011)

Table 2.3 (Continued).

Title	Study purpose	Method	Finding	References
Traffic noise contour mapping in Matara City - Sri Lanka	To address the increasing road traffic noise problem and prepare a noise map for noise control measures to prevent potential health issues for the population of Matara City in the Southern Province of Sri Lanka.	Measure noise levels using a handheld sound level meter and simulate a road traffic noise map using IMMI Software.	3. The noise levels exceed the maximum allowed level in more than half of Matara City, particularly in the suburban area. 3. Locations such as traffic lights and road crossings experience higher noise levels due to factors such as high-power engines, poor exhaust systems, and vehicle horns.	Sethunga et al. (2013)
Noise pollution analysis of wind turbines in rural areas	To analyze the wind turbine, the acoustical noise from frequency spectra and time history of the noise produced by the blade rotation and the wind turbine mechanical operations.	Simulate a wind farm noise map using CadnaA® software based on the annual mean wind speed dataset for each turbine.	1. Noise levels are around 55 dBA in average wind speed conditions and lower in the rest of the area. 3. Noise levels simulated in terms of high wind speed conditions, about 10-11m/s, it has been increased, but it seems to be compatible with daily human agricultural activity.	Ruggiero et al. (2015)

Table 2.3 (Continued).

Title	Study purpose	Method	Finding	References
The Evaluation of Noise Pollution at Samen District in Mashhad, Khorasan Razavi Province, Iran using Geographic Information System	To evaluate noise pollution in Samen district using GIS and comparison with existing standards.	Analyze and create a zoning map of the measured equivalent and maximum sound levels by developing a noise map using ArcGIS software.	1. The IDW interpolation method demonstrates higher accuracy compared to other interpolation methods. 2. Noise maps were analyzed using a one-sample t-test, The equivalent sound level during all times (morning, evening, and night) in summer and fall was found to be higher than the standard. 3. The L_{eq} was higher than the guideline values, indicating that noise pollution in the Samen district is a serious issue.	Moteallemi et al. (2017)
A GIS-based method for assessment and mapping of noise pollution in Ota metropolis, Nigeria	To assess and map noise pollution levels in Ota metropolis, Nigeria	Developed noise map based on the average equivalent noise (LAeq) values using ArcGIS 10.5 Software.	1. IDW provides satisfactory results when the number of elevation points in an area is large are uniformly Distributed. 2. The A-weighted sound level (LAeq), the background noise level (L_{10}), and the peak noise level (L_{90}) vary with location and period of the day due to traffic characteristics. 3. The results of this study are useful as a reference and guideline for future planning and regulations on noise limits to be implemented for urban areas like Ota Metropolis	Oyedepo et al. (2019)

Table 2.3 (Continued).

Title	Study purpose	Method	Finding	References
Noise predictions from elevated sources in industrial environments	To investigate the limitations and assumptions associated with the ISO 9613-2 prediction method for assessing noise impacts from industrial equipment installed in a free field and on top of floating screens.	Utilizes an idealized test case and a realistic industrial scenario that are modeled to examine the ISO 9613-2 ground effect calculation methods as implemented in SoundPLAN v7.3.	1. SoundPLAN v7.3 did not perform accurately, with corrections not exceeding 1 dB, when simulating short-range sound propagation over hard ground and at sharp sound angles for the considered source-receiver heights and distances. 2. ISO 9613-2 has limitations in scenarios with elevated and closely spaced sources and receivers.	Novkovic et al. (2017)
Computational sound propagation models: An analysis of the models Nord2000, CONCAWE, and ISO 9613-2 for sound propagation from a wind farm	To investigate this research by comparing the predicted to the measured sound pressure levels from a wind farm in northern Sweden.	Calculates and compares the results of sound level calculations using Nord2000, CONCAWE, and ISO 9613-2 with the measured sound levels.	1. CONCAWE and Nord2000 showed high accuracy for downwind conditions at 8 m/s. 2. Nord2000 is more accurate for upwind conditions at 8 m/s, as it better calculates the refraction. 3. ISO 9613-2 is not accurate for the specific site conditions of research. 4. The sound power level for wind speeds lower than 7 m/s differs depending on the wind speed and slightly sound measurements.	da Silva and Lorena (2017)

Table 2.3 (Continued).

Title	Study purpose	Method	Finding	References
Literature review of noise prediction (Machine learning)				
Vehicular traffic noise modeling using artificial neural network approach	To develop an accurate and relevant traffic noise prediction model for highways in India.	1. The ANN model trained with three key input parameters: total vehicle volume/hour, percentage of heavy vehicles, and average vehicle speed. 2. comparing the predicted noise descriptors with field measurements.	1. The ANN model accurately predicts the 10 Percentile exceeded sound from measurements. 2. The difference in results between the ANN approach and regression analysis is less than 5%.	Kumar et al. (2014)
Automated wind turbine noise analysis by machine learning	To determine the classification methods of sound sources using octave bands, the LASSO and RFE techniques will be employed.	1. measures sound over a long period to cover different weather conditions. 2. Train machine learning algorithms to classify the sound samples into different categories based on classification rules. 3. Optimize the features to improve speed and enhance the classification system's ability.	1. The classification algorithm able to classifying WTB noise, bird sounds, tractor noise, and explosion sound with 91% accuracy. 2. the processing speed can be improved by optimizing the features and training the classification system.	Välisuo (2017)

Table 2.3 (Continued).

Title	Study purpose	Method	Finding	References
Is it possible to predict background noise levels from measured meteorological data with machine learning techniques	1. To model the background noise levels using machine learning techniques. 2. To train the model for each dataset, and evaluate the accuracy of the regression algorithm.	1. Use a gradient boosting technique used in classification and regression based on the concept of ensemble. 2. Evaluate a performance by using the mean absolute error and the standard deviation error.	1. The results confirm the fact that the noise level depends on the measured location and meteorological conditions. 2. This technique will change the way we manage noise and meteorological data in acoustics and wind energy.	Bigot and Hochard (2019)
A machine learning approach for traffic-noise annoyance assessment	Develop models for predicting traffic-noise annoyance based on machine-learning techniques, including artificial ANN, SVM, MLR.	1. Apply machine-learning techniques (ANN, SVM, MLR) to obtain traffic-noise annoyance models. 2. Compare error rates and coefficient of determination (R^2) among the models. 3. Evaluate the accuracy and performance of traditional statistical models compared to machine-learning approaches.	1. The ANN model showed the best results, achieving significant error reduction compared to MLR and SVM models in both training and testing subsets. 2. R^2 increased significantly when using the ANN model compared to MLR and SVM models.	Bravo-Moncayo et al. (2019)

Table 2.3 (Continued).

Title	Study purpose	Method	Finding	References
Real-time machine learning for air quality and environmental noise detection	1. Improve indoor air quality and outdoor noise transmission by implementing real-time sensors and AI.	1. Implemented real-time detection and notification system with cloud and edge computing on website. 2. Applied machine learning techniques for air quality and noise level prediction and classification.	1. The real-time data on outdoor/indoor Air Quality Index (AQI) and noise levels, empowering building occupants to have control over Indoor Air Quality (IAQ). 2. The integration of AI-based real-time data with the existing building system facilitates the optimization of indoor environmental quality and supports occupant control.	Shah et al. (2020)
Noise prediction using machine learning with measurements analysis	1. Establish a machine learning model for noise prediction in workplaces with noise pollution. 2. Evaluate the impact of training data selection on the noise prediction model's performance.	1. Utilized the gradient boosting model (GBM) as the machine learning model for noise prediction. 2. Integrated past noise measurement records and various features into the proposed model for making predictions. 3. Investigated the effect of selecting different training data for model training.	1. The proposed method showed the ability to effectively predict noise levels. 2. Accurate prediction of future noise pollution is crucial for ensuring the health and well-being of laborers working in high-noise environments. 3. The proposed model offers a valuable tool for preventing people from working in harmful noise positions and maintaining employee health.	Wen and Huang (2020)

Table 2.3 (Continued).

Title	Study purpose	Method	Finding	References
Traffic Noise Modelling Using Land Use Regression Model Based on Machine Learning	1. Estimate the sound level during peak daily periods. 2. Determine the best approach to create prediction model based on ML, statistical regression, and GIS.	1. Utilized two computing methods: machine learning (decision tree, random forest algorithms) and statistical regression (linear regression, SVR). 2. Evaluated and compared the algorithms based on various performance metrics.	1. ML models, specifically random forest, outperformed the statistical regression-based models in predicting sound level for traffic noise modeling. 2. The combination of machine learning, LUR modeling, and GIS data provided superior performance in estimating sound pressure levels.	Adulaimi et al. (2021)
A Machine Learning for Environmental Noise Monitoring and Classification Using Matlab	1. To utilize ML for the classification of environmental sounds using MATLAB and compare the 2. Identifie noise pollution with international standards.	1. Collect environmental sounds using sound capture tools. 2. Utilize ML classification models for sound recognition and classifcation using MATLAB to ensure accurate analysis.	1. ML classification models from MATLAB are expected to provide accurate results regarding the identified noise types. 2. The findings will contribute to understanding and managing noise pollution in cities based on international standards and recommendations.	Albaji et al. (2021)

Table 2.3 (Continued).

Title	Study purpose	Method	Finding	References
Machine learning-based tools for wind turbine acoustic monitoring	Develop a tool capable of identifying the operating conditions of an acoustic source, specifically focusing on a WT near a sensitive receptor.	1. measurements of the noise in one-third octave bands to capture the frequency characteristics of the noise emitted by a wind turbine. 2. Developed a model based on a SVM for detecting the operating conditions of the WT	1. SVM model and ANN model, exhibited high precision in identifying the operating conditions of the wind turbine. 2. These models demonstrated their potential as effective tools for supporting the acoustic characterization of noise in environments near wind turbines.	Ciaburro et al. (2021)
Application of Machine Learning to Include Honking Effect in Vehicular Traffic Noise Prediction	1. Develop a road traffic noise prediction methodology using machine learning. 2. Assess the impact of traffic noise and improve predictions compared to standard models.	1. Apply decision trees, random forests, generalized linear models, and artificial neural networks for traffic noise prediction. 2. Compare the results using metrics like mean square error, correlation coefficient, coefficient of determination, and accuracy.	1. ML models that include the effects of honking demonstrate improved predictions of road traffic noise compared to standard models that neglect honking. 2. The honking is an essential parameter in future for traffic noise prediction models. 3. The study highlights the importance of using machine learning techniques and considering honking occurrences for more accurate modeling.	Singh et al. (2021)

Chapter III

METHOD

This study focuses on the propagation of sound level from a wind turbine to nearby community and develops a sound level prediction model. To test the hypothesis presented before sound level, wind speed, temperature, and moisture data was collected from study area at various distances from the wind turbines. The method of the study can divide into

1. Noise map generation: In this step, the study site had investigated for site description. The field measurement had measured sound level levels follows a guidance note on sound level assessment of wind turbine operations at EPA licensed sites (NG3) by USEPA. The secondary data such as base map picture and transportation route, had gather from ESRI's community and Thailand land development department. The field measurement and the secondary data had made a noise map generation. The noise map had generated by ArcGIS's interpolation and overlay analysis.

2. Statistical comparison: In this step, After the field measurement had measured, the sound levels had been compared relationship with distance, time, and wind speed. The T-test method used to determine a significant difference between sound level in day-time and night-time. The Pearson correlation used to measure of linear correlation between sound level and distance, time, and wind speed that it represents a relationship of two variables.

3. Developing a model to predict sound level levels cause by a wind turbine: In this step, IBM SPSS Modeler is a data mining and text analytics software application. IBM SPSS Modeler had selected from modeling program comparison. Other field measurement had measured sound level levels, wind speed, wind direction, temperature, and moisture follow a guidance note on sound level assessment of wind turbine operations at EPA licensed sites (NG3) by USEPA. field measurement data had input to IBM SPSS Modeler. Data had separated to training and testing partition. The prediction model had generated by the auto numeric node. The Auto Numeric node estimates and compares models for continuous numeric range outcomes using a number of different methods such as regression, generalized linear, SVM, C&R tree, CHAID, KNN algorithm. The model used to predict a sound level that nearby

community received at a worst case scenario. The worst case scenario had made from a Nakhon Ratchasima Climatological data for period 1990-2019.
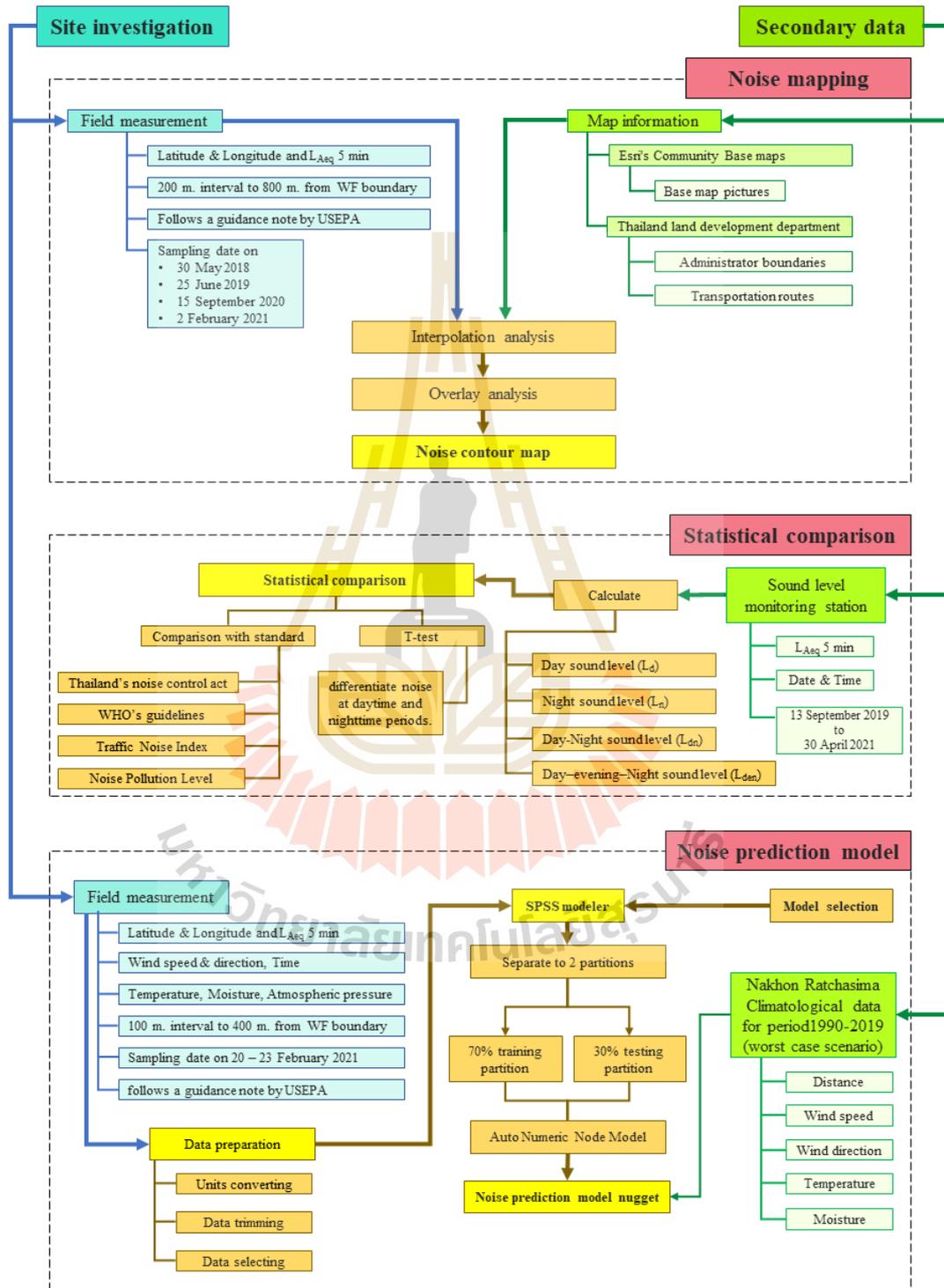
Figure 3.1 Conceptual framework

3.1   Noise mapping

3.1.1   Site description

The study area is in Nakhon Ratchasima province, Thailand, which has a tropical climate influenced by seasonal monsoon winds. The case study wind farm is located on a mountain ridge with 20-year wind speeds ranging from 0.87 to 1.29 m/s. The wind farm consists of 30 wind turbines with a hub height of 125 m and a rated power of 2.0 MW on an area of 3.25 sq. km. The study area is 800 m around the boundary of the wind farm and covers an area of 13 sq. km. The location of the wind farm, its layout, automated sound monitoring stations, and noise measurement points are shown in Figure 3.2.

The study area has three land use types: (1) residential; (2) industrial (a wind farm); and (3) agricultural. Figure 3.3 shows the satellite image of the study area and land use zones with the color code classification following the Department of Public Works and Town & Country Planning, Thailand. The area is dominated by dryland agriculture land use, such as cassava, cane, corn, etc., followed by industry and community. There are two rural communities close to the wind farm: Huai Bong village, located approximately 600 m to the southeast, has 326 households, and Noi Phatthana Village, located approximately 500 m to the northeast, has 200 households. Most households are cultivators. The study area consists of two major routes: (1) a highway, which is a two-lane road running east-west and located on the wind farm's south side; and (2) a rural road, which is a two-lane road running north-south and located on the wind farm's east side.

3.1.2   Study area

The study area was a wind farm boundary and its vicinity. To select the measuring point, the study area was divided into 8 directions with 800 m distance from the wind farm boundary. Locations was set at 5 locations for each at distances ranging every 200 m interval up to 800 m from the wind farm, totally 40 points. The surrounding area was rural, and agriculture was cultivated, including cassava, cane, and corn. The area was flat, with little difference in elevation. The majority of the area was unaffected by terrain features like hills, trees, and buildings that could activate sound propagation. The study area is shown in Figure 3.4.
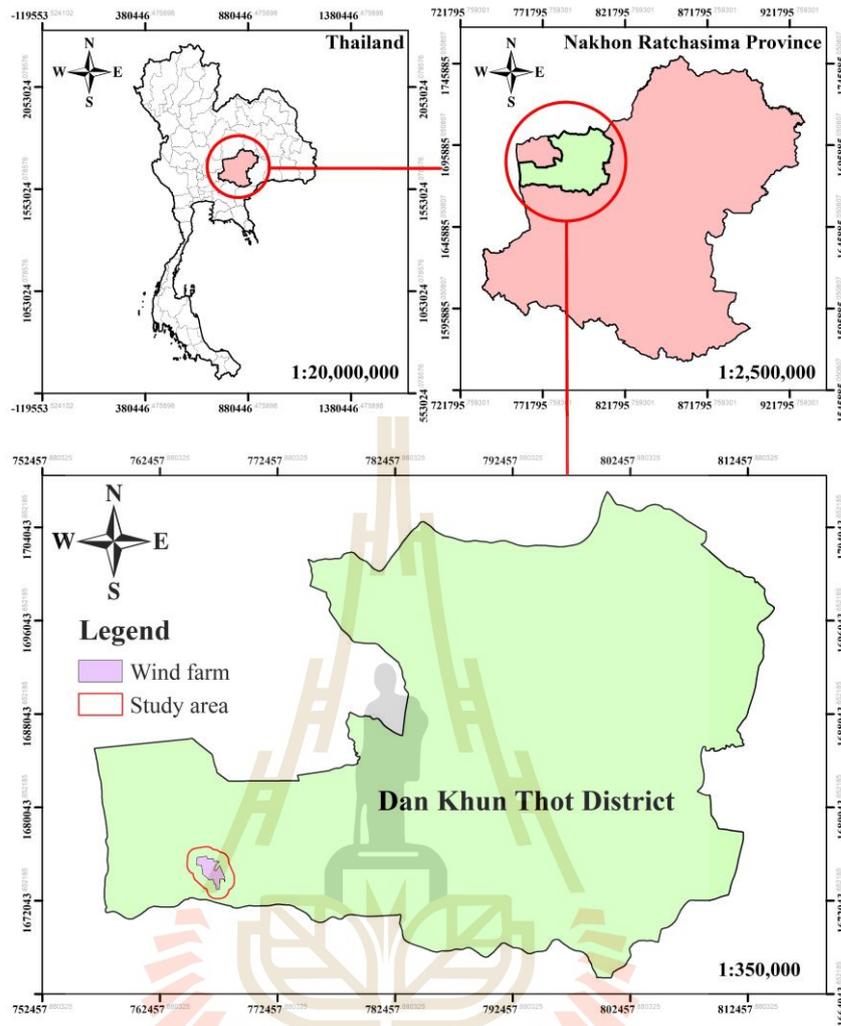
Figure 3.2 Study area
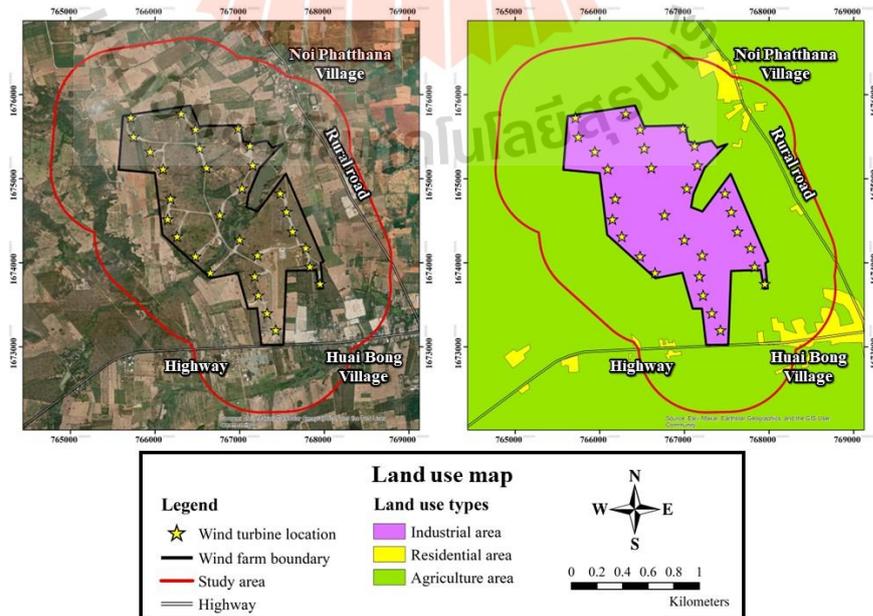
Figure 3.3 Satellite image and land use of the study area
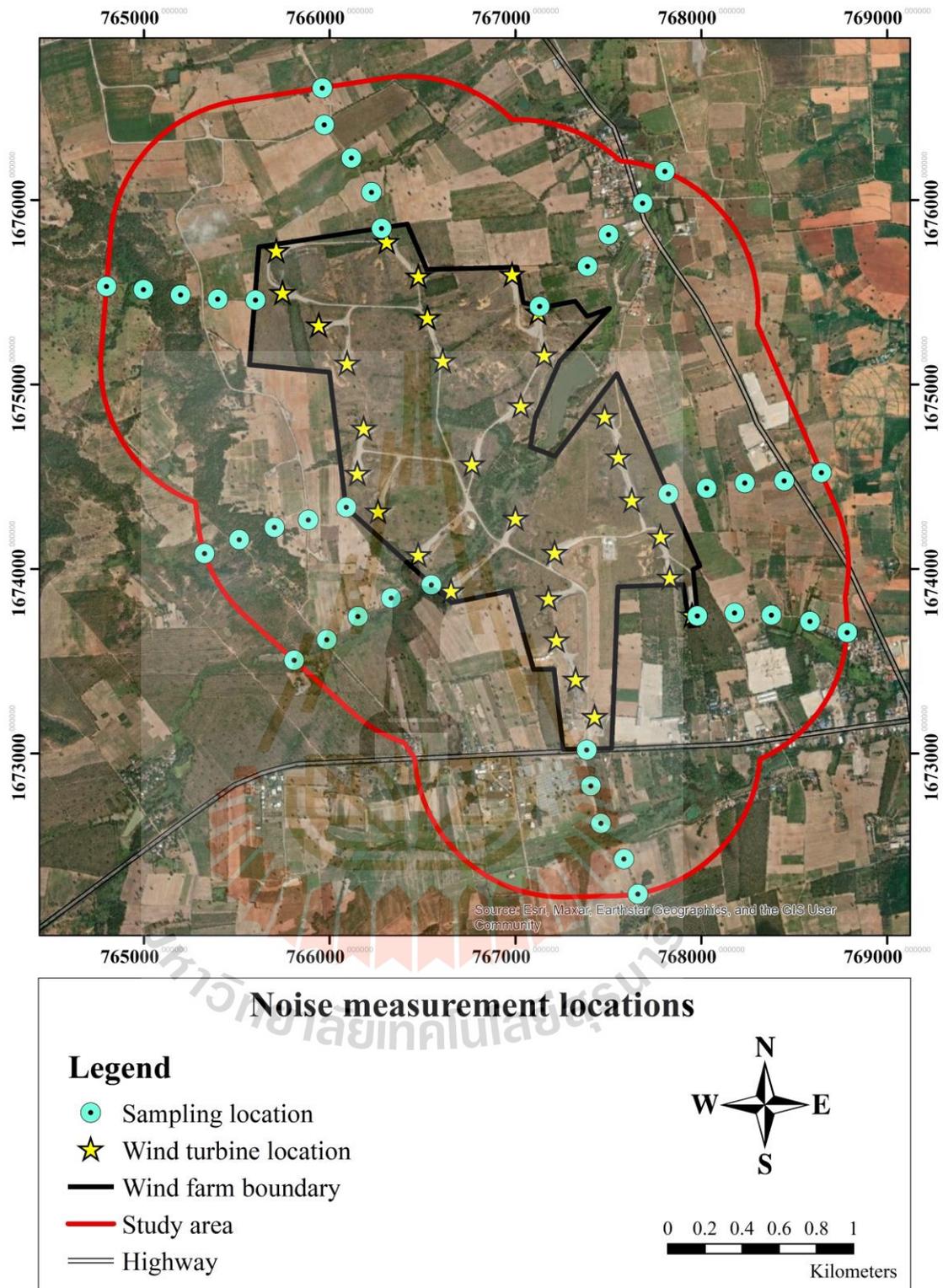
Figure 3.4 Study area and noise measurement locations for noise mapping

3.1.3 Method and equipment

The sound levels were measured for 10 minutes per location as an average L_{Aeq} (dB(A)) as described in (equation 11). The sampling time was around 13:00-

16:00 pm on May 30, 2018, June 25, 2019, September 15, 2020, and February 2, 2021. A noise measurement method follows guidance note on noise assessment of wind turbine operations at EPA licensed sites (NG3) by USEPA (McAleer & McKenzie, 2011). The sound level meter was set at 1.2-1.5 m above ground level on a tripod and positioned at least 3.5 m away from a reflecting surface to minimizing the impact of noise reflections such as a wall, building, or trees.
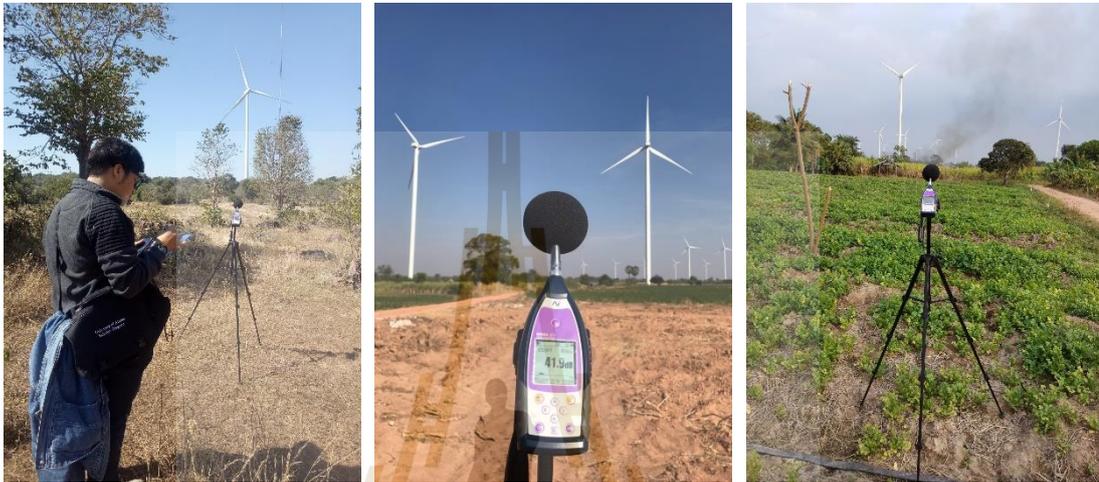
Figure 3.5 Noise measurement

The A-weighted continuous equivalent sound level (L_{Aeq}) is the logarithmic or energy-averaged noise level which is computed from the instantaneous noise levels. L_{Aeq} can be determined using equation 16.

$$L_{Aeq} = 10\log_{10}\left(\frac{1}{N}\sum_{i=1}^{N}10^{L_{A_i}/10}\right) \tag{eq.16}$$

When
N = the total number of readings

L_{A_i} = the ith A-weighted sound pressure level reading

$\overline{y}_i$ = The mean values

3.1.4 Sound level meter and global positioning system

The sound level was measured using the class 1 sound level meter BSWA 308 with the MPA231 microphone set from BSWA Technology Co., Ltd.- Productions. The sound level meter was set with a frequency weighting of "A" according to the international standard IEC 61672:2003 to represent human hearing. The global positioning system (GPS), as latitude and longitude, was measured with eTrex-10 from Garmin Ltd. The sound level meter and global positioning system meter are shown in Figure 3.6.

a)	b)

Figure 3.6 Measurement equipment

(a) Sound level meter and (b) Global positioning system meter

3.1.5 Noise map generation

The noise map was generated through an interpolation analysis using sound level sampling location coordinates, sound level distribution, and a base map (including country, city boundaries, or satellite imagery) in GIS. The sound level sampling locations were determined using GPS and represented as X, Y coordinates. The sound level distribution was analyzed using interpolation techniques, incorporating the sound level sampling location coordinates and field measurements. The base map was obtained from the Thailand Land Development Department. ArcGIS Desktop 10.5 software was utilized for creating the noise map, employing interpolation methods. The resulting interpolated surfaces can be visualized in ArcGIS as continuous color maps or contour lines, allowing for the identification of spatial patterns and trends in the noise levels across the study area.

The satellite imagery base map used in this study employed the WGS_1984_UTM_ZONE_48N coordinate system projection. Microsoft Excel software was utilized for the analysis and interpretation of tabular data, which included latitude, longitude, and sound level measurements from the field.

In this study, the kriging method was chosen for interpolation due to the anisotropic distribution of the sampling points, which is better suited for kriging. Additionally, kriging considers the spatial variability and offers variogram models that can improve interpolation accuracy. The noise maps were created by employing kriging interpolation within the Geostatistical analyst extension in ArcGIS. The Geostatistical analyst tool was utilized to generate statistics for comparative analysis and produce geospatial visualizations. The flowchart of interpolation process is shown in Figure 3.7.
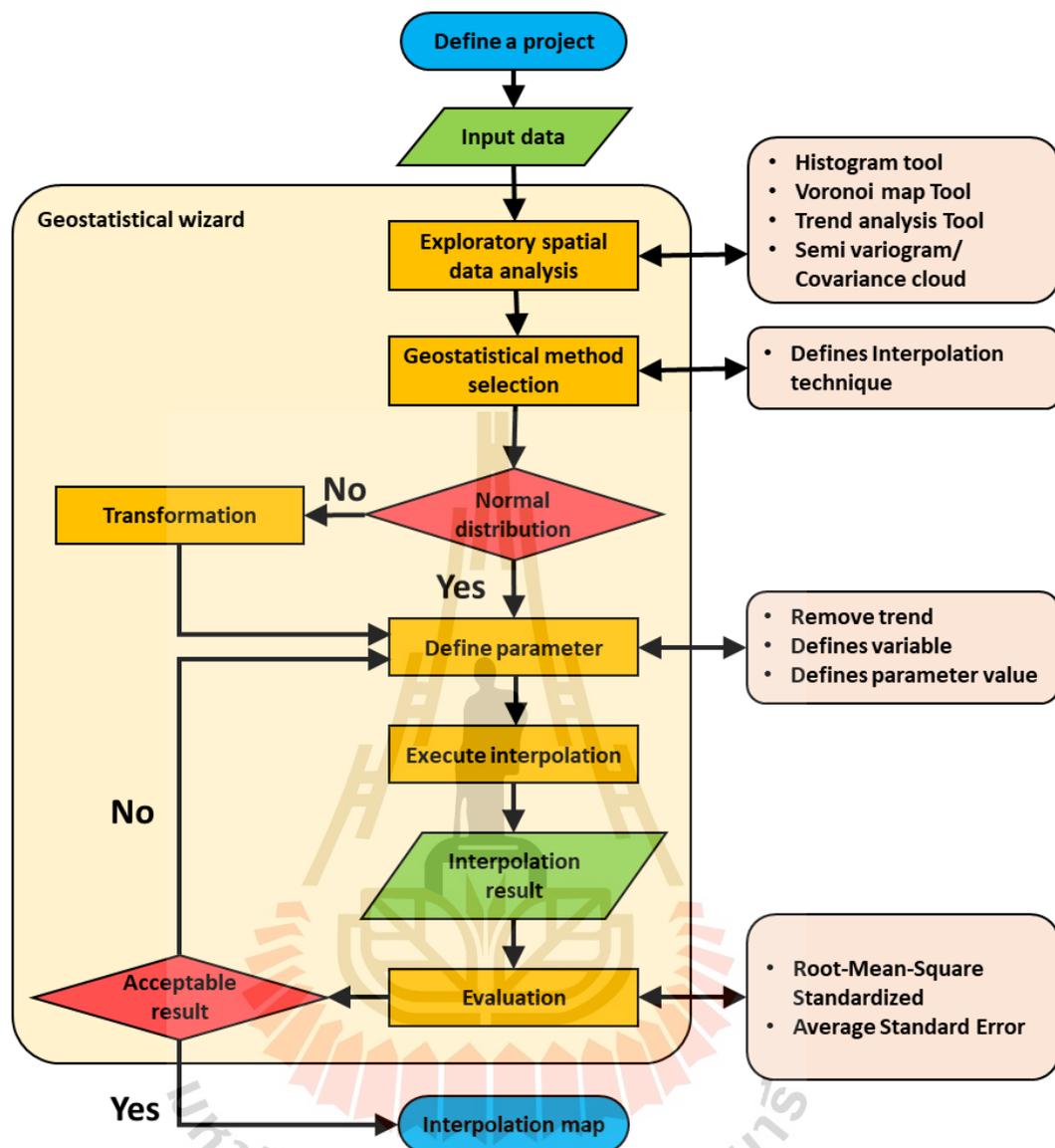
Figure 3.7 Noise map generation flow chart

The method involved several steps including:

In the first step, the noise datasets were imported into ArcGIS.

In the second step, exploratory spatial data analysis was conducted to examine the data and identify various statistics, including distribution, trends, directional components, and outliers. This analysis involved the use of different techniques, such as: Histogram analysis was performed to identify outliers and calculate the data distribution. Voronoi maps were utilized to analyze the spatial variability of neighborhood data. Semi-variogram/covariance cloud analysis was employed to assess spatial autocorrelation within the dataset and identify outliers. Trend analysis was conducted to identify global trends.

In the third step, the geostatistical methods were selected. The interpolation method defining as the kriging, setting the kriging type as simple and output surface type as prediction. The data was transformed to normal distributions. The second-order trend was removed. The variable was defined semivariogram variable, and a model type was chosen from options such as Circular, Spherical, Tetraspherical, Pentaspherical, Exponential, Gaussian, Rational Quadratic, Hole Effect, K-Bessel, J-Bessel, and Stable. The anisotropy was set to true, and the lag size and number of lags were determined. Through this process, the various combinations of parameters were executed.

In the fourth step, the interpolation was executed, and the results were assessed using cross-validation within the dataset.

In the fifth step, the statistical data was evaluated based on the root-mean-square standardized value, aiming for a value close to 1, and the average standard error, aiming for a minimum value (or close to 0). The values that yielded the most reliable and representative noise map were selected as the final settings for the combinations of parameter. If the evaluation indicated acceptable results, the noise map was considered generated. However, if the evaluation was not acceptable, the process required revisiting and redefining the combinations of parameters to achieve improved outcomes.

3.2 Noise Assessment

3.2.1 Automated sound monitoring stations

The wind farm operator installed two automated sound monitoring stations in a northeastern community and a southeastern community to continuously investigate the impact of noise from the wind farm. The monitoring stations were 500 m from the wind farm boundary (Figure 3.8). The instrument was the EM2030 Sound Level Monitor from Sonitus Systems Limited. The sound levels are measured automatically, analyzed, and uploaded with reports through the Sonitus Cloud platform. The monitoring stations measured a noise level every 5 minutes on average, L_{Aeq} (dB(A)). The data used in this study was measured from September 13th, 2019 to April 30th, 2021, covering a period of 596 days.
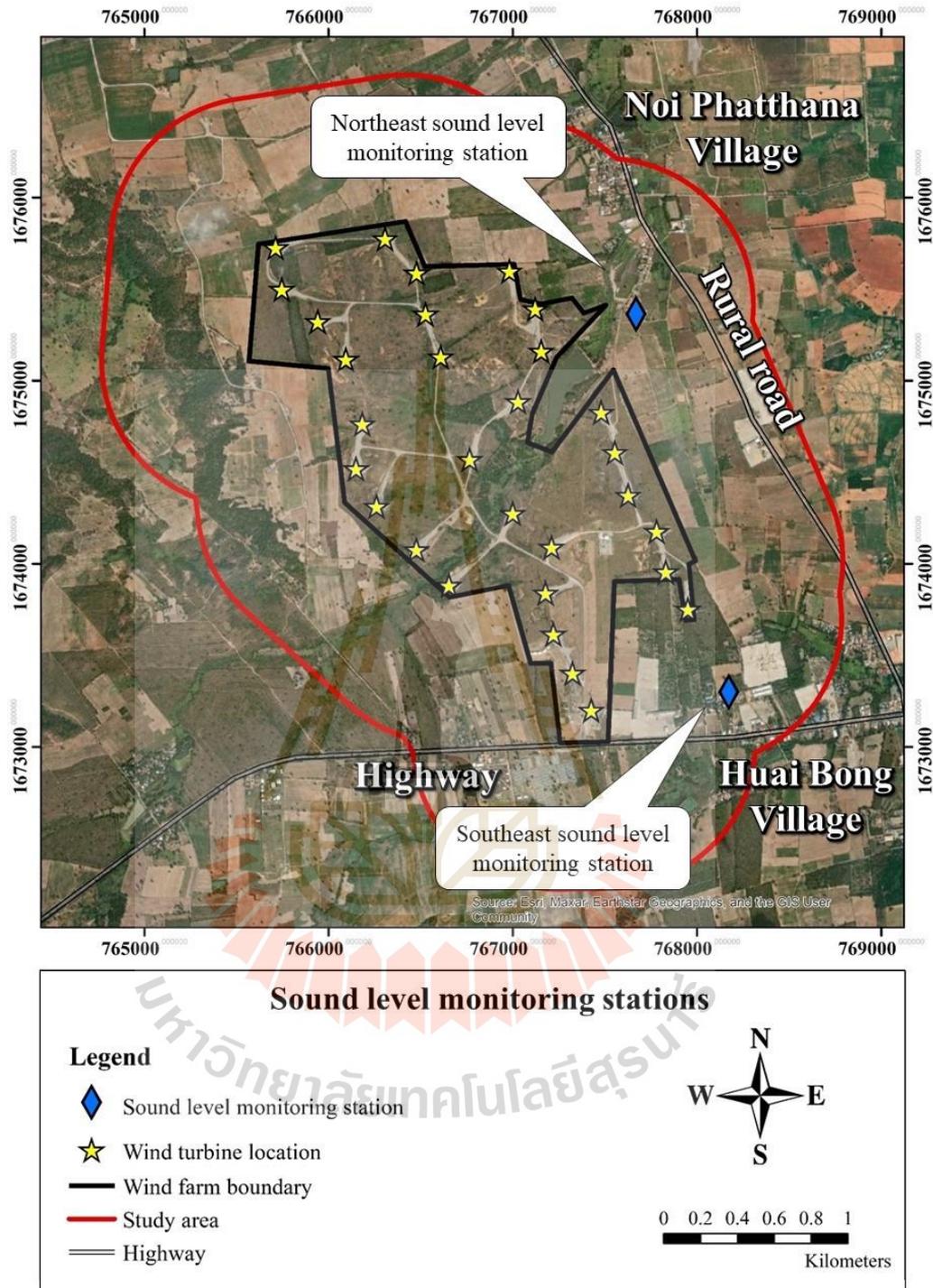
Figure 3.8 Automated sound monitoring station location

3.2.2 Statistical Analysis

Independent sample t-test analysis was performed to differentiate noise at daytime and nighttime periods.

The null hypothesis is that the means daytime and nighttime sound levels are equal.

The alternative hypothesis is that means of daytime and nighttime sound levels are unequal.

H_0: $\mu_1 = \mu_2$ (The means of daytime and nighttime sound levels are equal)

H_A: $\mu_1 \neq \mu_2$ (The means of daytime and nighttime sound levels are not equal)

In this work, the independent sample t-test method compared the difference with a 95% confidential interval.

3.2.3   Noise measurement metrics

Sound level descriptors are commonly used to measured how sound is heard, to determine the impact of noise on health and evaluate noise pollution, sound quality, and the potential for hearing damage. These descriptors are summarized by The U.S. The Environmental Protection Agency and the World Health Organization (United States. Office of Noise Abatement, 1974)

Day-Night sound level (L_{dn}): L_{dn} is the A-weighted equivalent sound level for a 24 hour period with an additional 10 dB weighted on the equivalent sound levels for nighttime to compensate for sleep interference and other disruptions, with separate weightings applied to:

-   Daytime that occurred within the 15 hour period of 7:00 a.m. and 10:00 p.m.
-   Nighttime that occurred within the 9 hour period of 10:00 p.m. and 7:00 a.m.

Day–evening–night sound level (L_{den}): L_{dn} is an average sound pressure level over a 24 hour period, evenings and nights in a year. For the daytime period, no additional weighting is applied, while a 5 dB penalty is added to the evening period and a 10 dB penalty is added to the nighttime period. The penalty reflects the increased noise sensitivity of people during these periods, with separate weightings applied to:

-   Daytime that occurred within the 12 hour period of 7:00 a.m. and 7:00 p.m.
-   Evening that occurred within the 3 hour period of 7:00 p.m. and 10:00 p.m.
-   Nighttime that occurred within the 9 hour period of 10:00 p.m. and 7:00 a.m.

L_{dn} and L_{den} can be determined using equation 17 and 18, respectively.

$$L_{dn} \quad = \quad 10\log\left(\frac{1}{24}\left(\left(15\times10^{\frac{L_d}{10}}\right)+\left(9\times10^{\frac{L_n+10}{10}}\right)\right)\right) \qquad \text{(eq.17)}$$

$$L_{den} \qquad 10\log\left(\frac{1}{24}\left(\left(12\times10^{\frac{L_d}{10}}\right)+\left(3\times10^{\frac{L_e+5}{10}}\right)+\left(9\times10^{\frac{L_n+10}{10}}\right)\right)\right) \qquad \text{(eq.18)}$$

When L_d =   Daytime equivalent sound level, dB(A)

L_e =   Evening equivalent sound level, dB(A)

L_n =   Nighttime equivalent sound level, dB(A)

Traffic noise index (TNI): TNI indicates the degree of variation in traffic noise levels and their effects on human annoyance (Langdon & Scholes, 1968). It shows the overall noise fluctuations over time by combining very noisy vehicles (L_{10}) and the general traffic noise (L_{90}) as described in equation 19.

Noise pollution level (L_{np}): L_{np} indicates the varying levels of noise that can cause physiological and psychological disturbances. L_{np} can be determined using equation 20.

$$TNI = 4 \times (L_{10} - L_{90}) + (L_{90} - 30) \qquad \text{(eq.19)}$$

$$L_{np} \quad \frac{L_{50} + (L_{10} - L_{90}) + (L_{10} - L_{90})^2}{60} \qquad \text{(eq.20)}$$

When L_{10} =   the sound level exceeded 10% of the time of the measurement period

         L_{50} =   the sound level exceeded 50% of the time of the measurement period

         L_{90} =   the sound level exceeded 90% of the time of the measurement period

3.3   Sound level prediction model

3.3.1   Study area

The study area was a vicinity wind farm located between wind farm and Noi Phatthana village. The measurement locations were at the northeast corner of the wind farm. The sound level measurement was performed at four points with distances of 100 m intervals up to 400 m. Additionally, meteorological ambient conditions were measured between these points, at a distance of 250 m. as shown in Figure 3.9.

3.3.2   Method and equipment

The sound level was measured with a calibrated PULSAR Model 44 S/N 1864 Sound Level Meter. The sound level meter was set with a frequency weighting of "A" according to the international standard IEC 61672:2003 to represent human hearing (International Electrotechnical Commission, 2013). Meteorology ambient conditions, including wind speed, direction, temperature, humidity, and atmospheric pressure, were measured with the NovaLynx Anemometer. The geographical positions of the measurement points were determined using a Garmin eTrex 10 handheld GPS. ESRI's ArcGIS 10.1 software was used to create the maps.

3.3.3 Data collection and data preparation

Field measurement data was taken in 5-min intervals for three days. The sampling date was chosen to cover the time when winter transitions into the dry season. This timeframe was selected due to the higher frequency of wind during this season compared to other seasons. The data collection period ranged from 1.00 pm on 20 February 20th, 2023, to 1.00 pm on February 23rd, 2023, a total of 864 data points per measurement location. In total, there are 3,456 datasets collected for the three-day period. measurement was taken at the minimum measurement frequency recommended by USEPA, fifty times per 10 minutes, to ensure sufficient data for modeling (McAleer & McKenzie, 2011). Measurement data was processed into a consistent and usable form. Data processing included data cleaning, data structuring, data transformation, and data filtering.
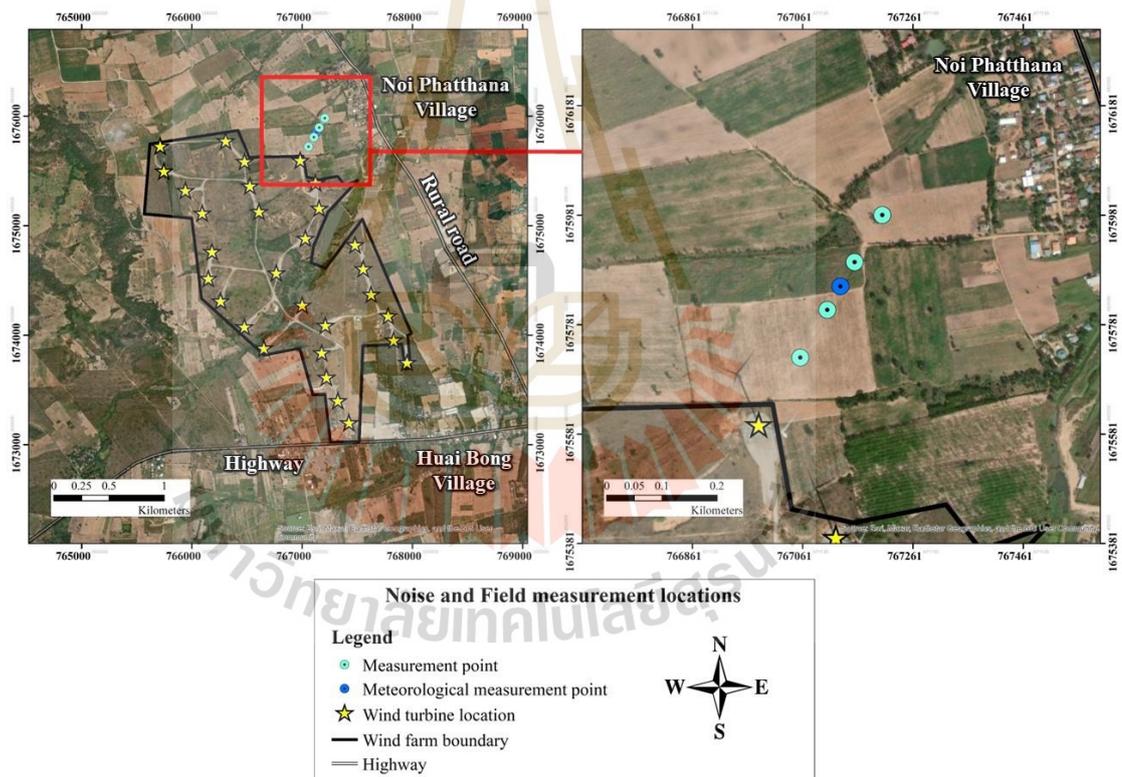
Figure 3.9 Study area and field measurement locations

3.3.4 SPSS Modeler

SPSS Modeler is data mining and analytics software used to build a predictive model. This research applied various algorithms to predict sound levels using field measurement data, including sound levels, wind speed, wind direction, temperature, humidity, and atmospheric pressure. The field measurement data was divided into two datasets, with a ratio of 70% for training and 30% for testing.

The auto-numerical node was used to generate a variety of algorithms in a single modeling run. The node explores every possible model and ranks each candidate model based on the correlation between predicted and observed values for each model. CHAID, CART, Linear, and Neural network models were possible to automatically create, and compare default models of continuous numerical outcomes from the auto-numerical node. Default values were set in the auto-numerical node. Four models were individual constructs that were then applied to construct ensemble models that were proposed for increasing accuracy. The brief descriptions of the prediction models used here are as follows:

CHAID (Chi-squared Automatic Interaction Detection) is a decision tree algorithm that builds a decision tree by recursively splitting the data into subsets based on the most significant differences between the target variable and predictor variables. CHAID is a popular algorithm for categorical target variables. It is used to identify the most important predictors that determine the target variable.

CART (Classification and Regression Trees) is another decision tree algorithm that builds a decision tree by recursively splitting the data into subsets based on the predictor variables that best predict the target variable. CART is used for categorical and continuous target variables. It can also be used for classification and regression tasks.

Linear regression is a statistical method for modeling the relationship between a dependent variable and one or more independent variables. In SPSS Modeler, linear regression models can be used for simple and multiple linear regressions that depend on the number of independent variables. The dependent variable is continuous, and the independent variables can be either continuous or categorical.

Neural networks are a type of machine learning algorithm that is designed to recognize patterns in data. In SPSS Modeler, neural network models can be used for classification and regression tasks. The neural networks are particularly useful when the relationships between the predictor variables and target variables are complex and non-linear. The neural network model in SPSS Modeler allows for the customization of the number of hidden layers and neurons in each layer, as well as the activation function used in the model.

An ensemble model is a machine learning technique that combines multiple individual models to improve the overall performance of the prediction. The idea behind ensemble models is that by combining multiple models, the strengths of each model can be leveraged, and the weaknesses can be mitigated.
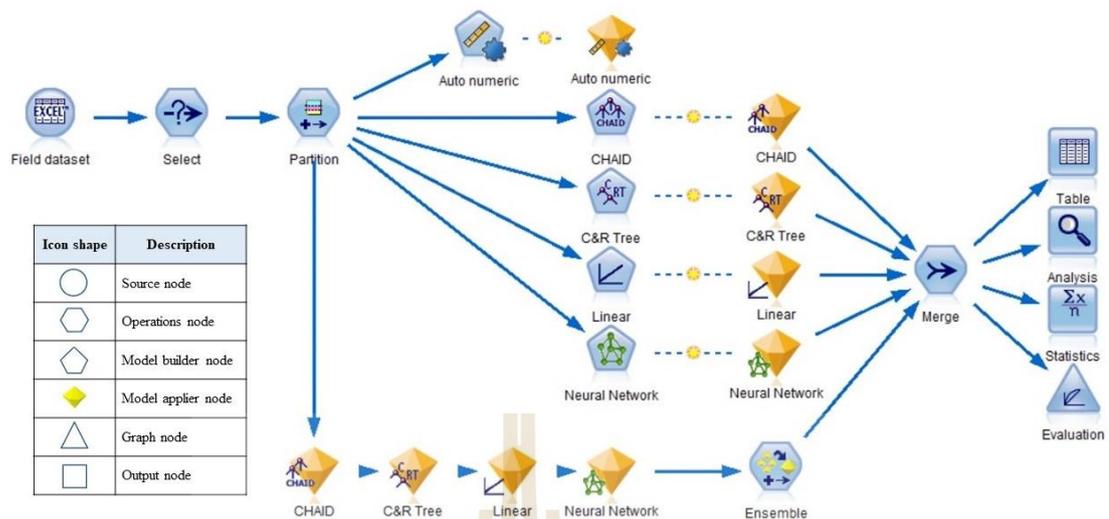
Figure 3.10 SPSS modeler flow chart

The modeling steps can be graphically illustrated as SPSS modeler flow, as shown in Figure 3.10. Nodes in the IBM SPSS Modeler are represented by a specific shape to indicate their function (The International Business Machines Corporation, 2021c). The source node (circle) imports data into the modeler from a different format. The operations node (hexagon) modifies the data in some way and returns the modified data to the modeler stream. The model builder node (pentagon) generates models from the data in the modeler. The model applier node (gold diamond) defines a container for the generated model that is returned to the modeler canvas. The graph node (triangle) generates a graph or report from the data in the modeler. The output node (rectangle) provides the means to obtain information about data and models. These node shapes work together to facilitate data processing and analysis in the IBM SPSS Modeler.

3.3.5 Model performance evaluation

To evaluate the prediction accuracy of the individual models and ensemble models, the predictor importance charts were produced to find the relative importance of each predictor in estimating the model. The most appropriate model was selected from 5 types of models by comparing the model's performance. The results of five models were merged, The performance error of the developed model was evaluated using R-squared (R^2), Root Mean Squared Error (RMSE), and the Mean Absolute Error (MAE), which expresses the average model-prediction error in the units of the variable of interest (Chicco et al., 2021). The smallest error model was selected as a prediction model (Ralević et al., 2014). The expressions of these parameters are given in equation 21, 22, and 23.

$$R^2 \quad = \quad 1 - \frac{\sum_i \left(x_i - y_i\right)^2}{\sum_i \left(x_i - \overline{y}_i\right)^2} \tag{eq.21}$$

$$MAE \quad = \quad \frac{1}{n}\sum_{i=1}^{n}\left|x_i - y_i\right| \tag{eq.22}$$

$$RMSE \quad = \quad \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - y_i\right)^2} \tag{eq.23}$$

When	x_i	=	The measured values
	y_i	=	The predicted values
	$\overline{y}_i$	=	The mean values

The gain chart is a visual representation of the performance of a predictive model. Gains are defined as the proportion of hits in each increment relative to the total number of hits in the tree (The International Business Machines Corporation, 2021a). Finally, the gain charts were plotted to evaluate the performance of the model.

3.3.6 Maximum sound level prediction for worst-case scenario

The prediction model has been selected based on evaluations for estimating the maximal sound levels generated in worst-case scenarios. In the study area, a wind turbine has already been constructed. The sound level produced by the turbine depends on various meteorological factors. Worst-case scenarios refer to environmental conditions that have the potential to cause the wind turbine to generate high sound levels that can propagate over long distances.

The data collection process involved selecting the highest or lowest values of various factors that affect the sound level and distance of propagation. These factors were obtained from meteorological data collected over 30 years in Nakhon Ratchasima province by the Meteorological Department of Thailand.

Table 3.1 Meteorological data in worst-case scenario

Parameters	Value	Descriptions
Wind speed, m/s	23.15	The highest average wind speed
Wind direction	45 (Northeast direction)	The downwind propagation from source to receiver is related to "worst-case"
Humidity, %	93	The speed of sound in air increases with the increase in humidity.
Temperature, °c	43.2	The speed of sound in air increases with the increase in temperature.
Pressure, hPa	1028	The speed of sound in air increases with the increase in air pressure.

Chapter IV
RESULTS

4.1   Site description

The study area covers the village of Huai Bong in Dankhuntod District, northwestern part of Nakhon Ratchasima Province. It is approximately 2.5 kilometer south of the Huai Bong Sub-District. The wind farm is located at the center of the study area. The wind farm has 30 wind turbines scattered around the site. The wind farm area is 3.25 square kilometers. This area is mainly covered by agriculture, such as cassava, cane, corn, and mixed deciduous forest. Two rural communities are nearby: Huai Bong Village locate approximately 1 kilometer to the southeast, and Noi Phatthana Village locate approximately 0.7 kilometer to the northeast. The agricultural area is mostly located around a wind farm and the two villages. There almost used for farming cassava, cane, and corn. Most of the area is flat land. Most of the area is not taking effect of terrain features such as hills, trees, and buildings that can affect sound propagation. There is a low bluff territory running from northwest to southwest. This area is covered with mixed deciduous forests. The tree can act as sound propagation, but there is no disadvantage to the community since the community is on the other side.

4.2   Noise map

4.2.1   Field measurement data

The averages and standard deviation of the sound levels at distances ranging every 200 m intervals up to 800 m from the wind farm at May 30, 2018, June 25, 2019, September 15, 2020, and February 2, 2021, as shown in Table 4.1.

Table 4.1 Field measurement data

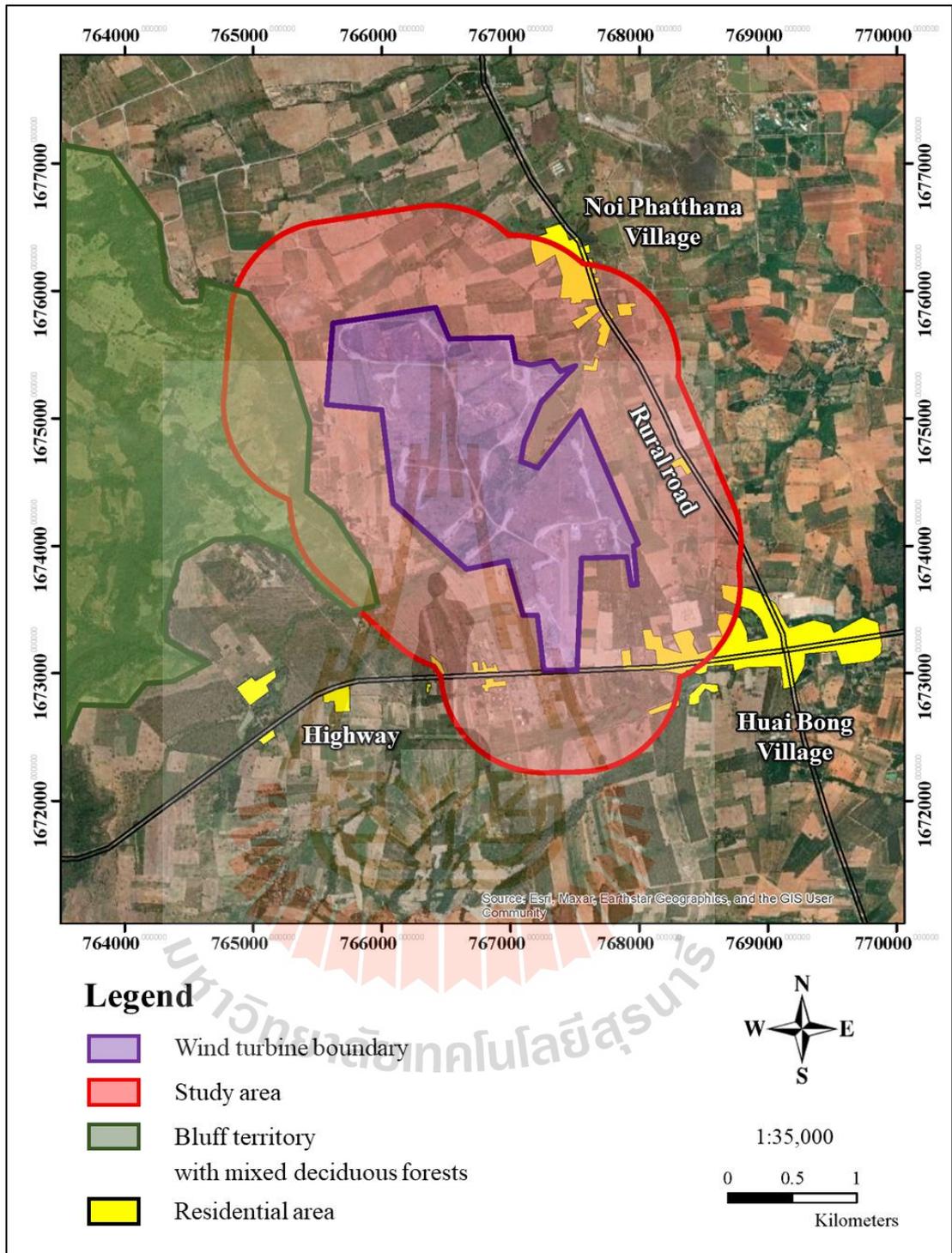| Parameters	Sound level, dB(A)				
	0 m	200 m	400 m	600 m	800 m
May 30, 2018	49.0±4.9	46.9±6.7	46.3±3.4	51.5±4.1	47.1±4.4
June 25, 2019	52.3±5.8	49.8±2.7	49.3±1.8	51.0±6.8	52.1±5.5
Sep 15, 2020	51.1±3.5	44.4±3.4	45.7±4.8	46.7±9.8	45.5±6.1
Feb 2, 2021	49.2±1.9	45.8±3.6	46.0±4.3	46.9±4.9	46.1±2.5

Figure 4.1 Site description

4.2.2 Evaluation of predictions

The exploratory spatial data was used to examine the frequency distribution of the data, checking for its normality, identify outliers, and explore spatial patterns, as shown in Figure 4.2.

Figure 4.3 includes the slope of the best fit line in the scatter plot, the remaining number of data points after removing outliers using cluster type in the voronoi map, the type of variogram models employed, and the statistics of the noise data collections.

On May 30, 2018, after removing outlier samples, 26 out of 40 samples remained. The analysis took circular methods. The resulting statistics were a root-mean-square standardized value of 1.588 and an average standard error of 2.515.

On June 25, 2019, after removing outlier samples, 24 out of 40 samples remained. The analysis took gaussian methods. The resulting statistics were a root-mean-square standardized value of 1.831 and an average standard error of 2.613.

On September 15, 2020, after removing outlier samples, 25 out of 40 samples remained. The analysis took stable methods. The resulting statistics were a root-mean-square standardized value of 1.440 and an average standard error of 2.656.

On February 2, 2021. after removing outlier samples, 34 out of 40 samples remained. The analysis took stable methods. The resulting statistics were a root-mean-square standardized value of 1.296 and an average standard error of 2.772.

4.2.3 Noise map around the wind farm

The noise maps were generated using kriging techniques to interpolate noise distribution from sampling locations. The spatial data used as the base map included satellite imagery and city boundaries. The noise map showed spatial distribution of sound level in areas of 1,2514,675 square m. The sound level areas generated by interpolation analysis separated every 2.5 dBA interval sound levels from 40 – 70 dBA. The noise maps of the study area present the noise levels of four measurements, as shown in Figure 4.4.

1) Noise map of May 30, 2018, Fig. 6 (A);

The predicted sound levels ranged from 40 to 60 dB(A). In the industrial area, sound levels were predicted to range between 45 and 60 dB(A), while in the agricultural area, sound levels were predicted to range from 40 to 50 dB(A). In the residential area, sound levels were predicted to range from 45 to 60 dB(A) in the south and southeast, and from 45 to 55 dB(A) in the northeast.

Significantly, higher predicted sound levels ranging from 55 to 60 dB(A) were predicted in the south, which is located in close proximity to a highway in

both the industrial and residential areas. On the other hand, the lower predicted sound levels ranged from 40 to 45 dB(A) in the northwest, west, and east, which are agricultural areas.

2) **Noise map of June 25, 2019, Fig. 6 (B);**

The predicted sound levels ranged from 40 to 70 dB(A). In the industrial area, sound levels were predicted to range between 50 and 70 dB(A), while in the agricultural area, sound levels were predicted to range from 40 to 55 dB(A). In the residential area, sound levels were predicted to range from 45 to 70 dB(A) in the south and southeast, and from 45 to 50 dB(A) in the northeast.

Significantly, higher predicted sound levels ranging from 55 to 70 dB(A) were predicted in the south, which is located in close proximity to a highway in both the industrial and residential areas. On the other hand, the lower predicted sound levels ranged from 40 to 45 dB(A) in the east, which are agricultural areas.

3) **Noise map of September 15, 2020, Fig. 6 (C);**

The predicted sound levels ranged from 40 to 55 dB(A). In the industrial area, sound levels were predicted to range between 45 and 55 dB(A), while in the agricultural area, sound levels were predicted to range from 40 to 55 dB(A). In the residential area, sound levels were predicted to range from 45 to 55 dB(A) in the south and southeast, and from 50 to 55 dB(A) in the northeast.

Significantly, higher predicted sound levels ranging from 50 to 55 dB(A) were predicted in the east, and northeast, which are both the industrial and residential areas. On the other hand, the lower predicted sound levels ranged from 40 to 45 dB(A) in the west, which are agricultural areas.

4) **Noise map of February 2, 2020, Fig. 6 (D);**

The predicted sound levels ranged from 40 to 55 dB(A). In the industrial area, sound levels were predicted to range between 45 and 55 dB(A), while in the agricultural area, sound levels were predicted to range from 40 to 50 dB(A). In the residential area, sound levels were predicted to range from 45 to 50 dB(A).

In conclusion, the noise map of four measurements show that the average sound level ranged between 40.0 and 70.0 dB(A). The sound levels vary based on the location and time of measurement. The industrial areas consistently show higher sound levels, ranging from 45 to 70 dB(A) in the maps. The residential areas show sound levels ranging from 45 to 60 dB(A) in most cases, but higher sound levels show in proximity to highways. The agricultural areas generally have lower sound levels, ranging from 40 to 55 dB(A).

The maps demonstrate that the south and southeast regions consistently have higher sound levels, mainly due to their close proximity to highways in both industrial and residential areas. It is possible that traffic is the primary noise source. The finding is similar to the study in Taiwan (Tsai et al., 2009), Malaysia (Segaran et al., 2020), and India (Manojkumar et al., 2019). On the other hand, the northwest, west, and east regions consistently exhibit lower sound levels, indicating their agricultural nature. The sound levels in these areas range from 40 to 50 dB(A), and can reach up to 50 to 55 dB(A). It is possible that the noise is caused by the sound of wind flowing through vegetation (Paulraj & Välisuo, 2017).

In this study, the standard deviation of the measured sound level ranged from -5.18 to 4.34 dB(A) at the same distance. These significant variations may be influenced by background noise. By the way, the sound level results obtained through the batch method represent specific time periods and should not be considered as a definitive representation of the overall sound level in the area.

30 May 2018

Count : 40 | Skewness : 1.0329
Min : 42.6 | Kurtosis : 3.4906
Max : 62.9 | 1-st Quartile : 44
Mean : 48.158 | Median : 47.15
Std. Dev. : 4.928 | 3-rd Quartile : 50.9

25 June 2019

Count : 40 | Skewness : 0.77558
Min : 40.3 | Kurtosis : 4.2436
Max : 65.8 | 1-st Quartile : 48
Mean : 50.9 | Median : 50.15
Std. Dev. : 4.8165 | 3-rd Quartile : 53.05

15 September 2020

Count : 40 | Skewness : 0.40971
Min : 35.7 | Kurtosis : 3.5015
Max : 65.3 | 1-st Quartile : 41.7
Mean : 46.675 | Median : 47.1
Std. Dev. : 6.1438 | 3-rd Quartile : 50.05

2 February 2021

Count : 40 | Skewness : 0.65905
Min : 41 | Kurtosis : 2.7562
Max : 56 | 1-st Quartile : 44.05
Mean : 46.797 | Median : 46
Std. Dev. : 3.6672 | 3-rd Quartile : 49

Figure 4.2 Histogram and statistics of noise data collections

30 May 2018
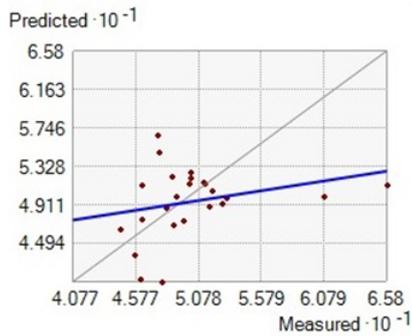
Sample	26
Model	Circular
Anisotropy	True
Mean	0.142
Root-Mean-Square	3.994
Mean Standardized	0.056
Root-Mean-Square Standardized	1.588
Average Standard Error	2.515
Regression function	0.563 * x + 20.436

25 June 2019
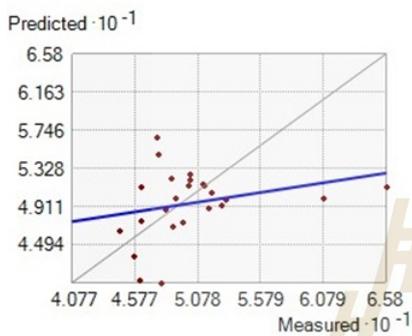
Sample	24
Model	Guassian
Anisotropy	True
Mean	-0.730
Root-Mean-Square	5.139
Mean Standardized	-0.205
Root-Mean-Square Standardized	1.831
Average Standard Error	2.613
Regression function	0.212 * x + 38.815

15 September 2020
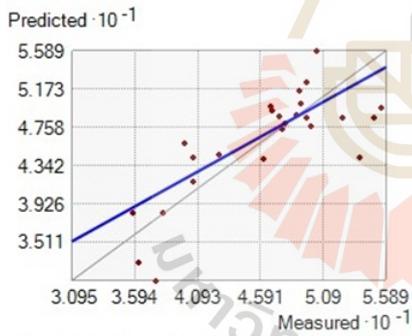
Sample	25
Model	Stable
Anisotropy	True
Mean	-0.142
Root-Mean-Square	3.824
Mean Standardized	-0.054
Root-Mean-Square Standardized	1.440
Average Standard Error	2.656
Regression function	0.758 * x + 11.712

2 February 2021
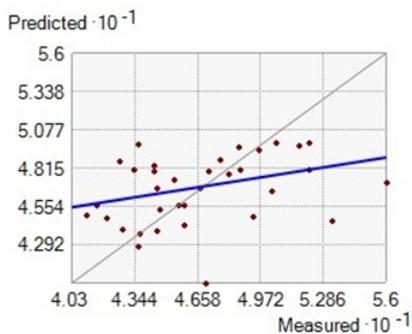
Sample	34
Model	Stable
Anisotropy	True
Mean	-0.043
Root-Mean-Square	3.594
Mean Standardized	-0.015
Root-Mean-Square Standardized	1.296
Average Standard Error	2.772
Regression function	0.217 * x + 36.747

Figure 4.3 The validation statistics of noise data collections
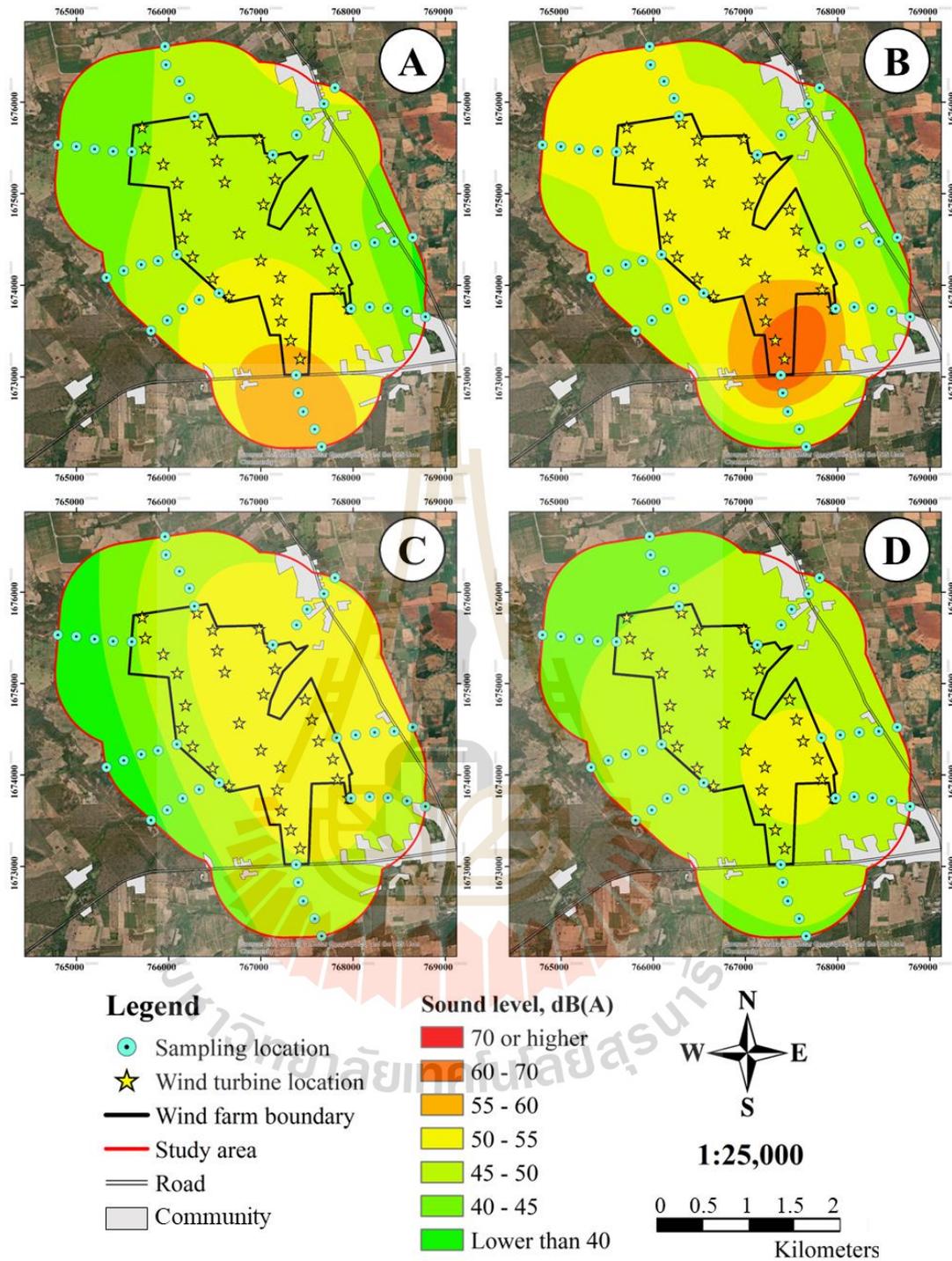
Figure 4.4 Noise distribution around the windfarm

4.2.4 Statistical analysis
1) Noise levels at different time

The sound level was obtained from the two automated sound monitoring stations in low-density residential areas. The plot of sound levels with the time of the day is presented in Figure 4.5. The field measurement data and statistical analysis are shown in Table 4.2 and Table 4.3. A T-test of the differences in sound level of daytime and nighttime periods shows significant differences (p>0.05) between the two stations.
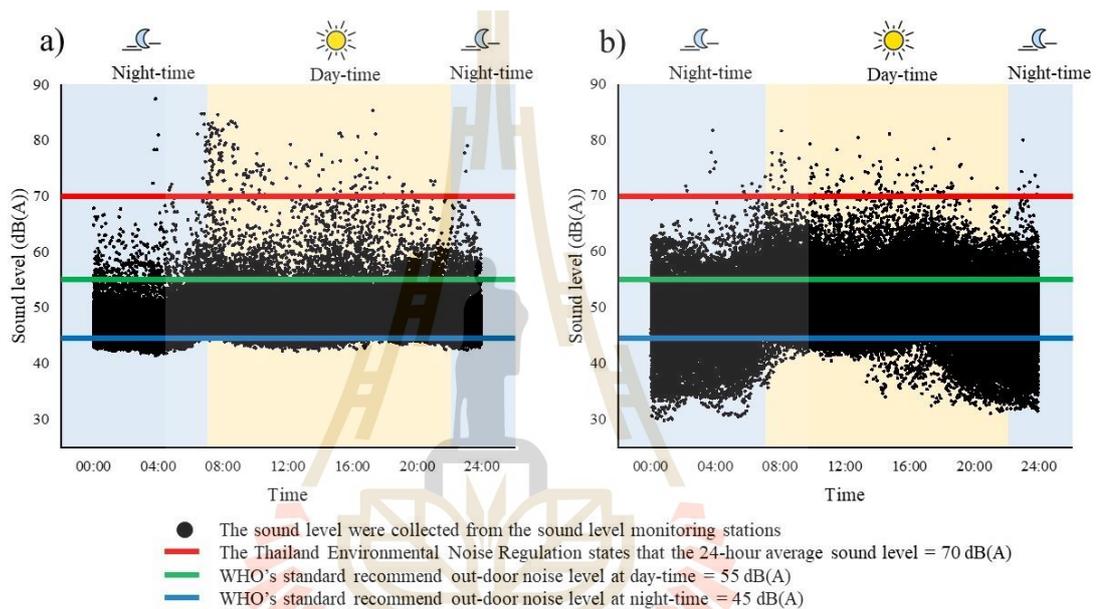
Figure 4.5 Variation of sound levels with time of the day
(Sep 13th, 2019 – Apr 30th, 2021) a) The southeast monitoring station,
b) The northeast monitoring station

A southeast monitoring station, Fig. 7(a); The equivalent sound level ($L_{Aeq,5min}$) was 41.51-87.56 dB(A), and the average sound level was 48.32+3.08 dB(A). The daytime sound level, with a mean of 48.98+3.07 dB(A), is higher than the nighttime sound level, with a mean of 47.20+2.77 dB(A). The results of the southeast monitoring station show that sound levels peaked between 7:00 am and 9:00 am and between 3:00 pm and 5:00 pm. Due to its proximity to the highway, the primary noise source at the southeast monitoring station is traffic during peak commuting hours (rush hour).

A northeast monitoring station, Fig. 7(b); The equivalent sound level ($L_{Aeq,5min}$) was 29.90-81.82 dB(A), and the average sound level was 49.51+4.85 dB(A). The daytime sound level, with a mean of 50.35+4.65 dB(A), is higher than the nighttime sound level, with a mean of 47.81+4.81 dB(A). The results of the northeast monitoring station show that sound levels were lower between 2:00 pm and 7:00 am (including

evening and nighttime). This is because most villagers come home and rest in the evening after work. There were no other activities in the residential area so that the surrounding environment could influence the sound level.

Table 4.2 The descriptive statistics and noise indicators

Descriptive statistics/Indicators	Southeast Monitoring Station			Northeast Monitoring Station		
	Total	Day	Night	Total	Day	Night
Count	162,351	101,769	60,582	166,596	111,382	55,214
Minimum, dB(A)	41.51	42.27	41.74	29.90	31.14	29.90
Maximum, dB(A))	87.56	85.50	87.56	81.82	81.72	81.82
Mean, dB(A)	48.32	48.98	47.20	49.51	50.35	47.81
Standard Deviation, dB(A)	3.08	3.07	2.77	4.85	4.65	4.81
L_{eq}, dB(A)	52.40	53.21	50.54	52.99	53.77	51.34
L_{dn}, dB(A)	57.43	-	-	58.19	-	-
L_{den}, dB(A)	57.68	-	-	58.46	-	-
L_{10}, dB(A)	51.73	52.28	50.31	55.63	56.28	53.91
L_{50}, dB(A)	47.94	48.58	46.79	49.12	49.91	47.24
L_{90}, dB(A)	44.98	45.81	44.32	44.25	45.20	42.90
TNI, dB(A)	41.98	41.69	38.28	59.77	59.52	56.94
L_{np}, dB(A)	55.45	55.74	53.38	62.65	63.03	60.27

Table 4.3 The t-test for the difference between the means of day-time and night-time sound levels

Stations		t	df	Sig. (2-tailed)	Mean	95% Confidence Interval	
						Lower	Upper
Southeast Monitoring Station	L_d	5,092.41	101,768	.000	48.97	48.96	48.99
	L_n	4,196.36	60,582	.000	47.20	47.18	47.22
Northeast Monitoring Station	L_d	3,617.73	111,381	.000	50.35	50.32	50.38
	L_n	2,336.37	55,213	.000	47.81	47.77	47.85

Additionally, most of the sound measurements taken at the northeast monitoring station are larger than the range measured from the southeast monitoring station, which indicates that the variability of the sound measurements at the northeast station was higher than the variability of sound pollution at the southeast station.

2) Comparison of measured noise levels with regulation standards

The existing noise levels monitored in this study were compared with the noise control standards set by the World Health Organization (WHO) guidelines and the Noise Control Act in Thailand, as shown in Table 4.2 and Table 4.3.

Thailand's noise control act: The comparison between equivalent sound levels (L_{eq}) (52.40 and 52.99 dB(A)) and maximum sound levels (87.56 and 81.82 dB(A)) of the two monitoring stations with standard shows that the noise levels are lower than a 24-hour exposure level of 70 dB(A).

WHO's guidelines for daytime sound levels: The comparison of daytime equivalent sound level (L_{eq}) of the two monitoring stations with the guidelines shows that the noise levels (53.21 and 53.77 dB(A)) exceed the recommended sound levels (53 dB(A)).

WHO's guidelines for nighttime sound levels: The comparison of the nighttime equivalent sound level (L_{eq}) of the two monitoring stations with the guidelines shows that the noise levels (50.54 and 51.34 dB(A)) exceed the recommended sound levels (45 dB(A)).

WHO's recommended levels for wind turbine noise sources: The comparison of the day–evening–night sound level (L_{den}) of the two monitoring stations with the guidelines shows that the noise levels (57.68 and 58.46 dB(A)) exceed the recommended sound levels (45 dB(A)). The measurement at the wind farm border (49.0±4.9 on May 30, 2018; 52.3±5.8 on June 25, 2019; 51.1±3.5 on September 15, 2020; 49.2±1.9 on February 2, 2021) shows that the sound level exceeds the recommended sound levels (45 dB(A)).

WHO's recommended levels for traffic noise sources: The comparison of the day–evening–night sound level (L_{den}) of the two monitoring stations (57.68 and 58.46 dB(A)) with standard shows that the noise levels were within the recommended sound levels (with 53 dB(A)). The comparison of the compared nighttime equivalent sound level (L_{eq}) (50.54 and 51.34 dB(A)) exceeds the recommended sound levels (45 dB(A)).

Traffic Noise Index (TNI): TNI indicates the degree of variations (degree of annoyance) for the traffic flow scenario. The higher value of TNI indicates more disturbances due to fluctuating noise concerning L_{10}. From the measurement, TNI was higher during the day at 41.69 and 59.52 dB(A), compared to nighttime at 38.28 and 56.94 dB(A) for southeast and northeast monitoring stations, respectively. Compared with standard, Both TNI are lower than the recommended sound levels (with 74 dB(A)). Comparing the two monitoring stations, TNI was higher at the northeast station, which is a residential colony next to a rural road. From observation, vehicles

on the rural road, which are cars, trucks, and motorcycles, were the major contributors to the noise pollution in the community. This result is similar to the result published by Ky et al. (Ky et al., 2021)

Noise Pollution Level (L_{np}): L_{np} indicates the degree of annoyance caused by fluctuating noise. From the measurement, L_{np} was higher during the day at 55.74 and 63.03 dB(A), compared to nighttime at 53.38 and 60.27 dB(A) for southeast and northeast stations. Comparing with standard, Both L_{np} are lower than the recommended sound levels (with 72 dB(A)). Respectively, Similar to TNI, L_{np} was higher at the northeast station.

4.2.5 Evaluation of noise risk zone and impact on human health

Land use in the study was differentiated into three categories, residential, Industrial, and agricultural. The noise map (Figure 4.4) shows that the residential zone is in the higher noise levels. Nighttime noise levels in the two villages exceed WHO recommendations. The noise level in the industrial zone or wind farm area was 45- 59.9 dB(A). The noise level in the agricultural area was 37.5 - 54.9 dB(A), and it occasionally reached up to 59.9 dB(A) because of the wind blowing on the vegetation (Paulraj & Välisuo, 2017). Based on the measurement of the automated sound monitoring station located in the residential area, the noise level of the roadside residential colonies is between 29.90 and 87.56 dB(A). The values of TNI and L_{np} were 38.28-59.77 and 53.38-63.03 dB(A), respectively. The maximum TNI and L_{np} values were 59.52 and 63.03 dB(A) during daytime at the side of the rural road (the northeast monitoring station).

To minimize the nuisance of noise pollution in this area, a mitigation measure, e.g., proper traffic management and strict enforcement of noise pollution control rules and regulations, is required. Many traffic noise management that is suitable for low-density residential, e.g., demarcation of noise-sensitive zones for speed reduction and increasing greeneries and open spaces along the roadside. The noise assessment shows that the noise levels measured at two monitoring stations are generally lower than the 24-hour time period and maximum permissible sound levels set by Thailand's noise and vibration control act. However, the daytime, nighttime, and day-evening-night sound levels are higher than the recommended sound levels set by WHO for the community, wind turbine, and traffic noise sources. Moreover, the measured traffic noise index (TNI) and noise pollution level (LNP) are both lower than the recommended sound levels. Overall, while the noise levels at the two monitoring stations comply with Thailand's Noise and vibration control act, they exceed WHO's recommended sound levels for some noise sources, indicating a need for further noise reduction measures. Nighttime noise can cause sleep disturbances, leading people to

suffer from daytime sleepiness, tiredness, annoyance, mood changes, and decreased short-term well-being and cognitive performance. Long-term sleep disturbance can lead to adverse cardiometabolic, psychiatric, and social outcomes (Halperin, 2014). This suggests that noise is potentially harmful to human health and well-being and may require further measures to reduce noise pollution.

WHO defines the noise levels and their impacts on humans as the following: more than 30 dB(A): not restful sleep; more than 30 dB(A): not restful sleep; more than 75 dB(A): harmful; more than 120 dB(A): painful. The average noise level at night in the villages was 47.20+2.77 and 47.81+4.81 dB(A), which exceeds the WHO recommendation. In this case, nighttime noise was caused by traffic, affects objectively measured sleep physiology, and subjectively assessed sleep disturbance in adults (World Health Organization, 2022). The sleep disturbance causes people to suffer from daytime sleepiness and tiredness, annoyance, mood changes, and decreased short-term well-being and cognitive performance the next day. The long-term sleep disturbance causes adverse outcomes of cardiometabolic, psychiatric, and social (Halperin, 2014).

4.3 Noise prediction model

4.3.1 Field measurement data

The ranges and averages of the field measurements from four measurement points are shown in Table 4.4. Comparing meteorological parameters between the field measurement and historical data obtained from the Thai Meteorological Department (TMD) of Nakhon Ratchasima province from 1990-2019 reveals that the measurement data is within the range of the historical data. The average sound level was higher at the measurement point closer to the wind turbine. A plot between sound level and time of the day for all measurement points over 72 hours is shown in Table 4.4.

Table 4.4 Field measurement data

Parameters	Units	Field measurement data				Historical data
		100 m	200 m	300 m	400 m	(1990-2019)
Sound level (Mean±SD)	dB(A)	46.8 ±5.4	47.8 ±6.3	43.0 ±4.9	43.6 ±4.6	-
Wind direction	Degree	62.1±54.8				-
Wind speed	m/s	1.2±1.1				0.9–1.3
Temperature	°C	28.5±2.7				24.4–30.1
Humidity	%	67.7±2.9				62.0–81.0
Pressure	hPa	998.5±0.4				997.7–1,013.8

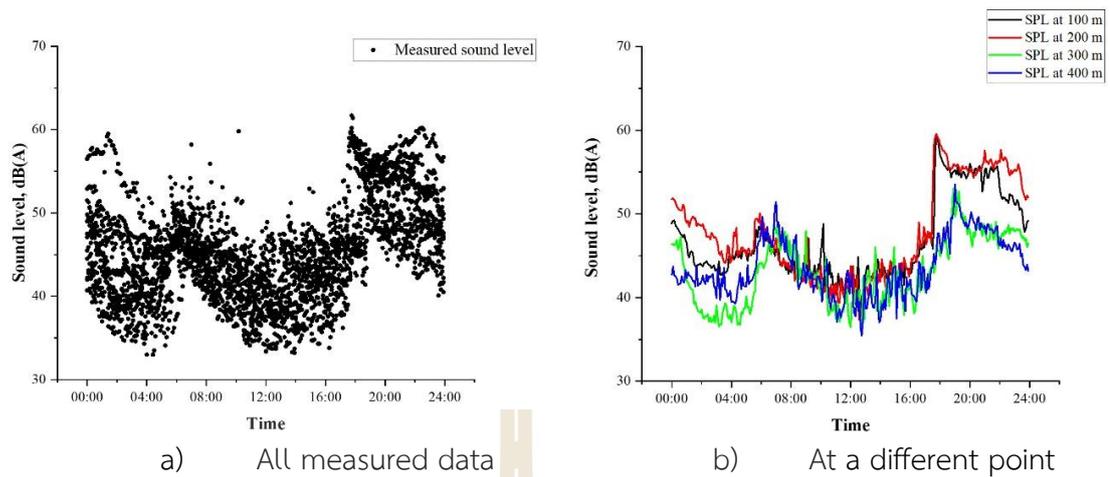| a) | All measured data | b) | At a different point |

Figure 4.6 Plots between sound level and time

From Figure 4.6, the difference between the sound at various times of the day can be seen. The higher sound level around morning and evening indicated the effect of human activity from the road and village nearby. The U.S. The Environmental Protection Agency (EPA) defines daytime sound levels as those that occur between the hours of 7.00 am and 10.00 pm and nighttime sound levels as those that occur between 10.00 pm and 7.00 am (United States Environmental Protection Agency Office of Noise Abatement and Control, 1974). The high noise levels in the daytime compared to the nighttime are typical for a quiet residential area.

The measured sound level, 33.0–61.7 dB(A), was lower than Thailand's standard, which sets an average level of 70 dB(A) for 24 hours and a maximum level of 115 dB(A). However, some measurements exceed the WHO's recommended value, 45 dB (A), for the wind turbine noise and the WHO's recommended value for community noise in outdoor living areas, 55 dB L_{Aeq} (World Health Organization, 2022). This means that noise in the study area could potentially be harmful to human health. Hence, mitigation measures should be implemented to protect residents in study area.

4.3.2 Data Preparation

The field measurement data used for model input was within a wind turbine's cut-in speed condition. The cut-in speed is when the wind turbine blades start to rotate and generate power. The wind turbines at the study site are the G114-2.0 MW model, which has a cut-in wind speed of 2.5 m/s. The remaining dataset (n = 576) was divided into training and testing. A ratio of 70/30 for training and testing datasets was a popular ratio, and it was considered the best ratio for training and validating the models (Nguyen et al., 2021). The number of training data was 399 (69.3%), and testing was 177 (30.7%). The distribution plot of the training and testing datasets with sound levels is shown in Figure 4.7.
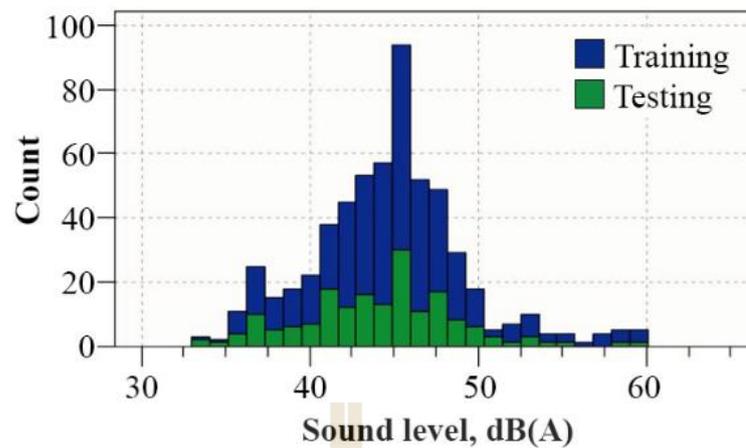
Figure 4.7 The distribution of the training and testing dataset

4.3.3 Modeling

The modified datasets were used to generate models from the auto-numerical node with default values. When an automated modeling node is executed, the node estimates candidate models. The model candidate provided four modeling methods: CHAID, CART, Linear, and Neural Network. The ensemble model combines the other models to produce one optimal predictive model. The default ensemble method is voting. The voting operates by counting how many times each predicted value is selected and then choosing the value with the greatest cumulative count.

4.3.4 Predictor Importance

The predictor importance chart helps indicate the relative importance of each predictor in estimating the model. In Figure 4.6, the predictor importance chart of the CHAID, CART, Linear, and Neural network models reveal that distance is the primary predictor, followed by temperature, time, and wind speed.

4.3.5 Model performance evaluation

Table 4.5 shows the comparison of the statistical analysis for model evaluation. Considering the R-Squared (R^2), the top 3 best performances were the Ensemble model (0.613), CHAID (0.608), and CART (0.608). Comparing the RMSE and MAE values of the models in Table 4.5 indicates the Ensemble as the premier model with the lowest values of 2.919 and 2.328, respectively. Therefore, the Ensemble model was selected as a prediction model. The ensemble model was further validated using cross-validation, splitting a dataset into training and testing subsets.

The Ensemble model was further validated using cross-validation by splitting a dataset into training and testing subsets. In this paper, RMSE and MAE are utilized to assess the performance of the forecasting model. As shown in Table 4.6, The percentage difference between training and testing, RMSE (10.08%) and MAE (5.89%) is low.

It indicates that the model is not overfitting (Kim & Simon, 2014). Thus, the proposed model could forecast the sound level with a reasonable level of accuracy. The metrics RMSE and MAE also validate the effectiveness of the model.

Table 4.5 Comparison of performance metrics of five models

Model	R^2	RMSE	MAE
CHAID	0.608	2.871	2.437
CART	0.608	2.871	2.564
Linear	0.276	3.903	3.053
Neural network	0.372	3.848	3.011
Ensemble	0.613	2.919	2.328

Table 4.6 Ensemble model validation performance metric

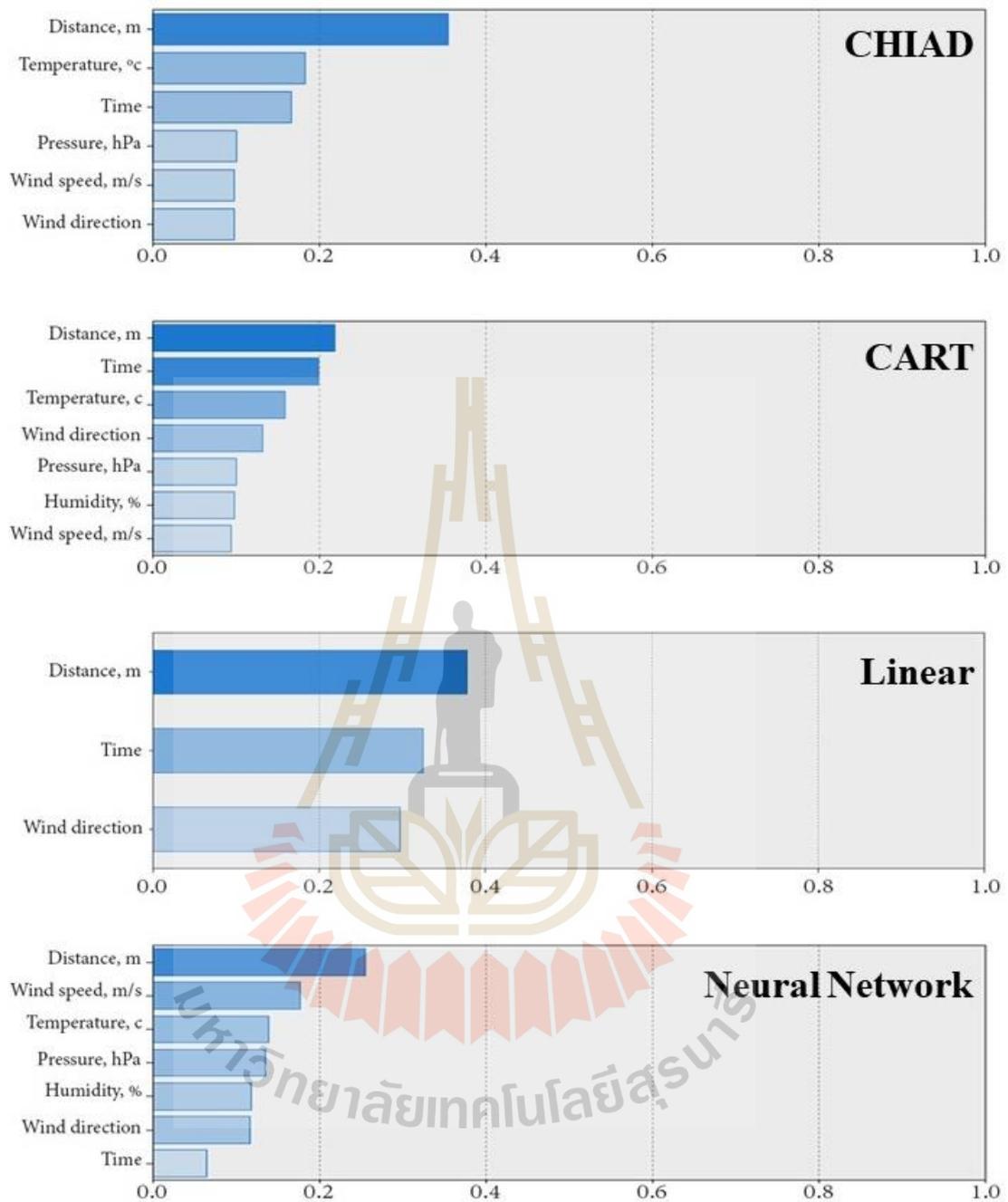| Partition	RMSE	MAE
Training	2.818	2.191
Testing	3.134	2.328
% Difference	10.08	5.89

Figure 4.8 Predictor importance chart

The performances of the models were visually compared using the gain chart plots. The plot presents accumulated gains % to percentile for training and testing datasets. The gain chart in Figure 4.9 indicates that the models are exemplary because the charts rise steeply toward 100% approximately and then level off.
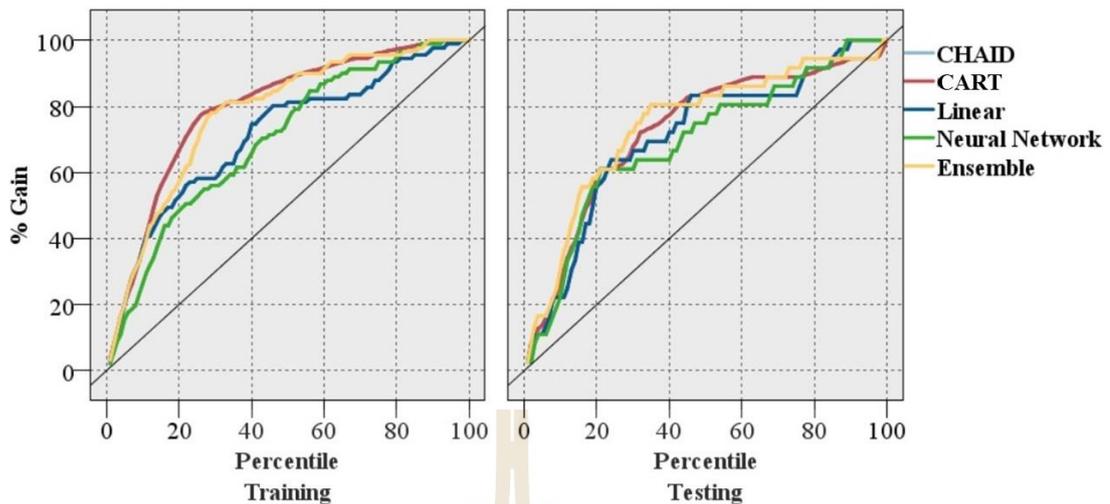
Figure 4.9 Gain charts

4.3.6   Maximum sound level prediction for worst-case scenario

The maximum sound level is predicted using an ensemble model. The result is obtained through a voting mechanism that combines the predictions from CHAID, CART, Linear, and Neural algorithms. The modeling steps can be graphically illustrated as SPSS modeler flow, as shown in Figure 4.10.
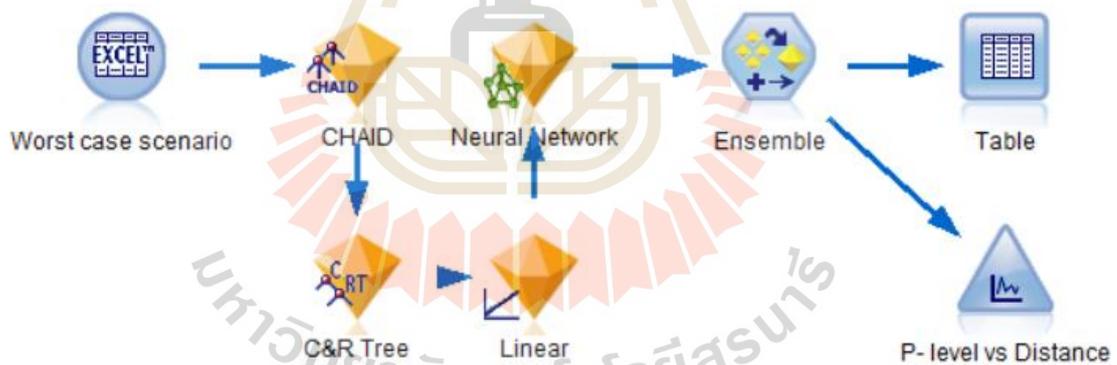
Figure 4.10 Prediction flow chart

The sound level prediction model results reveal patterns in the relationship between the predicted sound level and distance, as shown in Figure 4.11. Overall, the predicted sound level during nighttime is higher than during daytime. Focusing on the predicted sound level during nighttime, it initially increases as the distance from the source increases, reaching a peak of 52.2 dB at 160 m, representing the maximum sound level. Then, the predicted sound level starts to decrease at a distance of 200 m. However, it eventually reaches a stationary state at a distance of 360 m, maintaining a constant sound level of 43.9 dB even as the distance increases.

Figure 4.11 Relationship between predicted sound level and distance

Chapter V
CONCLUSION AND RECOMMENDATIONS

This chapter presents the conclusions and recommendations of the study based on the research objectives, which are as follows: (1) To study the propagation of wind turbine noise by generating a noise map, and (2) To investigate sound level with a noise prediction model. The study has concluded these objectives and provides future research and development recommendations.
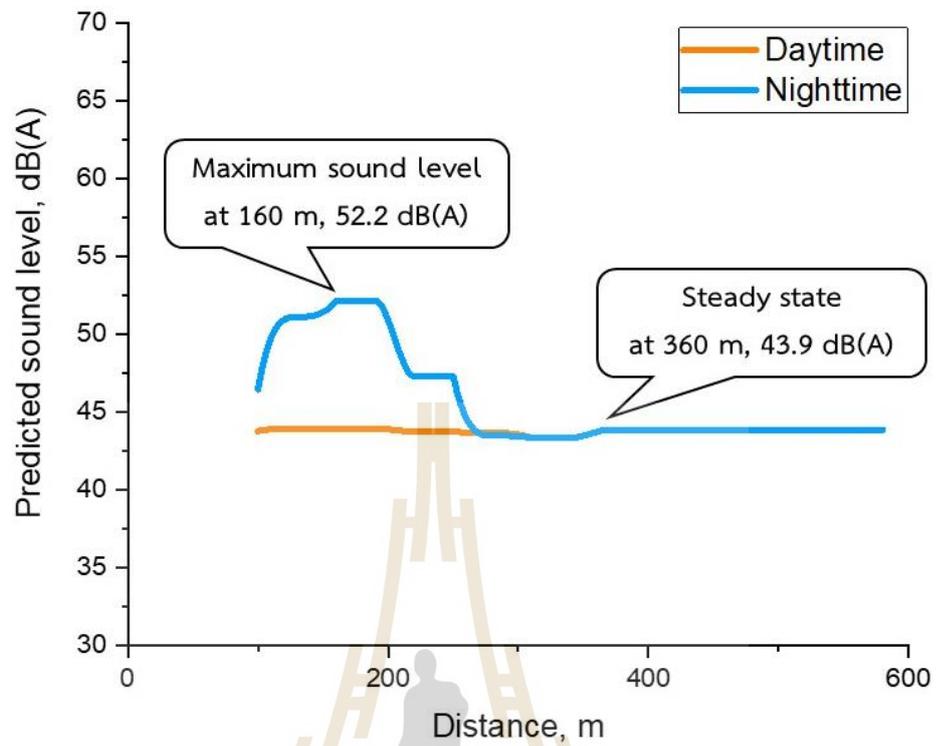
5.1   Overview of the Study

This research aims to study the potential effect of wind farm noise on the community at 800 m. radius from the boundary of the wind farm located in Huai Bong Sub-District, Dankhuntod District, Nakhon Ratchasima Province in Thailand. The study can be divided into 2 parts; 1. Noise map generation and 2. Developing a model to predict sound levels caused by a wind turbine.

For the noise map generation, the sound level was measured On-site. The field measurement following a guidance note on sound level assessment of wind turbine operations at EPA-licensed sites (NG3) by USEPA. The field measurement and the secondary data, aerial photo, and transportation route are used to generate noise contour and map. The techniques employed for this purpose included interpolation and overlay analysis. Specifically, noise maps were generated using kriging as the interpolation technique. The process of developing these noise maps was facilitated through the use of ArcGIS Desktop 10.5 software.

For the sound level prediction generation, A model to predict sound was performed using IBM SPSS Modeler. IBM SPSS Modeler had selected from modeling program comparison. Other field measurements had measured sound levels, wind speed, wind direction, temperature, and moisture. The prediction model was generated by estimates and compares models for continuous numeric range outcomes using a number of different methods such as regression, generalized linear, SVM algorithm, C&R tree, CHAID algorithm, KNN algorithm, Neural network, and Ensemble algorithm. The model was used to predict a sound level that a nearby community received in a worst-case scenario. The worst-case scenario had made from Nakhon Ratchasima Climatological data for the period 1990-2019.

5.2 Conclusion

5.2.1 Noise mapping

In this study, sound levels were measured at distances ranging every 200 m intervals up to 800 m from the wind farm, totaling 40 points. A noise measurement method followed the guidance note on noise assessment of wind turbine operations at EPA's licensed sites (NG3) by USEPA. The development of noise maps using GIS for the area around the wind farm, based on field data measured yearly from 2018 to 2021, is presented. ArcGIS desktop 10.5 software was used in this study to develop noise maps and land use maps. These noise maps were generated using kriging interpolation techniques on geostatistical analyst.

The noise map from the four measurements indicates that the average sound level was between 30.0 and 70.0 dB(A). In agricultural areas, the sound level ranged from 30.0 to 44.9 dB(A), with peaks reaching between 55.0 and 59.9 dB(A) due to the sound of the wind flowing through vegetation, which served as the background noise (Paulraj & Välisuo, 2017). High noise levels between 55.0 and 70.0 dB(A) were found around the roads, indicating that traffic is the primary noise source. This finding is consistent with studies conducted on traffic noise in urban noise mapping environments in Taiwan (Tsai et al., 2009), noise mapping in urban environments in India (Manojkumar et al., 2019), and noise mapping in residential environments in Malaysia (Segaran et al., 2020).

Theoretically, sound levels are measured on a logarithmic scale. Doubling the distance from a wind turbine reduces the sound level by six decibels (Alberts, 2006). However, in this study, doubling the distance from a wind turbine led to sound level variations ranging from -5.18 to 4.34 dB(A). Sound levels at the same distance showed significant variations influenced by background noise. The wind turbine generates a noise level that becomes equal to the background noise level when the wind speed is approximately 12 m/s and the distance exceeds 100 m from the receiver (Katinas et al., 2016), However, the 30-year wind speeds ranged from 0.9 to 1.3 m/s at Nakhon Ratchasima Province, significant lower. which means that the noise level that generated from wind turbine becomes equal to the background noise level even when the distance does not exceed 100 m from the receiver. There is a possibility that a sound source other than wind turbine noise could be dominating this area. The dominating noise could be generated by wind blowing on the microphone or vegetation (Bolin, 2006), or even from traffic noise.

5.2.2 Noise assessment

The sound level was obtained from the two automated sound monitoring stations located southeast and northeast of the wind farm. The monitoring stations measured a noise level every 5 minutes on average, L_{Aeq} (dB(A)). The data used in this study was measured from September 13th, 2019 to April 30th, 2021, covering 596 days. Independent sample t-test analysis was performed to differentiate between noise levels during daytime and nighttime periods. And sound level descriptors were calculated to compare measured noise levels with standards such as Thailand's noise control act, World Health Organization guidelines, Traffic Noise Index, and Noise pollution level, to determine the impact of noise on health and evaluate noise pollution.

The results of the t-test analysis conducted from both stations indicate significant differences in sound levels between daytime and nighttime periods ($p > 0.05$), with sound levels during daytime periods being significantly higher than those during nighttime periods. The sound levels peaked between 7:00 am and 9:00 am and again between 3:00 pm and 5:00 pm, There is a possibility that the primary noise source during rush hours is traffic, and several factors contribute to the reinforcement of sound levels during this period, including both traffic and human activities. But some sound levels are higher during off-peak hours than during rush hours due to increased traffic flow, allowing cars to travel at higher speeds that generate higher levels of traffic noise (Yang et al., 2020).

The comparison results from both stations were the equivalent sound levels (L_{eq}) and maximum sound level (L_{max}) with standard, which shows that the noise levels are lower than a 24-hour exposure level (70 dB(A)) stated in Thailand's noise control act. However, they exceed the WHO's guidelines for sound levels during daytime periods (53 dB(A)) and during nighttime periods (45 dB(A)). The day–evening–night sound levels (L_{den}) also exceed the WHO's recommended sound levels for wind turbines and traffic noise sources (45 dB(A)). Although traffic noise index (TNI) and noise pollution levels (LNP) are lower than the recommended sound levels, nighttime noise still can cause sleep disturbances, leading people to suffer from daytime sleepiness, tiredness, annoyance, mood changes, and decreased short-term well-being and cognitive performance (World Health Organization, 2022). Long-term sleep disturbance can lead to a range of adverse outcomes, including cardiometabolic, psychiatric, and social repercussions. The long-term sleep disruption has been increased risks in various health. (Halperin, 2014).

5.2.3 Noise prediction modeling

In this study, the field measurement data including sound levels, wind direction, wind speed, temperature, humidity, and pressure were measured in 5-min intervals for three days (From 1.00 pm, 20 February 20th, 2023, to 1.00 pm, February 23rd, 2023) with distances of 100 m intervals up to 400 m at the northeast corner of the wind farm, a total of 864 times per point. To investigate sound level with a noise prediction model. Field measurement data was processed into a consistent and usable form. Data processing included data cleaning, data structuring, data transformation, and data filtering. They divided it into two datasets, with a ratio of 70% for training and 30% for testing. SPSS Modeler is used to build a prediction model from the auto-numerical node with default values. To explores possible model and ranks each candidate model based on the correlation between predicted and observed values for each model. The performance error of the developed model was evaluated using R- squared (R^2) , Root Mean Squared Error (RMSE) , and the Mean Absolute Error (MAE),

The candidate model provided five modeling methods, including CHAID, CART, Linear, Neural network, and an Ensemble model that combines the other models to produce an optimal predictive model. The predictor importance chart reveals that distance is the primary predictor, followed by temperature, time, and wind speed. The results of the model evaluation show that the Ensemble model has the highest R-Squared value (0.613) and the lowest values for RMSE (2.919) and MAE (2.328). The Ensemble model proves to be the most suitable technique, as it involves weighing several individual models and combining them to improve predictive performance (Sagi & Rokach, 2018). Several researchers have observed better prediction performance with Ensemble models compared to others (Xiao et al., 2018). The performance of the models was visually compared using gain chart plots. The chart of the Ensemble model rises steeply to a faster rate than other algorithms in both the training and testing sections, reaching a 100% gain, and then levels off. Additionally, the Ensemble model underwent cross-validation by splitting the dataset into training and testing subsets. The percentage difference between training and testing for Root Mean Square Error (RMSE) (10.08%) and Mean Absolute Error (MAE) (5.89%) is low, indicating that the model is not overfitting (Kim & Simon, 2014). Overfitting occurs when the model cannot generalize and fits too closely to the training dataset instead. The Ensemble model was ensured to be capable of being the prediction model.

5.2.4 Maximum sound level prediction for worst-case scenario

The prediction model has been chosen based on evaluations conducted to estimate the maximal sound levels generated in worst-case scenarios. This aims to address the gap in noise regulations for wind turbine noise sources in Thailand. The ensemble model predicts the maximum sound level through a voting mechanism that combines the predictions from CHAID, CART, Linear, and Neural network algorithms. Worst-case scenarios refer to environmental conditions that have the potential to cause the wind turbine to generate high sound levels that can propagate over long distances. These factors were obtained from meteorological data collected over a 30-year period in Nakhon Ratchasima province by the Meteorological Department of Thailand. The results reveal patterns in the relationship between the predicted sound level and distance. Nighttime sound levels are higher than daytime levels. As the distance from the source increases, the nighttime sound level initially rises, reaching a peak of 52.2 dB(A) at a distance of 160 m. Afterward, at a distance of 200 m, the sound level begins to decline and eventually reaches a stationary state at 360 m, maintaining a constant level of 43.9 dB(A).

5.3 Recommendations

1) The noise maps can also be used to identify the vulnerable area compared to the local and the WHO's acceptable thresholds., the decision-makers can identify the areas that require mitigation measures to minimize the nuisance of noise pollution.

2) Implementing Internet of Things (IoT) technology for noise mapping, utilizing continuous noise sensors to generate real-time noise maps accessible through websites or applications. This approach improves data accuracy, enabling more effective noise management and mitigation strategies.

3) The findings from the study on maximum sound level prediction for worst-case scenarios will support the development of future noise regulations for wind turbines in Thailand. The current regulations, which state that regulations on sound impact that the maximum allowable noise level must not exceed 10 dB(A) and the 24-hour A-weighted equivalent continuous sound level must not exceed 70 dB(A) for unknown sound sources, may be considered too high for wind turbine noise. The results of this study can help establish more specific noise limits for wind turbines, which will ensure effective management and mitigation of noise pollution from the wind energy projects in Thailand.

4) Additional research on various machine learning algorithms, such as AdaBoost, Random Forest, Extremely Randomized Trees, and other related algorithms, is recommended for further exploration and investigation in this field.

REFERENCES

Adulaimi, A. A. A., Pradhan, B., Chakraborty, S., & Alamri, A. (2021). Traffic Noise Modelling Using Land Use Regression Model Based on Machine Learning, Statistical Regression and GIS. *Energies, 14*(16), 5095.

Alam, W. (2011). GIS based assessment of noise pollution in Guwahati city of Assam, India. *International journal of environmental sciences, 2*(2), 731-740.

Albaji, A., Rashid, R., Sarijari, M., Salam, Z., Hamid, S. Z. A., & Ali, Y. H. (2021). A Machine Learning for Environmental Noise Monitoring and Classification Using Matlab. In: Universiti Teknologi Malaysia, Johor, Malaysia.

Alberts, D. J. (2006). Addressing wind turbine noise. *Report from Lawrence Technological University*.

Ambika, P. (2020). Chapter Thirteen - Machine learning and deep learning algorithms on the Industrial Internet of Things (IIoT). *Advances in Computers, 117*(1), 321-338. doi:https://doi.org/10.1016/bs.adcom.2019.10.007

Athanasiadis, I. N., Kaburlasos, V. G., Mitkas, P. A., & Petridis, V. (2003). *Applying machine learning techniques on air quality data for real-time decision support.* Paper presented at the First international NAISO symposium on information technologies in environmental engineering (ITEE'2003), Gdansk, Poland.

Attenborough, K. (2002). Sound propagation close to the ground. *Annual Review of Fluid Mechanics, 34*(1), 51-82.

Bechtel, R. B., & Churchman, A. (2003). *Handbook of environmental psychology*: John Wiley & Sons.

Bekele, A., Downer, R., Wolcott, M., Hudnall, W., & Moore, S. (2003). Comparative evaluation of spatial prediction methods in a field experiment for mapping soil potassium. *Soil Science, 168*(1), 15-28.

Bellinger, C., Mohomed Jabbar, M. S., Zaïane, O., & Osornio-Vargas, A. (2017). A systematic review of data mining and machine learning for air pollution epidemiology. *BMC Public Health, 17*(1), 907. doi:10.1186/s12889-017-4914-3

Bigot, A., & Hochard, G. (2019). *Is it possible to predict background noise levels from measured meteorological data with machine learning techniques.* Paper presented at the Proceedings of the 8th International Conference on Wind Turbine Noise, Lisbon, Portugal.

Bishop, T., & McBratney, A. (2001). A comparison of prediction methods for the creation of field-extent soil property maps. *Geoderma, 103*(1-2), 149-160.

Blanchard, T., & Samanta, B. (2019). Prediction of wind turbine noise propagation. *Wind Engineering, 43*(3), 233-246. doi:10.1177/0309524x18780397

Bolin, K. (2006). *Masking of wind turbine sound by ambient noise.* KTH,

Bozkurt, T. S. (2021). Preparation of Industrial Noise Mapping and Improvement of Environmental Quality. *Current Pollution Reports, 7*(3), 325-343.

Bravo-Moncayo, L., Lucio-Naranjo, J., Chávez, M., Pavón-García, I., & Garzón, C. (2019). A machine learning approach for traffic-noise annoyance assessment. *Applied Acoustics, 156*, 262-270.

Bruxella, J. M. D., Sadhana, S., & Geetha, S. (2014). Categorization of data mining tools based on their types. *International Journal of Computer Science and Mobile Computing, 3*(3), 445-452.

Chancham, C., Waewsak, J., Chaichana, T., Landry, M., & Gagnon, Y. (2014). Assessment of Onshore Wind Energy Potential Using Regional Atmospheric Modeling System (RAMS) for Thailand. *Energy Procedia, 52*, 487-496. doi:https://doi.org/10.1016/j.egypro.2014.07.102

Chauhan, A., & Pande, K. K. (2010). Study of noise level in different zones of Dehradun City, Uttarakhand. *Report and opinion, 2*(7), 65-68.

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science, 7*, e623.

Chou, J.-S., Ho, C.-C., & Hoang, H.-S. (2018). Determining quality of water in reservoir using machine learning. *Ecological informatics, 44*, 57-75.

Ciaburro, G., Iannace, G., Puyana-Romero, V., & Trematerra, A. (2021). Machine learning-based tools for wind turbine acoustic monitoring. *Applied Sciences, 11*(14), 6488.

da Silva, G., & Lorena, J. (2017). Computational sound propagation models: An analysis of the models Nord2000, CONCAWE, and ISO 9613-2 for sound propagation from a wind farm. In.

Department of Alternative Energy Development Efficiency. (2015). Alternative Energy Development Plan: AEDP2015. In: Ministry of Energy Bangkok.

Department of Alternative Energy Development Efficiency. (2020). The Alternative Energy Development Plan 2018-2037: AEDP 2018-2037. In: Ministry of Energy Bangkok.

Fahy, F., & Thompson, D. (2015). *Fundamentals of sound and vibration*: CRC press.

Farcaş, F., & Sivertunb, Å. (2010). Road traffic noise: GIS tools for noise mapping and a case study for Skåne region. *Sweden: Citeseer*.

Field, H. L., & Long, J. M. (2018). *Introduction to agricultural engineering technology: a problem solving approach*: Springer.

Forouhid, A. E., Khosravi, S., & Mahmoudi, J. (2023). Noise Pollution Analysis Using Geographic Information System, Agglomerative Hierarchical Clustering and Principal Component Analysis in Urban Sustainability (Case Study: Tehran). *Sustainability, 15*(3), 2112.

Fyhri, A., & Aasvang, G. M. (2010). Noise, sleep and poor health: Modeling the relationship between road traffic noise and cardiovascular problems. *Science of The Total Environment, 408*(21), 4935-4942. doi:https://doi.org/10.1016/j.scitotenv.2010.06.057

García-Gonzalo, E., Fernández-Muñiz, Z., Garcia Nieto, P. J., Bernardo Sánchez, A., & Menéndez Fernández, M. (2016). Hard-rock stability analysis for span design in entry-type excavations with learning classifiers. *Materials, 9*(7), 531.

Gray PhD, L. (2000). Properties of Sound. *Journal of Perinatology, 20*(1), S6-S11. doi:10.1038/sj.jp.7200442

Halperin, D. (2014). Environmental noise and sleep disturbances: A threat to health? *Sleep science, 7*(4), 209-212.

Hansen, C. H. (2001). Fundamentals of acoustics. *Occupational Exposure to Noise: Evaluation, Prevention and Control. World Health Organization, 1*(3), 23-52.

Harman, B. I., Koseoglu, H., & Yigit, C. O. (2016). Performance evaluation of IDW, Kriging and multiquadric interpolation methods in producing noise mapping: A case study at the city of Isparta, Turkey. *Applied Acoustics, 112*, 147-157.

International Electrotechnical Commission. (2013). IEC61672-1 Electroacoustics Sound level meters Part 1: Specifications. Retrieved from https://webstore.iec.ch/publication/5708

International Organization for Standardization. (1996). ISO 9613-2: Attenuation of Sound during Propagation Outdoors- Part 2- General Method of Calculation. *International Organization for Standardization*.

ISO, A. (1996). Attenuation of Sound During Propagation Outdoors–Part 2: A General Method of Calculation (ISO 9613-2). *ISO, Geneva, Switzerland*.

Johnston, K., Ver Hoef, J. M., Krivoruchko, K., & Lucas, N. (2001). *Using ArcGIS geostatistical analyst* (Vol. 380): Esri Redlands.

Katinas, V., Marčiukaitis, M., & Tamašauskienė, M. (2016). Analysis of the wind turbine noise emissions and impact on the environment. *Renewable and Sustainable Energy Reviews, 58*, 825-831.

Kim, K. I., & Simon, R. (2014). Overfitting, generalization, and MSE in class probability estimation with high-dimensional data. *Biometrical Journal, 56*(2), 256-269.

Kondili, E., & Kaldellis, J. K. (2012). Environmental-social benefits/impacts of wind power. In A. Sayigh (Ed.), *Comprehensive Renewable Energy* (pp. 515). Oxford: Elsevier.

Kosarenko, Y. (2021). How to Create Decision Trees for Business Rules Analysis. Retrieved from https://why-change.com/2021/11/13/how-to-create-decision-trees-for-business-rules-analysis/

Kragh, J. (2011). *Traffic noise prediction with Nord2000-an update.* Paper presented at the Proceedings of ACOUSTICS.

Kravchenko, A. (2003). Influence of spatial structure on accuracy of interpolation methods. *Soil Science Society of America Journal, 67*(5), 1564-1571.

Kumar, P., Nigam, S., & Kumar, N. (2014). Vehicular traffic noise modeling using artificial neural network approach. *Transportation Research Part C: Emerging Technologies, 40*, 111-122.

Ky, N. M., Lap, B. Q., Hung, N. T. Q., Thanh, L. M., & Linh, P. G. (2021). Investigation and assessment of road traffic noise: a case study in Ho Chi Minh City, Vietnam. *Water, Air, & Soil Pollution, 232*(7), 259.

Langdon, F. J., & Scholes, W. (1968). The Traffic Noise Index: A Method of Controlling Noise Nuisance.

Licitra, G., & Fredianelli, L. (2013). *Which limits for wind turbine noise? A comparison with other types of sources using a common metric.* Paper presented at the Proceedings of the 5th international conference on wind turbine noise.

Long, M. (2014). Fundamentals of acoustics. *Architectural acoustics*, 39-79.

Longley, P. (2005). *Geographic information systems and science*: John Wiley & Sons.

Lu, G. Y., & Wong, D. W. (2008). An adaptive inverse-distance weighting spatial interpolation technique. *Computers & geosciences, 34*(9), 1044-1055.

Madhavan, S. S., M Kienzler, R. (2019). Introduction to machine learning. Retrieved from https://developer.ibm.com/learningpaths/learning-path-machine-learning-for-developers/machine-learning-overview/?mhsrc=ibmsearch_a&mhq=machine%20learning

Manojkumar, N., Basha, K., & Srimuruganandam, B. (2019). Assessment, prediction and mapping of noise levels in Vellore city, India. *Noise Mapping, 6*(1), 38-51.

McAleer, S., & McKenzie, A. (2011). Guidance note on noise assessment of wind turbine operations at EPA licensed sites (NG3). *Environmental Protection Agency, Office of Environmental Enforcement*.

Michaud, D. S., Marro, L., & McNamee, J. P. (2018). Derivation and application of a composite annoyance reaction construct based on multiple wind turbine features. *Canadian Journal of Public Health = Revue Canadienne de Santé Publique, 109*, 242 - 251.

Moteallemi, A., Bina, B., Minaei, M., & Mortezaie, S. (2017). The Evaluation of Noise Pollution at Samen District in Mashhad, Khorasan Razavi Province, Iran using Geographic Information System. *International Journal of Occupational Hygiene, 9*(4), 179-185.

Mueller, T., Pierce, F., Schabenberger, O., & Warncke, D. (2001). Map quality for site-specific fertility management. *Soil Science Society of America Journal, 65*(5), 1547-1558.

Nanthavanij, S., Boonyawat, T., & Wongwanthanee, S. (1999). Analytical procedure for constructing noise contours. *International Journal of Industrial Ergonomics, 23*(1-2), 123-127.

Năstase, E.-V. (2017). *Influence of the material used to build the blades of a wind turbine on their starting conditions.* Paper presented at the MATEC Web of Conferences.

Nguyen, Q. H., Ly, H.-B., Ho, L. S., Al-Ansari, N., Le, H. V., Tran, V. Q., . . . Pham, B. T. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Mathematical Problems in Engineering*.

Novkovic, D., Stojiljkovic, M., & Lloyd, S. (2017). Noise Predictions from Elevated Sources in Industrial Environments.

NSW Wind Energy Handbook. (2002). Sustainable Energy Development Authority of NSW (SEDA). *Sustainable Energy Development Authority*, 18.

Oerlemans, S., & Schepers, J. G. (2009). Prediction of Wind Turbine Noise and Validation against Experiment. *International Journal of Aeroacoustics, 8*(6), 555-584. doi:10.1260/147547209789141489

Oyedepo, S. O., Adeyemi, G. A., Olawole, O., Ohijeagbon, O., Fagbemi, O., Solomon, R., . . . Efemwenkiekie, U. (2019). A GIS–based method for assessment and mapping of noise pollution in Ota metropolis, Nigeria. *MethodsX, 6*, 447-457.

Pandya, G. (2003). Assessment of traffic noise and its impact on the community. *International Journal of Environmental Studies, 60*(6), 595-602.

Pantazopoulou, P. (2010). Wind turbine noise measurements and abatement methods. *Wind power generation and wind turbine design*, 641-660.

Paulraj, T., & Välisuo, P. (2017). *Effect of wind speed and wind direction on amplitude modulation of wind turbine noise.* Paper presented at the INTER-NOISE and NOISE-CON congress and conference proceedings.

Ralević, N., Glišović, N. S., Djaković, V. D., & Andjelić, G. B. (2014). *The performance of the investment return prediction models: Theory and evidence.* Paper presented at the 2014 IEEE 12th International Symposium on Intelligent Systems and Informatics (SISY).

Ruggiero, A., Quartieri, J., Guarnaccia, C., & Hloch, S. (2015). Noise pollution analysis of wind turbines in rural areas. *International Journal of Environmental Research, 9*(4), 1277-1286.

Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8*(4), e1249.

Schloeder, C., Zimmerman, N., & Jacobs, M. (2001). Comparison of methods for interpolating soil properties using limited data. *Soil science society of America journal, 65*(2), 470-479.

Segaran, V., Tong, Y., Abas, N., Daniel, B. D., Nagapan, S., & Kelundapyan, R. (2020). *Traffic noise assessment among residential environment in batu pahat, johore, Malaysia.* Paper presented at the IOP Conference Series: Materials Science and Engineering.

Servick, K. (2014). Eavesdropping on ecosystems (vol 343, pg 834, 2014). *Science, 343*(6175), 1077-1077.

Sethunga, S., Bodhika, J., & Dharmaratna, W. (2013). Traffic noise contour mapping in Matara city-Sri Lanka. *Traffic*.

Shah, S. K., Tariq, Z., Lee, J., & Lee, Y. (2020). *Real-time machine learning for air quality and environmental noise detection.* Paper presented at the 2020 IEEE International Conference on Big Data (Big Data).

Sharma, A. (2020). Random Forest vs Decision Tree | Which Is Right for You? Retrieved from https://www.analyticsvidhya.com/blog/2020/05/decision-tree-vs-random-forest-algorithm/

Singh, D., Francavilla, A. B., Mancini, S., & Guarnaccia, C. (2021). Application of machine learning to include honking effect in vehicular traffic noise prediction. *Applied Sciences, 11*(13), 6030.

Sugimoto, T., Koyama, K., Kurihara, Y., & Watanabe, K. (2008). *Measurement of infrasound generated by wind turbine generator.* Paper presented at the 2008 SICE Annual Conference.

Syed Muzamil, B., & Dharmendra Singh, R. (2019). Chapter 9 - Survey on Evaluating the Performance of Machine Learning Algorithms: Past Contributions and Future Roadmap. 153-164. doi:https://doi.org/10.1016/B978-0-12-816718-2.00016-6

The International Business Machines Corporation. (2018). K-Nearest Neighbors Algorithm. Retrieved from https://www.ibm.com/topics/knn#:~:text=The%20k%2Dnearest%20neighbors%20algorithm%2C%20also%20known%20as%20KNN%20or,of%20an%20individual%20data%20point.

The International Business Machines Corporation. (2021a). Gains Charts. Retrieved from https://www.ibm.com/docs/en/spss-modeler/saas?topic=gains-charts

The International Business Machines Corporation. (2021b). Modeling Overview. Retrieved from https://www.ibm.com/docs/en/spss-modeler/saas?topic=nodes-modeling-overview

The International Business Machines Corporation. (2021c). Overview of Nodes. Retrieved from https://www.ibm.com/docs/en/spss-modeler/saas?topic=SS3RA7_sub/modeler_mainhelp_client_ddita/clementine/clef_nodes_intro.htm

Tonin, R. (2012). SOURCES OF WIND TURBINE NOISE AND SOUND PROPAGATION. *Acoustics Australia, 40*(1).

Tsai, K.-T., Lin, M.-D., & Chen, Y.-H. (2009). Noise mapping in urban environments: A Taiwan study. *Applied Acoustics, 70*(7), 964-972.

Tunpaiboon, N. (2021). Industry Outlook 2021-2023: Power Generation. Retrieved from https://www.krungsri.com/en/research/industry/industry-outlook/Energy-Utilities/Power-Generation/IO/io-power-generation-21

United States Environmental Protection Agency Office of Noise Abatement and Control. (1974). Information on levels of environmental noise requisite to protect public health and welfare with an adequate margin of safety. Retrieved from https://www.nonoise.org/library/levels74/levels74.htm

United States. Office of Noise Abatement. (1974). *Information on levels of environmental noise requisite to protect public health and welfare with an adequate margin of safety*: US Government Printing Office.

Välisuo, P. O. (2017). *Automated wind turbine noise analysis by machine learning.* Paper presented at the INTER-NOISE and NOISE-CON congress and conference proceedings.

Wen, P.-J., & Huang, C. (2020). Noise prediction using machine learning with measurements analysis. *Applied Sciences, 10*(18), 6619.

Wolff, R. (2022). 10 Best Data Mining Tools in 2022. Retrieved from https://monkeylearn.com/blog/data-mining-tools/

World Health Organization. (2010, 27 April 2010). Noise. Retrieved from https://www.who.int/europe/news-room/fact-sheets/item/noise

World Health Organization. (2022). *Compendium of WHO and other UN guidance on health and environment*. Retrieved from

Wu, Y.-H., & Hung, M.-C. (2016). Comparison of spatial interpolation techniques using visualization and quantitative assessment. *Applications of spatial statistics*, 17-34.

Xiao, Y., Wu, J., Lin, Z., & Zhao, X. (2018). A deep learning-based multi-model ensemble method for cancer prediction. *Computer methods and programs in biomedicine, 153*, 1-9.

Yang, W., He, J., He, C., & Cai, M. (2020). Evaluation of urban traffic noise pollution based on noise maps. *Transportation Research Part D: Transport and Environment, 87*, 102516.

Yilmaz, G., & Hocanli, Y. (2006). Mapping of Noise by Using Gis in ŞAnliurfa. *Environmental Monitoring and Assessment, 121*(1), 103-108. doi:10.1007/s10661-005-9109-1

Zhou, L., Tian, Y., Baidya Roy, S., Dai, Y., & Chen, H. (2013). Diurnal and seasonal variations of wind farm impacts on land surface temperature over western Texas. *Climate Dynamics, 41*(2), 307-326. doi:10.1007/s00382-012-1485-y
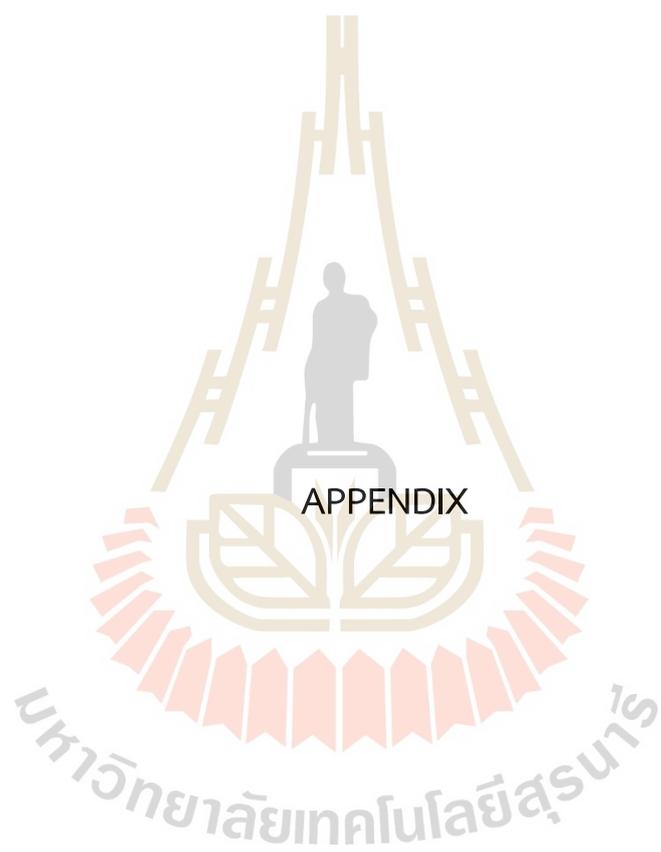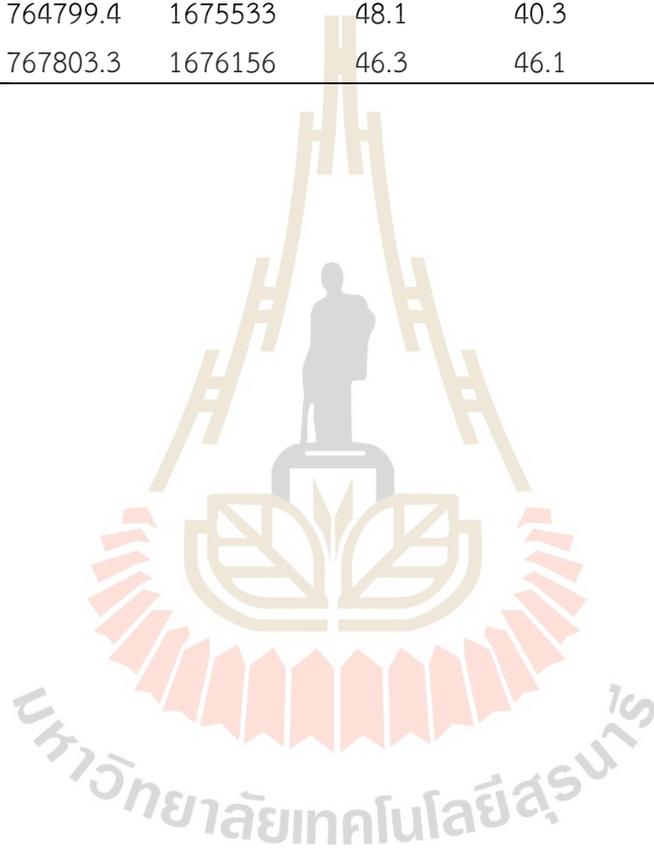
APPENDIX
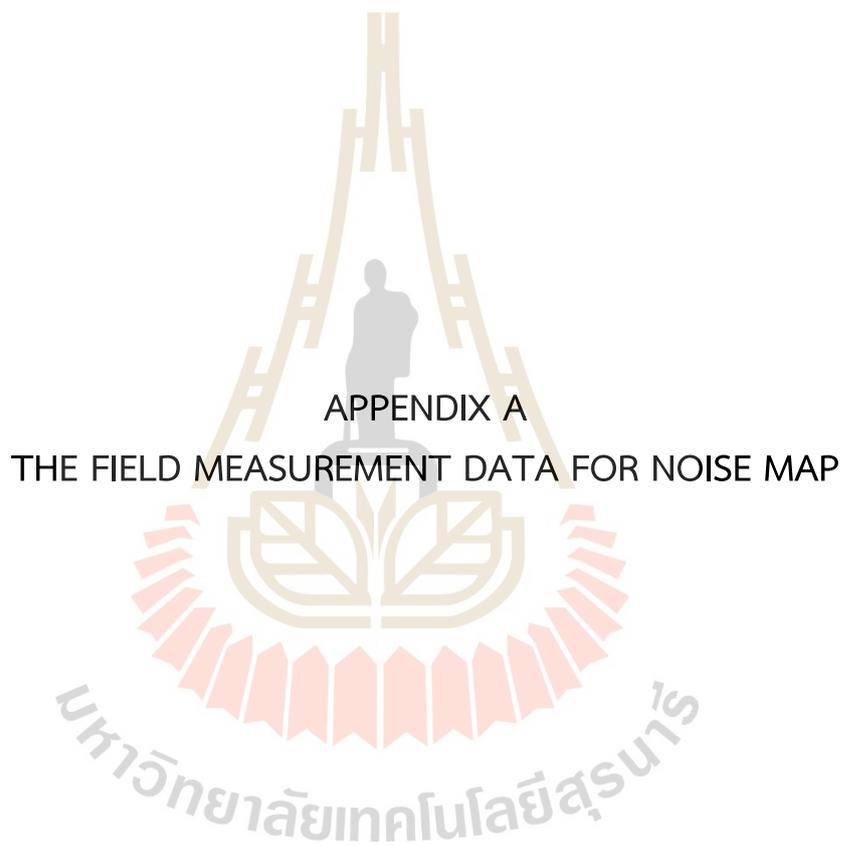
APPENDIX A

THE FIELD MEASUREMENT DATA FOR NOISE MAP

Table A1 Field measurement data of various date

Point	Coordinates		Measurement date			
	X	Y	2018-5	2019-6	2020-9	2021-2
N00	766279	1675847	47.1	52.6	54.9	49.3
N01	766225	1676043	43.7	48.6	46.7	41.5
N02	766117	1676228	45.3	51.6	47.8	41.0
N03	765971	1676408	53.0	58.2	48.3	42.8
N04	765961	1676608	45.4	56.6	46.1	47.0
E00	767822	1674407	50.8	47.6	53.1	52.1
E01	768028.6	1674438	42.9	44.5	42.7	50.5
E02	768234	1674467	44.1	49.8	49.5	49.6
E03	768446	1674478	55.8	55.3	55.4	56.0
E04	768646	1674523	43.2	53.1	52.4	43.6
S00	767383	1673020	58.0	65.8	53.0	52.1
S01	767405	1672824	62.9	51.1	48.7	47.1
S02	767459	1672620	52.3	50.2	49.8	46.7
S03	767583	1672428	53.9	46.3	42	45.6
S04	767658	1672237	43.7	45.7	47.4	44.5
W00	766090	1674335	51.7	50.0	53.7	48.1
W01	765885	1674267	46.3	53.6	40.5	50.3
W02	765702.2	1674226	44.1	49.0	43.2	54.6
W03	765513.2	1674159	56.6	49.6	35.7	45.9
W04	765327	1674085	49.8	48.3	36.2	45.9
NE00	767130.5	1675424	47.2	53	46.8	48.7
NE01	767387.9	1675641	47.2	48.2	40.6	42.7
NE02	767500.9	1675813	50.2	46.2	50.3	43.4
NE03	767684	1675983	51	58.8	65.3	53.3
NE04	767979	1673746	56.2	60.8	52.3	51.6
SE00	768179.1	1673762	43.2	47.5	48.9	48.6
SE01	768377	1673750	45.2	50.5	49.0	47.7
SE02	768584.3	1673716	43.7	47.8	46.8	44.4
SE03	768785	1673656	48.2	54.9	47.6	43.5
SE04	766547	1673914	48.6	56.1	49.5	44.6
SW00	766331.1	1673842	50.4	50.1	45.7	47.0
SW01	766152	1673742	44.5	50.4	42.6	44.4
SW02	765984.9	1673616	47.7	48.8	40.5	44.5

Table A1 (Continued)

Point	Coordinates		Measurement date			
	X	Y	2018-5	2019-6	2020-9	2021-2
SW03	765809.3	1673505	45.2	44.5	38.1	44.7
SW04	765601	1675458	43.9	50.1	39.8	46.1
NW00	765397	1675464	43.3	51.9	52.6	47.8
NW01	765197	1675487	42.6	51.2	44.6	42.0
NW02	764998.8	1675516	43.0	51.3	37.5	43.7
NW03	764799.4	1675533	48.1	40.3	41.4	43.6
NW04	767803.3	1676156	46.3	46.1	40.0	45.4
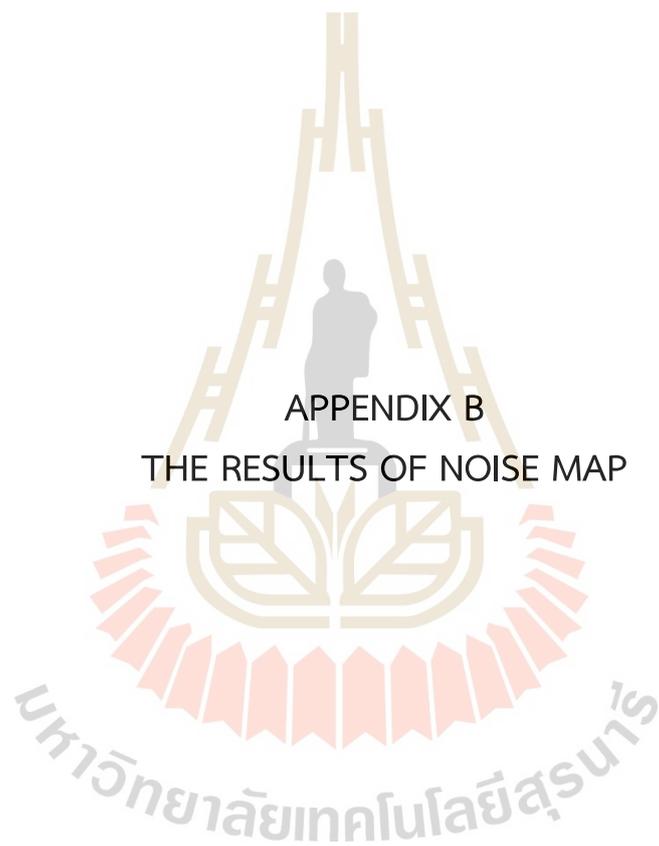
APPENDIX B
THE RESULTS OF NOISE MAP
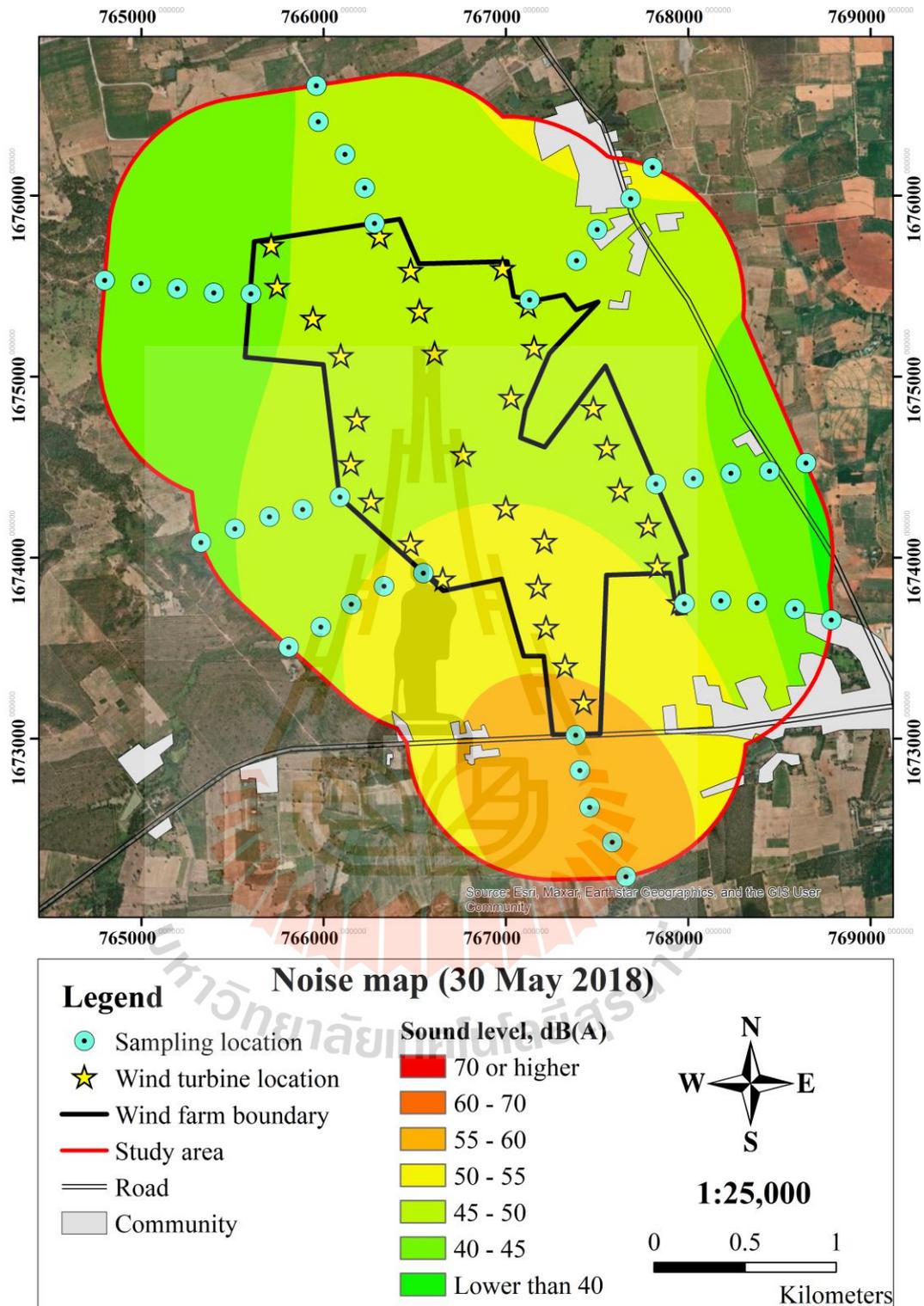
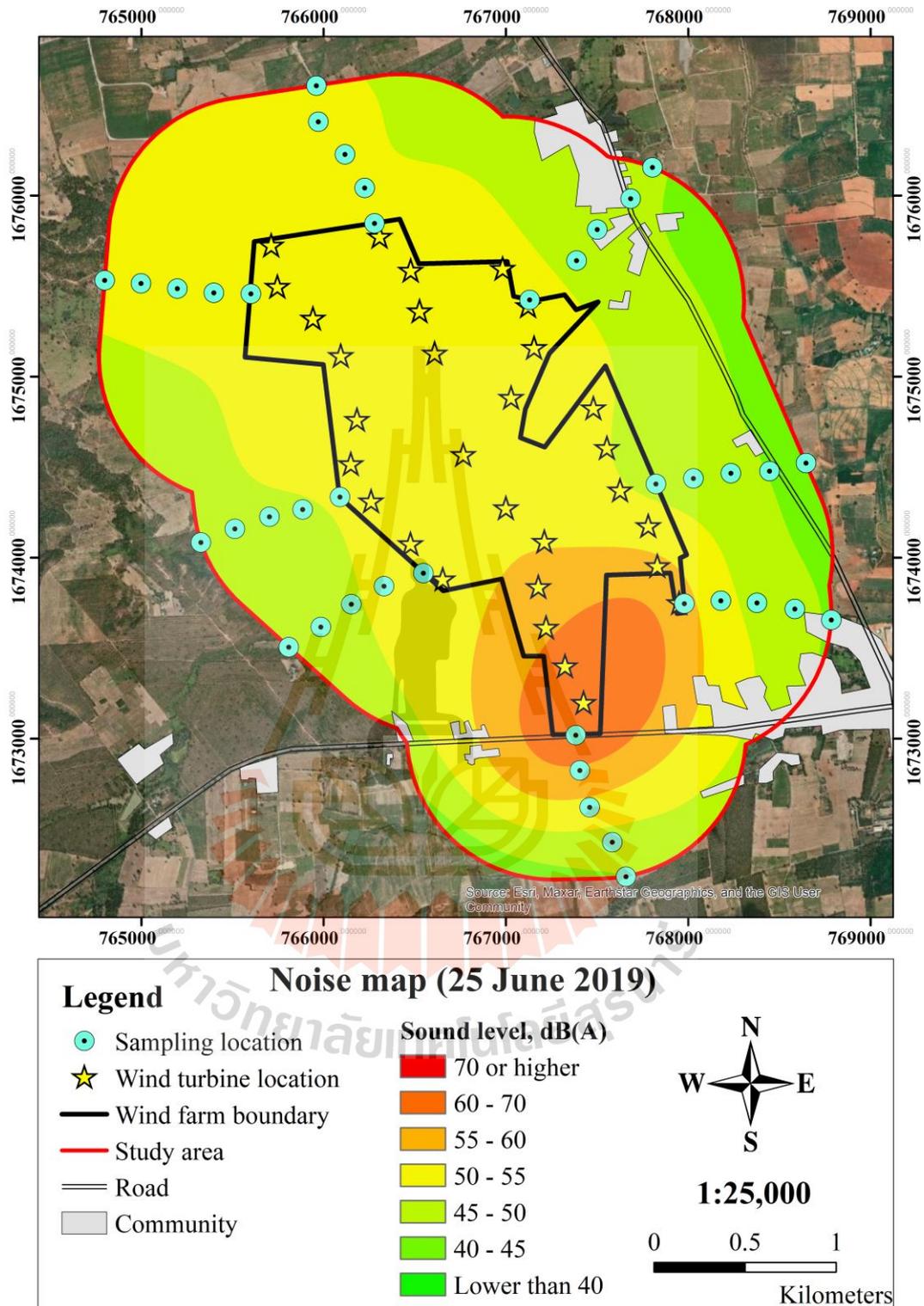Figure B1 Noise map at May 30, 2018
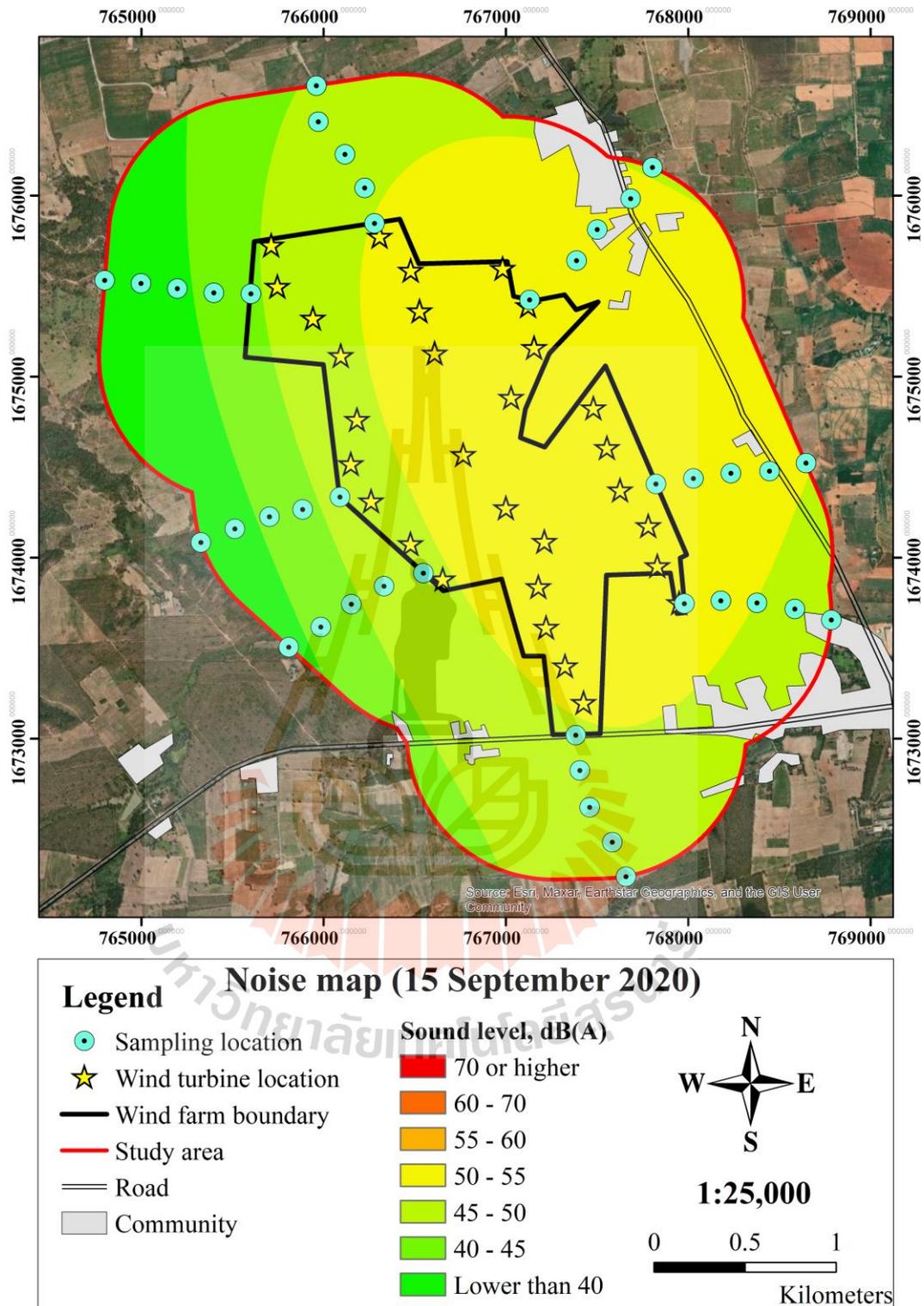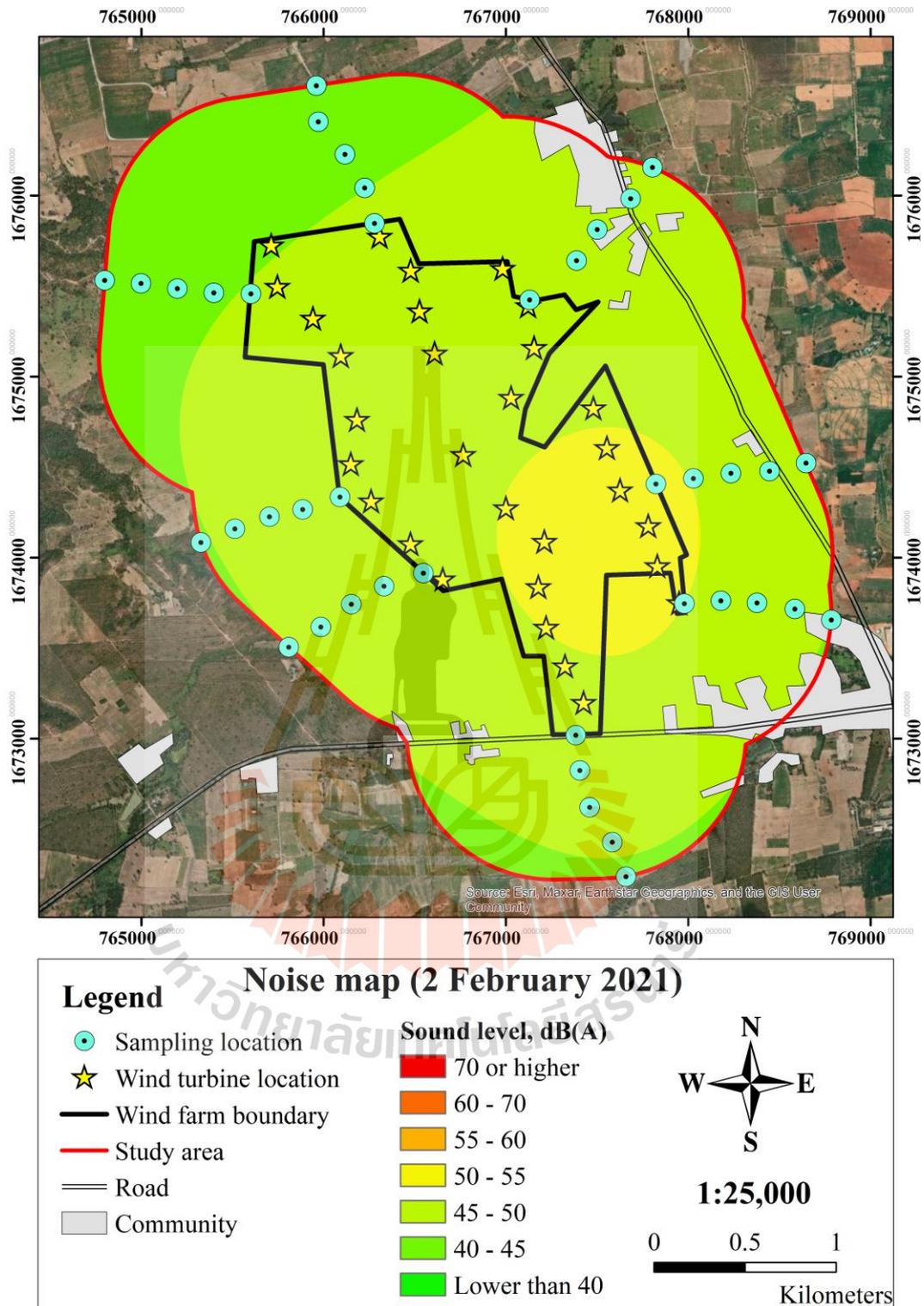
Figure B2 Noise map at June 25, 2019

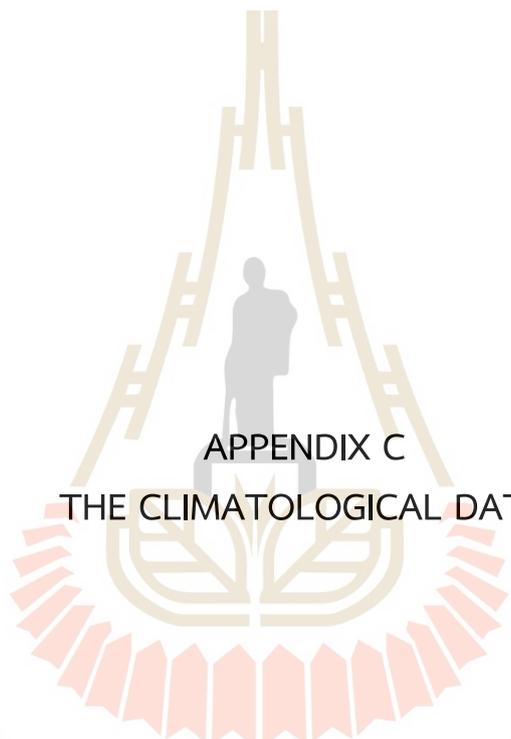Figure B3 Noise map at September 15,2020

Figure B4 Noise map at February 2, 2021

APPENDIX C
THE CLIMATOLOGICAL DATA

Table C1 Climatological data for the period 1990-2019 at Nakhon Ratchasima

Elements		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	Annual
Pressure (hPa)	Mean	1,013	1,011	1,009	1,008	1,007	1,006	1,005	1,006	1,007	1,010	1,012	1,013	1,009
	Mean Daily Range	5.80	6.20	6.00	5.60	4.80	4.20	4.10	4.30	4.70	4.90	5.10	5.50	5.10
	Ext.Max.	1,026	1,024	1,028	1,020	1,017	1,012	1,013	1,012	1,016	1,020	1,022	1,026	1,028
	Ext.Min.	1,003	1,002	999	998	998	998	997	998	998	1,000	1,002	1,001	997
Temperature (Celsius)	Mean Max.	30.9	33.5	35.6	36.7	35.3	34.7	33.9	33.4	32.3	31.4	30.8	29.9	33.2
	Ext.Max.	37.7	39.4	41.5	43.2	41.8	40.3	38.5	37.7	36.2	36.1	36.3	36.0	43.2
	Mean Min.	19.1	21.1	23.6	25.1	25.4	25.4	25.0	24.8	24.3	23.6	21.5	19.2	23.2
	Ext.Min.	10.8	12.4	14.8	17.8	21.8	22.0	22.2	21.7	20.8	16.7	13.7	8.3	8.3
	Mean	24.7	27.0	29.1	30.1	29.4	29.3	28.8	28.3	27.6	27.1	26.0	24.4	27.7
Relative Humidity (%)	Mean	66	62	63	66	73	73	74	76	81	78	71	66	70.8
	Mean Max.	85	83	83	85	88	87	88	90	93	92	88	85	87.3
	Mean Min.	44	41	41	45	53	54	55	58	63	60	53	47	51.0
	Ext.Min.	20	15	15	19	28	32	33	37	36	32	26	21	15.0
Visibility (Km.)	Mean	7.8	7.4	7.8	8.7	9.6	10.1	10.0	9.9	9.6	8.8	8.9	8.6	8.9
	07.00LST	6.4	6.2	7.0	8.1	9.0	9.7	9.5	9.4	8.9	7.8	7.9	7.5	8.1
Cloud Amount (1-10)	Mean	3.8	3.9	4.9	5.6	6.9	7.5	8.0	8.3	7.9	6.3	4.6	3.9	6.0

Table C1 (Continued)

Elements		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	Annual
Wind (Knots)	Prev.Wind	NE	NE	E, S	SW	SW	SW	W	W	W	NE	NE	NE	-
	Mean	1.8	1.8	1.8	1.8	1.9	2.4	2.5	2.3	1.7	2.1	2.5	2.4	2.1
	Max.	21.0	21.0	26.0	45.0	42.0	25.0	30.0	24.0	24.0	26.0	24.0	28.0	45.0
Pan Evaporation (mm.)	Total	133.7	133.8	172.5	170.4	162.1	154.2	155.4	144.1	117.0	124.1	124.6	135.3	1,727
Rainfall (mm)	Total	9.3	13.2	47.4	76.8	147.1	112.2	127.6	173.2	230.3	133.0	19.7	2.8	1,092
	Num. of Days	1.8	2.5	5.9	8.7	14.4	13.8	15.0	18.2	18.7	11.2	3.7	1.4	115.3
	Daily Max.	37.0	59.8	93.2	92.0	89.4	68.7	116.3	121.3	129.7	116.3	61.3	23.7	129.7
Phenomena (Days)	Fog	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.1	0.0	0.4
	Haze	22.9	24.6	24.9	17.7	4.7	2.1	1.2	1.0	1.6	9.6	13.0	18.2	141.5
	Hail	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
	Thunder storm	0.6	1.1	3.6	8.0	10.1	6.2	5.4	6.9	8.4	4.8	0.6	0.0	55.7
	Squall	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.2
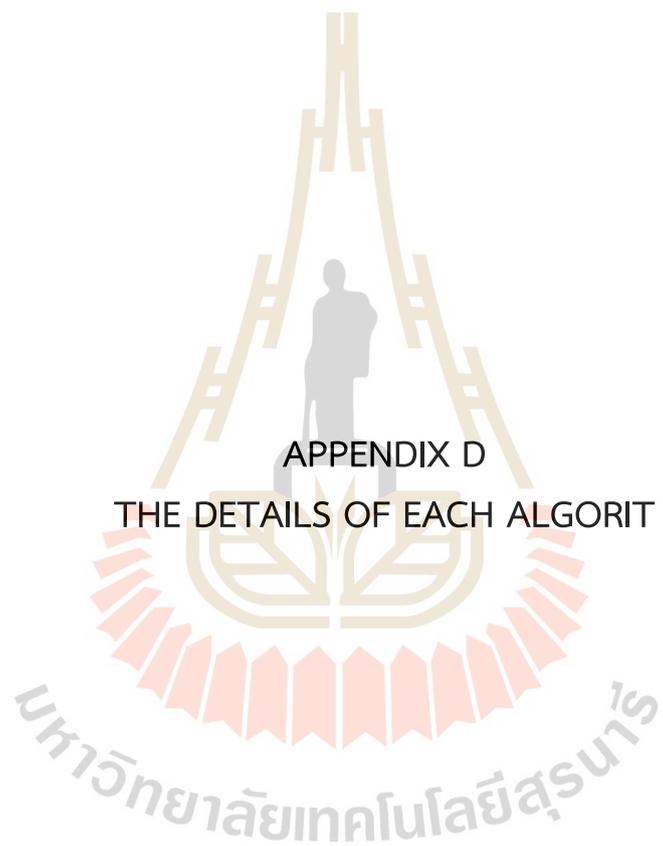
APPENDIX D

THE DETAILS OF EACH ALGORIT

Figure D1 The decision tree of CHIAD

Figure D2 The decision tree of CART

Model Summary

Target	Leq dB(A)
Automatic Data Preparation	On
Model Selection Method	Forward Stepwise
Information Criterion	609.536

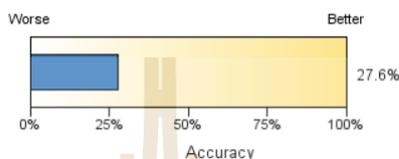The information criterion is used to compare to models. Models with smaller information criterion values fit better.

Worse — Better — 27.6%
Accuracy
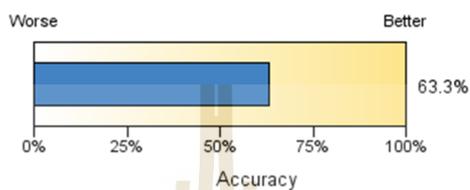
Figure D3 The model summary of Linear

Predicted by Observed
Target: Leq dB(A)
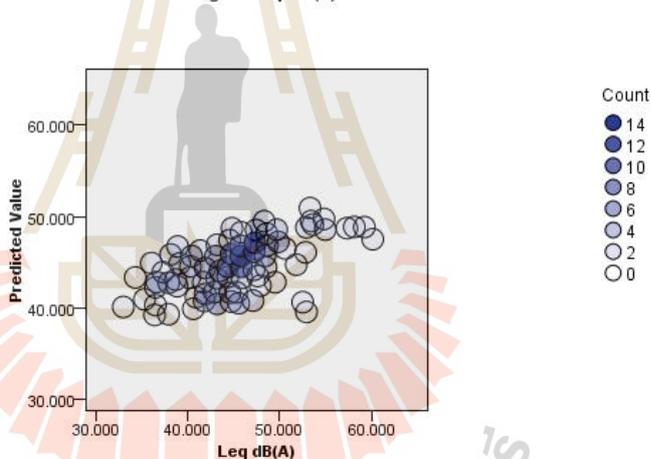
Figure D4 The scatterplot of the predicted values of Linear

Residuals
Target: Leq dB(A)

Mean = -0.001
Std.Dev. = 1.002
N = 220
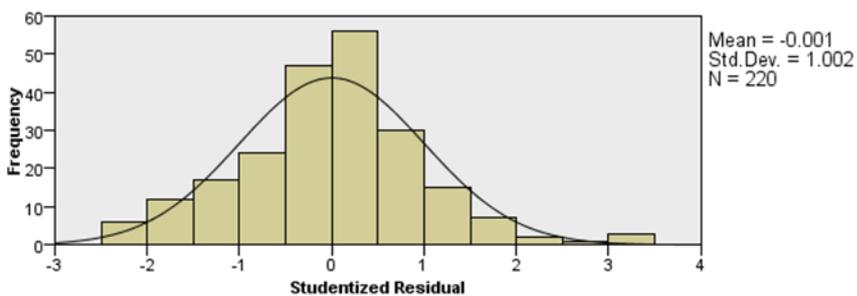
The histogram of Studentized residuals compares the distribution of the residuals to a normal distribution. The smooth line represents the normal distribution. The closer the frequencies of the residuals are to this line, the closer the distribution of the residuals is to the normal distribution.

Figure D5 The binned histogram of the studentized residuals of Linear

Model Summary

Target	Leq dB(A)
Model	Multilayer Perceptron
Stopping Rule Used	Error cannot be further decreased
Hidden Layer 1 Neurons	5

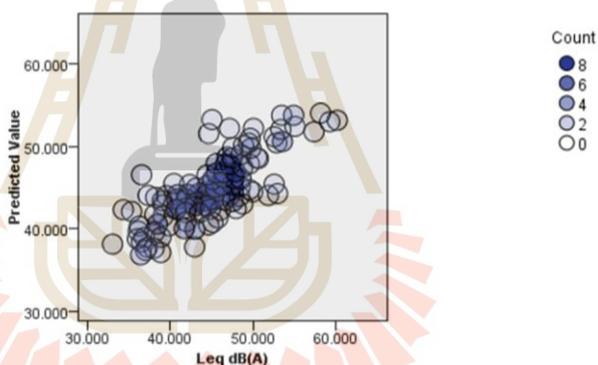Figure D6 The model summary of Neural Network

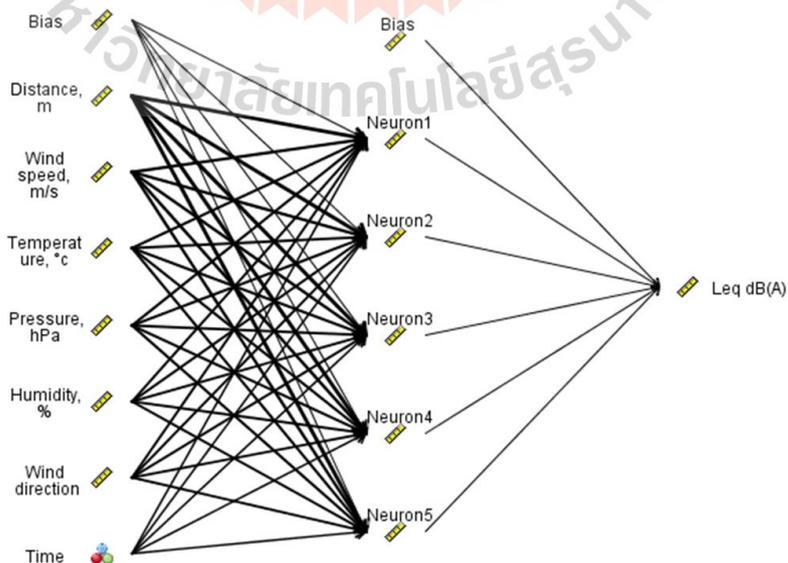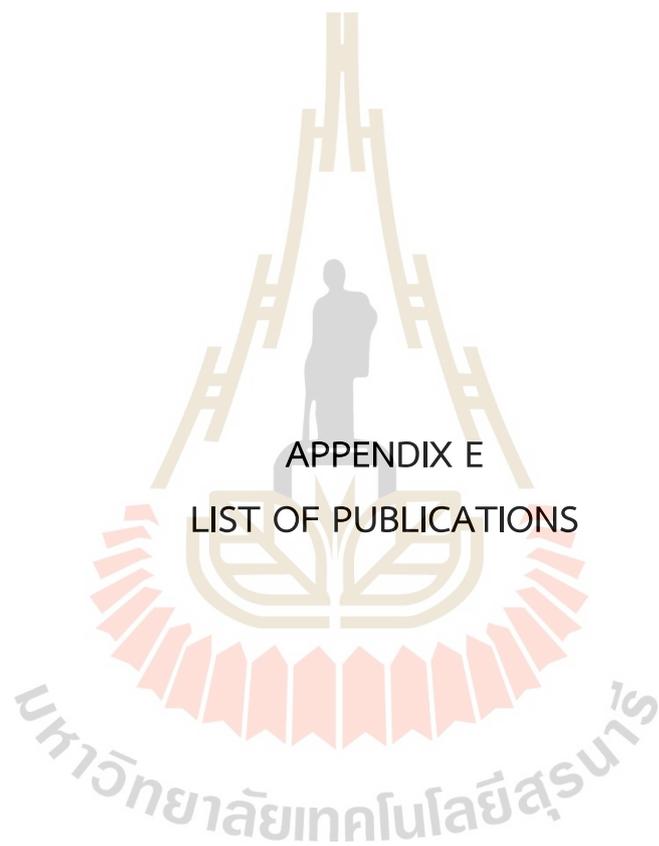Figure D7 The scatterplot of the predicted values of Neural Network

Figure D8 The network structure of Neural Network

APPENDIX E

LIST OF PUBLICATIONS

LIST OF PUBLICATIONS

Charoentangprasert, N., Tantamsapya, T., Yossapol, C., (2023) Prediction of Wind Turbine Noise using SPSS Modeler. KKU Research Journal (Graduate Studies), Volume 23rd, Issue: 4th, October to December 2023.
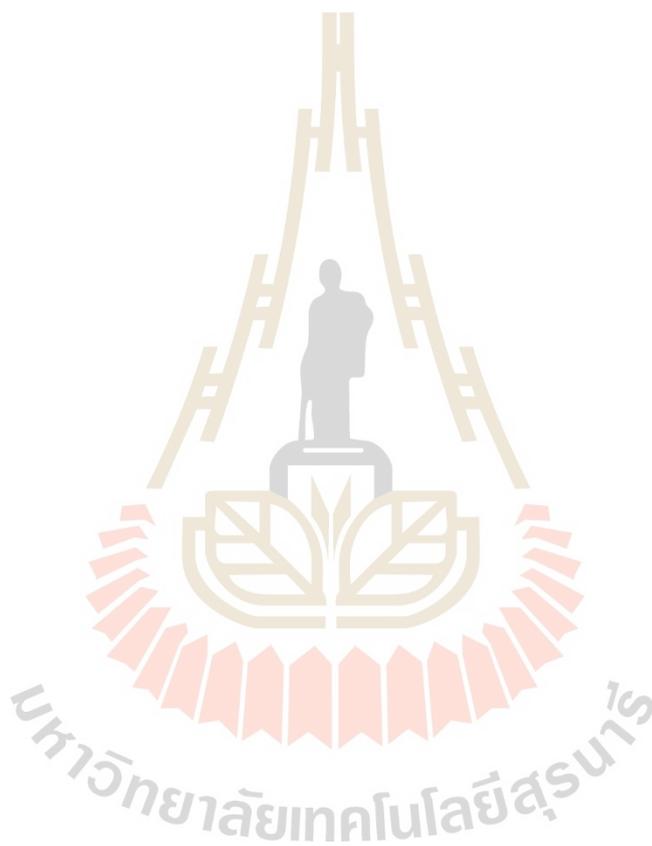
BIOGRAPHY

Mr. Nattapat Charoentangprasert, born on October 9, 1995, in Nakhon Ratchasima Province, Thailand. After graduated with the Science-Mathematics Program at Triamudomsuksanomkao Nakhon Ratchasima School in 2014, he graduated with First Class Honors in Environmental Engineering from Suranaree University of Technology and earned the SUT Kittibandit Scholarship in 2018. During his studies, he presented at the SUT International Virtual Conference on Science and Technology 2020 (IVCST 2020) with the article titled "Spatial Distribution of Noise Around Wind Farm in Huai Bong sub-district, Dankhontod district, Nakhon Ratchasima province, Thailand". He also published the article "Prediction of Wind Turbine Noise using SPSS Modeler" in the KKU Research Journal (Graduate Studies), Volume 23rd, Issue 4, from October to December 2023. He worked as an Assistant Environmental Engineer and Project Engineer at the Suranaree Environmental Technology Research & Consulting Unit from 2018 to 2022, making contributions to various environmental engineering projects.