*Research Article*

# A Fuzzy Soft Model for Haze Pollution Management in Northern Thailand

**Parkpoom Phetpradap** [1,2]

[1]*Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand*
[2]*Centre of Excellence in Mathematics, CHE, Si Ayutthaya Rd., Bangkok 10400, Thailand*

Correspondence should be addressed to Parkpoom Phetpradap; parkpoom.phetpradap@cmu.ac.th

In this article, we propose fuzzy soft models for decision making in the haze pollution management. The main aims of this research are (i) to provide a haze warning system based on real-time atmospheric data and (ii) to identify the most hazardous location of the study area. PM10 is used as the severity index of the problem. The efficiency of the model is justified by the prediction accuracy ratio based on the real data from 1st January 2016 to 31st May 2016. The fuzzy soft theory is modified in order to make models more suitable for the problems. The results show that our fuzzy models improve the prediction accuracy ratio compared to the prediction based on PM10 density only. This work illustrates a fuzzy analysis that has the capability to simulate the unknown relations between a set of atmospheric and environmental parameters. The study area covers eight provinces in the northern region of Thailand, where the problem severely occurs every year during the dry season. Seven principle parameters are considered in the model, which are PM10 density, air pressure, relative humidity, wind speed, rainfall, temperature, and topography.

## 1. Introduction

Pollution problems are inevitably a global concern of the 21st century. Over the past decade, polluted haze has become a major problem in the northern region of Thailand and surrounded countries. In March 2019, the problem reached a crisis when the daily average PM2.5 and PM10 (particulate matter of 2.5 microns and 10 microns in diameter or smaller) density rates were well beyond the national standard of 25 $\mu$g/m$^3$ and 50 $\mu$g/m$^3$ for several days according to local environmental data sources such as Pollution Control Department [1], Climate Change Data Centre of Chiang Mai University [2], and Smoke Haze Integrated Research Unit [3]. This situation has occurred every year on dry season, from January to May, and generally reached its peak in March. During this period, a large amount of particulate matters are released into the atmosphere, including carbon monoxide, carbon dioxide, volatile organic compounds, and carcinogenic polycyclic aromatic hydrocarbons [4]. The main emission source is biomass open burning, such as forest fires, solid waste burning, and agricultural residue field burning [5, 6].

This problem has a significant effect on human health, local traveling industry, and the economy as a whole, especially in Chiang Mai province, a popular tourist destination.

The public health ministry of Thailand has reported an increase in bronchial asthma and respiratory diseases in people living in these areas. In addition, these fine particles contain carcinogenic polycyclic aromatic hydrocarbons that can induce lung cancer [7]. The smoke haze episodes also reduce visibility and cause a variety of environmental effects which eventually leads to decline in various economic sectors such as tourism, transportation, and agriculture. Thai government has launched various policies to get the smoke haze problem under control. However, the problem still continues to grow, even with the enforcement of outdoor burning ban issued by Thai government during February to April period.

Apparently, the atmospheric parameters and topography play the key parts of the problem. The air pollutants are

trapped near ground level due to the meteorological conditions (e.g., stagnant air), and the basin-like topography surrounded by high mountain ranges results in restricted pollution dispersion. Moreover, low rainfall in dry season also adds on to the severity of the haze problem. For this reason, the leaching of smoke or dust particles in the air is low [6]. These conditions caused the air pollutants to flow out difficultly and the particle cannot be easily escaped from the area. Notably, there are some technologies that mitigate the pollution problem. However, the costs of devices are considerably expensive.

Undoubtedly, an efficient warning system would become a major help in the haze problem management. The system will significantly improve public safety and mitigate damage caused. The Goddard Earth Observing System Model Version 5 (GEOS-5) is currently one of the widely used pollution prediction models developed by NASA's research team.

In this article, the potential use of fuzzy soft set theory in real-time haze warning is investigated. The main aims of this research are (i) to provide a haze warning system based on real-time atmospheric data and (ii) to identify the most hazardous location of the study area. The benefits are to create the awareness for people in the affected area and to suggest the location to establish pollution mitigation devices. Molodtsov [8–10] initiated the concept of soft set theory as a new mathematical tool for dealing with uncertainties. Soft set theory has rich potential for applications in several directions, a few of which had been shown by Molodtsov [8]. The idea of applying fuzzy soft set theory in atmospheric models is already considered concerning the applications to air pollution management [11–14] and water management [15–20]. However, it is believed that air pollution models may be different for each region due to many several factors [21, 22]. Therefore, existing models still need to be restudied. Up to our knowledge, there are only a few prediction models in the region of study since the main concerns are on the site of environmental science. The prediction results from GEOS-5 model are popular choices to be used as benchmarks for environmental scientist. The regional-developed models include a logistic regression model [23] and Geographic Information System- (GIS-) based model [24]. Therefore, our model would offer an alternative prediction model for the haze pollution problem.

The study location covered eight provinces in the northern part of Thailand where haze problem has severely occurred: Mae Hong Son, Chiang Mai, Lamphun, Chiang Rai, Phayao, Lampang, Phrae, and Nan. The density of PM10 is used as severity index of the haze pollution level. Additionally, seven principle parameters are considered in the model: six are atmospheric parameters—PM10, air pressure, relative humidity, wind speed, rainfall, and temperature—and the other one is the topographic parameter. All atmospheric data are obtained from the Pollution Control Department [1]. The obtained data period is from 1st January 2016 to 31st May 2016.

The rest of this article is organized as follows. In Section 2, we explain the methodology and present some examples. In Section 3, we describe the setup of the model, which includes the study location, the data, and the parameters.

Then, we present our decision-making results and discussion in Section 4. Finally, the conclusion is given in Section 5.

## 2. Methodology

*2.1. Fuzzy Soft Theory.* In this section, we provide useful notations of soft sets and fuzzy soft sets. Let $U = \{L_1, L_2, \ldots, L_m\}$ be an initial universal set and let $E = \{P_1, P_2, \ldots, P_n\}$ be a set of parameters.

*Definition 1* (see [8]). Let $P(U)$ denote the power set of $U$ and $A \subset E$. A pair $(F, A)$ is called a soft set over $U$, where $F$ is a mapping given by $F: A \longrightarrow P(U)$.

*Example 1.* Let the initial universe $U = \{L_1, L_2, \ldots, L_8\}$ be the eight selected provinces in the northern region of Thailand: Mae Hong Son, Chiang Mai, Lamphun, Chiang Rai, Phayao, Lampang, Phrae, and Nan. Moreover, let $E = \{P_1, P_2, P_3, P_4\}$ be atmospheric parameters: PM10 density, air pressure, relative humidity, and wind speed, respectively. Then, an example of possible soft set is

$$
\begin{aligned}
F(P_1) &= \{L_1, L_2, L_3\}, \\
F(P_2) &= \{L_2, L_4, L_7, L_8\}, \\
F(P_3) &= \{L_1, L_5, L_6, L_8\}, \\
F(P_4) &= \{L_4, L_7\}.
\end{aligned}
\tag{1}
$$

Note that each approximation has two parts, predicate $p$ and approximate value set. For example, the predicate is PM10 density and the approximate value set is $\{L_1, L_2, L_3\}$ for $F(P_1)$. Additionally, the summary information of this soft set is represented in Table 1.

*Definition 2* (see [25]). Let $\Psi(U)$ denote the set of all fuzzy sets of $U$ and let $A_i \subset E$. A pair $(F_i, A_i)$ is called a fuzzy soft set over $U$, where $F_i$ is a mapping given by $F_i: A_i \longrightarrow \Psi(U)$.

*Example 2.* We consider the same setup as in Example 1. An example of a fuzzy soft set is

$$
F(P_1) = \left\{ \left(\frac{L_1}{0.7}\right), \left(\frac{L_2}{0.9}\right), \left(\frac{L_3}{1}\right), \left(\frac{L_4}{0.4}\right), \left(\frac{L_5}{0.2}\right), \left(\frac{L_6}{0.3}\right), \left(\frac{L_7}{0.4}\right), \left(\frac{L_8}{0.5}\right) \right\},
$$

$$
F(P_2) = \left\{ \left(\frac{L_1}{0.5}\right), \left(\frac{L_2}{0.8}\right), \left(\frac{L_3}{0.1}\right), \left(\frac{L_4}{1}\right), \left(\frac{L_5}{0.2}\right) L_5/0, \left(\frac{L_6}{0.3}\right), \left(\frac{L_7}{0.6}\right), \left(\frac{L_8}{1}\right) \right\},
$$

$$
F(P_3) = \left\{ \left(\frac{L_1}{0.8}\right), \left(\frac{L_2}{0.2}\right), \left(\frac{L_3}{0.1}\right), \left(\frac{L_4}{0}\right), \left(\frac{L_5}{0.9}\right), \left(\frac{L_6}{0.7}\right), \left(\frac{L_7}{0.2}\right), \left(\frac{L_8}{0.8}\right) \right\},
$$

$$
F(P_4) = \left\{ \left(\frac{L_1}{0.2}\right), \left(\frac{L_2}{0.4}\right), \left(\frac{L_3}{0.4}\right), \left(\frac{L_4}{0.7}\right), \left(\frac{L_5}{1}\right), \left(\frac{L_6}{0.5}\right), \left(\frac{L_7}{0.6}\right), \left(\frac{L_8}{0.4}\right) \right\}.
\tag{2}
$$

Table 2 provides the summary information of this fuzzy soft set.

*Definition 3.* For a given fuzzy soft set with a universal set $U$ and parameter set $P$, we denote $l_{ij}$ as the membership value of $L_i$ in $F(P_j)$.

TABLE 1: Tabular presentation of the soft set in Example 1.

| Label | PM10 density | Air pressure | Humidity | Wind speed |
|---|---|---|---|---|
| $L_1$ | 1 | 0 | 1 | 0 |
| $L_2$ | 1 | 1 | 0 | 0 |
| $L_3$ | 1 | 0 | 0 | 0 |
| $L_4$ | 0 | 1 | 0 | 1 |
| $L_5$ | 0 | 0 | 1 | 0 |
| $L_6$ | 0 | 0 | 1 | 0 |
| $L_7$ | 0 | 1 | 0 | 1 |
| $L_8$ | 0 | 1 | 1 | 0 |

TABLE 2: Tabular presentation of the fuzzy soft set in Example 2.

| Label | PM10 density | Air pressure | Humidity | Wind speed |
|---|---|---|---|---|
| $L_1$ | 0.7 | 0.5 | 0.8 | 0.2 |
| $L_2$ | 0.9 | 0.8 | 0.2 | 0.4 |
| $L_3$ | 1 | 0.1 | 0.1 | 0.4 |
| $L_4$ | 0.4 | 1 | 0 | 0.7 |
| $L_5$ | 0.2 | 0 | 0.9 | 0.1 |
| $L_6$ | 0.3 | 0.3 | 0.7 | 0.5 |
| $L_7$ | 0.4 | 0.6 | 0.2 | 0.6 |
| $L_8$ | 0.5 | 1 | 0.8 | 0.4 |

*Definition 4* (see [25]). For a given fuzzy soft set, *the choice value* of $L_i$ is defined by

$$c_i = \sum_{j=1}^{n} l_{ij}. \tag{3}$$

*Definition 5* (see [25]). For a given fuzzy soft set, *the comparison table* is the $n \times n$ table, in which the entry $e_{ij}$ is the number of parameters for which the membership value of $L_i$ exceeds or equals the membership value of $L_j$. Both row and column of the table are labelled by the elements of the universal set.

*Remark 1*

(1) Each main diagonal element of a comparison table is always equal to $n$.

(2) $0 \le e_{ij} \le n$ for all $i, j$

*Definition 6* (see [25])

(i) *Impact indicator* of $L_i$ is the sum of all values on row $i$ on the comparison table. This can be calculated by the following formula:

$$I_i = \sum_{j=1}^{n} e_{ij}. \tag{4}$$

(ii) *Divider indicator* of $L_j$ is the sum of all values on column $j$ on the comparison table. This can be calculated by the following formula:

$$D_i = \sum_{i=1}^{n} e_{ij}. \tag{5}$$

(iii) The *score value* of $L_i$ is defined as

$$c_i = I_i - D_i. \tag{6}$$

Both values can be used as evaluations in a decision making. However, according to Kong et al. [26], it is possible that these values may lead to different decision results. Therefore, they introduced grey relational grade, a new evaluation indicator that combines both information of score values and choice values, to make the decision making more robust. The calculation algorithm of grey relational grade is briefly presented.

*Algorithm 1* (see [26]). Decision making based on grey relational grade.

(1) Input the choice value sequence $(c_1, c_2, \ldots, c_m)$ and the score sequence $(s_1, s_2, \ldots, s_m)$ where $c_i$ and $s_i$ are the choice value and the score value of $L_i$, respectively.

(2) Calculate grey relational generating values:

$$c'_i = \frac{c_i - \min_{i}\{c_i\}}{\max_{i}\{c_i\} - \min_{i}\{c_i\}},$$

$$s'_i = \frac{s_i - \min_{i}\{s_i\}}{\max_{i}\{s_i\} - \min_{i}\{s\}}. \tag{7}$$

(3) Calculate grey difference information:

$$c'_{max} = \max_{i}\{c'_i\},$$
$$\Delta c'_i = \left| c'_{max} - c'_i \right|,$$
$$s'_{max} = \max_{i}\{s'_i\},$$
$$\Delta s'_i = \left| s'_{max} - s'_i \right|,$$
$$\Delta_{max} = \max_{i}\{\Delta c'_i, \Delta s'_i\},$$
$$\Delta_{min} = \min_{i}\{\Delta c'_i, \Delta s'_i\}. \tag{8}$$

(4) Calculate grey relative coefficients:

$$\gamma_c(c_i) = \frac{\Delta_{min} + 0.5\Delta_{max}}{\Delta c'_i + 0.5\Delta_{max}},$$

$$\gamma_s(s_i) = \frac{\Delta_{min} + 0.5\Delta_{max}}{\Delta s'_i + 0.5\Delta_{max}}. \tag{9}$$

(5) Calculate grey relational grade:

$$\gamma(L_i) = 0.5\gamma_c(c_i) + 0.5\gamma_s(s_i). \tag{10}$$

(6) The decision is $L_k$ if $\gamma(L_k) = \max_i\{\gamma(L_i)\}$. Optimal choices may have more than one if there are more than one element corresponding to the maximum.

Decision making based on score values and choice values relies on the assumption that the parameters are equally important. However, in some decision-making problems, some parameters can be more important than the others. Therefore, we propose new definitions of choice values and score values based on weight information. Note that idea of weighted score value is briefly discussed in Maji et al. [9].

Define a weight $w := (w_1, w_2, \ldots, w_n)$ as weight sequence of parameters where $w_i$ is the weight associated with the parameter $P_i$.

*Definition 7.* For a given fuzzy soft set and a weight $w$, *the weighted choice value* of $L_i$ is defined by

$$c_{\mathbf{w},i} = \sum_{j=1}^{n} w_j l_{ij}. \tag{11}$$

*Definition 8.* For a given fuzzy soft set and a weight $w$, *the weighted comparison table* is the $n \times n$ table, in which the entry $e_{ij}$ is calculated by the following formula:

$$e_{ij} = \sum_{i=1}^{n} w_j 1\left(l_{ik} \geq l_{jk}\right), \tag{12}$$

where $1(\cdot)$ is an indicator function defined by

$$1(A) = \begin{cases} 1, & \text{if the statement A is true,} \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

In other words, this is the weighted sum of parameters which the membership value of $L_i$ exceeds or equals the membership value of $L_j$. Both row and column of the table are labelled by the elements of the universal set.

*Remark 2*

(1) Each main diagonal element of a weighted comparison table is always equal to the sum $\sum_k w_k$.

(2) $0 \leq e_{ij} \leq \sum_k w_k$ for all $i, j$.

*Definition 9*

(i) A *weighted impact indicator* is the impact indicator that is calculated based on the weighted comparison table associated with the weight $w$.

(ii) A *weighted divider indicator* is the divider indicator that is calculated based on the weighted comparison table associated with the weight $w$.

(iii) The *weighted score values* of $L_i$ are the score values that are calculated based on the weighted impact

indicator and weighted divider indicator associated with the weight $w$.

Apparently, if $w = (1, 1, 1, \ldots, 1)$, the weighted choice values and the weight score values are equal to the choice values and the score values defined in Definitions 4 and 6.

*Example 3.* In a decision-making problem with $U = \{L_1, L_2, \ldots, L_8\}$, $E = \{P_1, P_2, \ldots, P_8\}$, define the weight $w = (3, 2, 1, 2, 2, 1, 1, 2)$. Suppose that a fuzzy soft set is

$$F(P_1) = \left\{\left(\frac{L_1}{0.7}\right), \left(\frac{L_2}{0.9}\right), \left(\frac{L_3}{1}\right), \left(\frac{L_4}{0.4}\right), \left(\frac{L_5}{0.2}\right), \left(\frac{L_6}{0.3}\right), \left(\frac{L_7}{0.4}\right), \left(\frac{L_8}{0.5}\right)\right\},$$

$$F(P_2) = \left\{\left(\frac{L_1}{0.5}\right), \left(\frac{L_2}{0.8}\right), \left(\frac{L_3}{0.1}\right), \left(\frac{L_4}{1}\right), \left(\frac{L_5}{0}\right), \left(\frac{L_6}{0.3}\right), \left(\frac{L_7}{0.6}\right), \left(\frac{L_8}{1}\right)\right\},$$

$$F(P_3) = \left\{\left(\frac{L_1}{0.8}\right), \left(\frac{L_2}{0.2}\right), \left(\frac{L_3}{0.1}\right), \left(\frac{L_4}{0}\right), \left(\frac{L_5}{0.9}\right), \left(\frac{L_6}{0.7}\right), \left(\frac{L_7}{0.2}\right), \left(\frac{L_8}{0.8}\right)\right\},$$

$$F(P_4) = \left\{\left(\frac{L_1}{0.2}\right), \left(\frac{L_2}{0.4}\right), \left(\frac{L_3}{0.4}\right), \left(\frac{L_4}{0.7}\right), \left(\frac{L_5}{1}\right), \left(\frac{L_6}{0.5}\right), \left(\frac{L_7}{0.6}\right), \left(\frac{L_8}{0.4}\right)\right\},$$

$$F(P_5) = \left\{\left(\frac{L_1}{0.4}\right), \left(\frac{L_2}{0.2}\right), \left(\frac{L_3}{0.8}\right), \left(\frac{L_4}{0.9}\right), \left(\frac{L_5}{0.2}\right), \left(\frac{L_6}{0.4}\right), \left(\frac{L_7}{0.7}\right), \left(\frac{L_8}{0.7}\right)\right\},$$

$$F(P_6) = \left\{\left(\frac{L_1}{0.5}\right), \left(\frac{L_2}{0.1}\right), \left(\frac{L_3}{0.1}\right), \left(\frac{L_4}{0.3}\right), \left(\frac{L_5}{0.5}\right), \left(\frac{L_6}{0.7}\right), \left(\frac{L_7}{1}\right), \left(\frac{L_8}{0.5}\right)\right\},$$

$$F(P_7) = \left\{\left(\frac{L_1}{0.5}\right), \left(\frac{L_2}{0.5}\right), \left(\frac{L_3}{0.4}\right), \left(\frac{L_4}{0.8}\right), \left(\frac{L_5}{0.2}\right), \left(\frac{L_6}{0.4}\right), \left(\frac{L_7}{0.2}\right), \left(\frac{L_8}{0.7}\right)\right\},$$

$$F(P_8) = \left\{\left(\frac{L_1}{0.75}\right), \left(\frac{L_2}{1}\right), \left(\frac{L_3}{1}\right), \left(\frac{L_4}{0.75}\right), \left(\frac{L_5}{0.75}\right), \left(\frac{L_6}{0.5}\right), \left(\frac{L_7}{0.75}\right), \left(\frac{L_8}{0.75}\right)\right\}. \tag{14}$$

Our aim is to choose the optimum choice according to the weight $w$. By Definition 7, the weighted choice value sequence is $(7.6, 8.3, 8.2, 9, 4.3, 6.1, 7.9, 8.7)$. Next, we calculate weighted score value. By Definition 8, the weighted comparison table is shown in Table 3. Then, by Definition 9, the weighted impact indicator, the weighted divider indicator, and the weighted score values are shown in Table 4. The weighted score value sequence is $(-1, 13, 18, 31, -60, -29, 9, 19)$. Finally, we make a decision based on grey relational grade. The calculation of Algorithm 1 is shown in Table 5. Therefore, $L_4$ is the optimal choice.

*2.2. Particle Swarm Optimization.* The particle swarm optimization (PSO) algorithm is a metaheuristic algorithm based on the concept of swarm intelligence. The algorithm was proposed in 1995 by Kennedy and Eberhart [27]. PSO is metaheuristic as it makes few or no assumptions about the problem being optimized and can search very large spaces of candidate solutions. Also, PSO does not use the gradient of the problem being optimized, which means PSO does not require that the optimization problem be differentiable as is required by classic optimization methods such as gradient descent and quasi-Newton methods. Also, it is capable of solving complex mathematics problems existing in engineering [28].

TABLE 3: The weighted comparison table in Example 3.

| | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_5$ | $L_6$ | $L_7$ | $L_8$ |
|---|---|---|---|---|---|---|---|---|
| $L_1$ | 14 | 5 | 5 | 7 | 13 | 11 | 7 | 6 |
| $L_2$ | 10 | 14 | 9 | 6 | 12 | 8 | 9 | 7 |
| $L_3$ | 9 | 10 | 14 | 6 | 12 | 8 | 8 | 9 |
| $L_4$ | 9 | 8 | 8 | 14 | 12 | 12 | 12 | 9 |
| $L_5$ | 4 | 4 | 2 | 4 | 14 | 3 | 4 | 3 |
| $L_6$ | 5 | 6 | 7 | 2 | 11 | 14 | 2 | 4 |
| $L_7$ | 9 | 6 | 6 | 7 | 13 | 12 | 14 | 7 |
| $L_8$ | 9 | 9 | 7 | 7 | 11 | 12 | 9 | 14 |

TABLE 4: The weighted impact indicator, the weighted divider indicator, and the weighted score values of each choice in Example 3.

| | $I_{w,i}$ | $D_{w,i}$ | $s_{w,i}$ |
|---|---|---|---|
| $L_1$ | 68 | 69 | −1 |
| $L_2$ | 75 | 62 | 13 |
| $L_3$ | 76 | 58 | 18 |
| $L_4$ | 84 | 53 | 31 |
| $L_5$ | 38 | 98 | −60 |
| $L_6$ | 51 | 80 | −29 |
| $L_7$ | 74 | 65 | 9 |
| $L_8$ | 78 | 59 | 19 |

TABLE 5: The calculation of grey relational grade based on the weighted choice value sequence and the weighted score value sequence in Example 3.

| | $c'_{w,i}$ | $s'_{w,i}$ | $\Delta c_{w,i}$ | $\Delta s_{w,i}$ | $\gamma_{c_w}(c_i)$ | $\gamma_{s_w}(s)$ | $\gamma_w(L_i)$ |
|---|---|---|---|---|---|---|---|
| $L_1$ | 0.702 | 0.648 | 0.298 | 0.352 | 0.627 | 0.587 | 0.607 |
| $L_2$ | 0.851 | 0.802 | 0.149 | 0.198 | 0.770 | 0.717 | 0.744 |
| $L_3$ | 0.830 | 0.857 | 0.170 | 0.143 | 0.746 | 0.778 | 0.762 |
| $L_4$ | 1 | 1 | 0 | 0 | 1 | 1 | 1 |
| $L_5$ | 0 | 0 | 1 | 1 | 0.333 | 0.333 | 0.333 |
| $L_6$ | 0.383 | 0.341 | 0.617 | 0.659 | 0.448 | 0.431 | 0.439 |
| $L_7$ | 0.766 | 0.758 | 0.234 | 0.242 | 0.681 | 0.674 | 0.678 |
| $L_8$ | 0.936 | 0.868 | 0.064 | 0.132 | 0.887 | 0.791 | 0.839 |

This method is now available to use in computer packages such as Matlab or R.

## 3. Model Construction

*3.1. The Study Area.* Our study area is in the northern region of Thailand, the haze pollution affected area. The region, approximately 94,000 km$^2$ in size and six million in population, consists of nine provinces: Mae Hong Son, Chiang Mai, Lamphun, Chiang Rai, Phayao, Lampang, Phrae, Nan, and Uttaradit. For this case study, Uttaradit was excluded since its haze problem was not severe. The study area is geographically characterised by several mountain ranges, which continue from the Shan Hills in bordering Myanmar to Laos, and the river valleys which cut through them. The basins of rivers Ping, Wang, Yom, and Nan run from north to south. The basins cut across the mountains of two great ranges, the Thanon Thong Chai Range in the west and the

Phi Pan Nam in the east. All studied provinces lie between these basins. The elevations are generally moderate, a little above 2,000 metres (6,600 ft) for the highest summit. Table 6 provides the geographic information summary of each province. The latitudes and longitudes shown are the locations of meteorology stations where atmospheric data are collected. The basin sizes are divided into five categories: no basin, wide, normal, moderate, and narrow, and we set the airflow difficulty level of each category to be 0, 1, 2, 3, and 4, respectively. The narrow basin implies that the flow of the air is more difficult. The location map of study area is shown in Figure 1.

*3.2. The Data.* The hourly atmospheric data of PM10 density ($\mu$g/m$^3$, at 3 m from ground), air pressure (mmHg, at 2 m), relative humidity (%RH, at 2 m), wind speed (m/s at 30 m), rainfall (mm at 3 m), and temperature (°C at 2 m) from 1$^{st}$ January 2016 to 31$^{st}$ May 2016 were obtained with authorization from the Pollution Control Department [1]. About 3% of data was missing from the record. The missing data were replaced by the same data at the preceding time. Figure 2 represents the daily fluctuation of PM10 density of the eight selected locations during the study period. Table 7 represents the summary statistics of PM10 density of the eight selected locations.

*3.3. The Parameters.* Based on environmental research studies [30–33], the climate and the topography of the study area play significant roles in the pollution problem. Therefore, the parameter set consists of seven parameters in this application, which are PM10 density, air pressure, relative humidity, wind speed, rainfall, temperature, and airflow difficulty level. The first six parameters are atmospheric parameters, while the last parameter is topographic parameter. Additionally, the effects of each atmospheric components on the PM10 density, the severity index, can be categorized into two types; positive and negative. A positive atmospheric component is the component such that increasing in its value will lead to the increase of the PM10 density, while a negative atmospheric component is the component such that increasing in its value will lead to the decrease of the PM10 density. The parameter information is summarised in Table 8.

## 4. Results and Discussion

*4.1. Haze Warning System.* The first aim of this research is to create a warning system based on real-time atmospheric data. The system predicts whether the PM10 density will exceed the crisis level or not in the following 4 hours. Note that the length of warning period can be adjusted. In this article, we choose the period of 4 hours since the period of time is reasonable enough to do some safety mitigation such as buying protection masks, completing necessary outdoor activities, or evacuating to public designated safe zones. The warnings will be set to be announced at 12 a.m., 4 a.m., 8 p.m., 4 p.m., 8 p.m., and 12 p.m. each day. The PM10 crisis

TABLE 6: Geographic information summary of the study area.

| Label | Province | Latitude | Longitude | Approximate height above sea level (m) | Basin size | Airflow difficulty level |
|---|---|---|---|---|---|---|
| $L_1$ | Mae Hong Son | 19.137805 | 97.975261 | 330 | Moderate | 3 |
| $L_2$ | Chiang Mai | 18.837539 | 98.971612 | 310 | Narrow | 4 |
| $L_3$ | Lamphun | 18.657448 | 99.020208 | 290 | Narrow | 4 |
| $L_4$ | Chiang Rai | 19.932397 | 99.799329 | 410 | Moderate | 3 |
| $L_5$ | Phayao | 19.211871 | 100.209360 | 400 | Moderate | 3 |
| $L_6$ | Lampang | 18.278623 | 99.506985 | 235 | Normal | 2 |
| $L_7$ | Phrae | 19.714564 | 100.180450 | 155 | Moderate | 3 |
| $L_8$ | Nan | 18.814173 | 100.781905 | 250 | Normal | 2 |
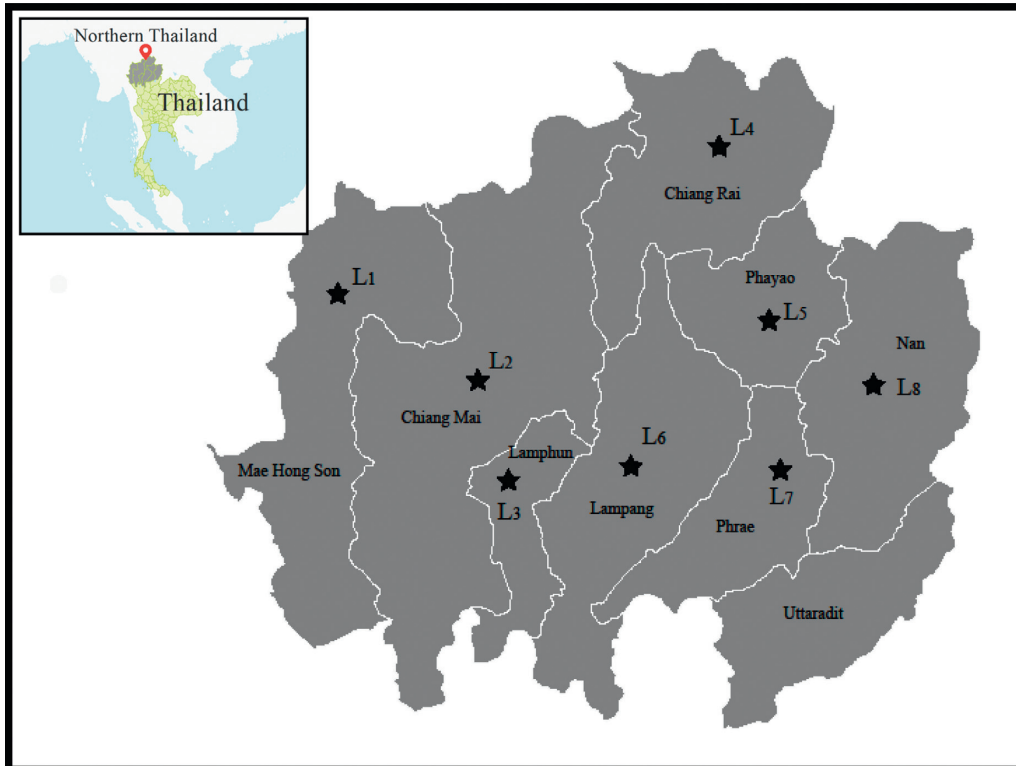


FIGURE 1: The study area: eight selected provinces in the northern region of Thailand. The figure is obtained from Samphutthanon et al. [29].
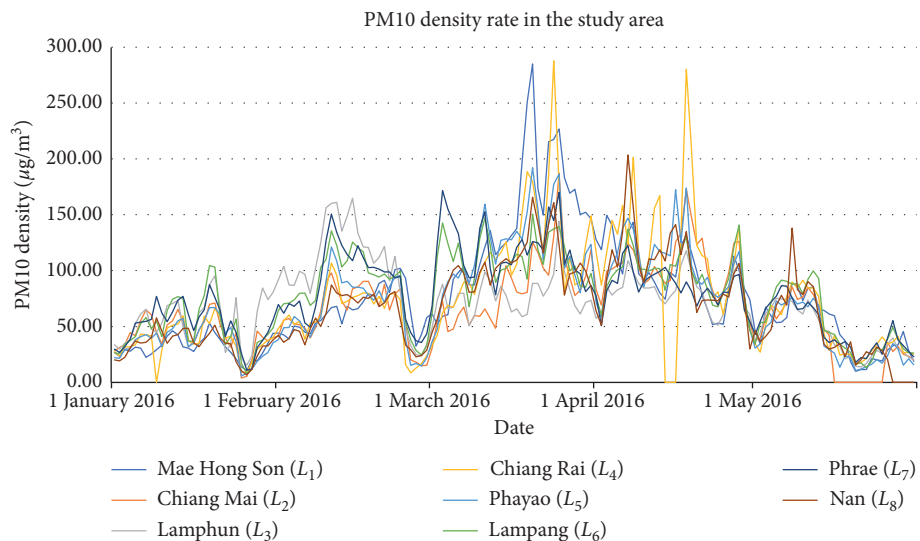


FIGURE 2: PM10 density rate of the eight selected locations from 1st January 2016 to 31st May 2016.

TABLE 7: The summary statistics of PM10 density of the eight selected locations from 1st January 2016 to 31st May 2016.

| Location | Min | Median | Max | Average | S.D. |
|---|---|---|---|---|---|
| $L_1$ | 4.25 | 61.75 | 385.00 | 76.18 | 57.86 |
| $L_2$ | 2.00 | 66.25 | 256.50 | 69.98 | 37.44 |
| $L_3$ | 1.00 | 69.75 | 225.50 | 71.11 | 36.52 |
| $L_4$ | 3.00 | 66.38 | 373.00 | 75.59 | 55.01 |
| $L_5$ | 1.50 | 67.50 | 280.00 | 73.24 | 47.31 |
| $L_6$ | 3.50 | 79.50 | 206.75 | 78.14 | 38.92 |
| $L_7$ | 8.25 | 77.50 | 230.75 | 77.84 | 38.35 |
| $L_8$ | 3.75 | 64.75 | 252.50 | 69.58 | 42.21 |
| *All locations* | *1.00* | *69.00* | *385.00* | *74.01* | *44.98* |

TABLE 8: Parameter information summary of the haze pollution problem.

| Label | Parameters | Parameter types | Effects (for atmospheric parameters) |
|---|---|---|---|
| $P_1$ | PM10 density | Atmospheric | Positive |
| $P_2$ | Air pressure | Atmospheric | Positive |
| $P_3$ | Relative humidity | Atmospheric | Positive |
| $P_4$ | Wind speed | Atmospheric | Negative |
| $P_5$ | Rainfall | Atmospheric | Negative |
| $P_6$ | Temperature | Atmospheric | Negative |
| $P_7$ | Airflow difficulty level | Topographic | N/A |

level is set at $120 \, \mu g/m^3$ based on Thailand national ambient air quality standard [34].

*4.1.1. Warning System Based on PM10 Density.* The trivial warning system is a warning that relies on the information of the PM10 density only. That is, a warning is signaled when the PM10 density at current time exceeds a certain threshold value. The warning system is generated by Algorithm 2.

*Algorithm 2.* Haze warning system based on PM10 density.

    (1) At the warning time, input PM10 density data of each location. The inputted data are the average of hourly data of the components in the preceding 4 hours.

    (2) A warning is signaled if the PM10 density of the location exceeds the threshold value $\alpha$.

The efficiency of the algorithm is evaluated by the accuracy ratio compared to the real data. The prediction is counted as accurate if the warning is signaled and the PM10 density in the next 4 hours exceeds $120 \, \mu g/m^3$ or the warning is not signaled and the PM10 density in the next 4 hours does not exceed $120 \, \mu g/m^3$. We test the algorithm with $\alpha = 84, 96, 108, 114$ and $118$ which are, respectively, 70%, 80%, 90%, 95%, and 98% of the crisis level. The accuracy ratios of each threshold values are shown in Table 9. The plot between the average accuracy ratio of all eight locations and the threshold values is shown in Figure 3. It can be seen that the best threshold value for these data is 118 (98% of the crisis level) with 90.99% accuracy ratio.

*4.1.2. Warning System Based on Fuzzy Soft Set with Weighted Information.* To improve the efficiency of the warning

TABLE 9: The accuracy ratio of the haze warning system by Algorithm 3.

| Location | The threshold value $\alpha$ out of $120 \, \mu g/m^3$ | | | | |
|---|---|---|---|---|---|
| | 70% | 80% | 90% | 95% | 98% |
| $L_1$ | 80.46 | 85.84 | 89.46 | 90.34 | 90.89 |
| $L_2$ | 79.36 | 86.39 | 90.12 | 92.21 | 92.86 |
| $L_3$ | 76.29 | 86.39 | 91.66 | 92.32 | 92.97 |
| $L_4$ | 78.59 | 85.07 | 89.02 | 88.91 | 91.88 |
| $L_5$ | 78.49 | 85.29 | 89.02 | 90.12 | 89.79 |
| $L_6$ | 67.29 | 76.51 | 84.96 | 86.39 | 87.27 |
| $L_7$ | 69.81 | 82.11 | 88.04 | 90.12 | 91.22 |
| $L_8$ | 77.06 | 83.32 | 88.14 | 90.78 | 91.00 |
| *Average* | *75.92* | *83.86* | *88.80* | *90.15* | *90.99* |

system, the fuzzy soft set with weighted information can be comprised. Note that the fuzzy soft set without weights is not suitable for this model. This is due to the fact that the importance of the parameters is not the same. For instance, PM10 density parameter is the most important parameter than the other parameters for the reason that no haze problem will occur if the PM10 density amount is low. It should be noted that the membership values of the atmospheric parameters change in every warning based on the real-time data, while the topographic parameter remains the same throughout the time period. Additionally, when the weighted information is $w = (1, 0, \ldots, 0)$, this warning system turns out to be the warning system based on PM10 density defined in Section 4.1.1.

The choice values are used in decision making. For this system, a warning is signaled when the weighted choice values at current time exceed a certain threshold value.

Our proposed decision making for the warning system with weighted information is as follows.
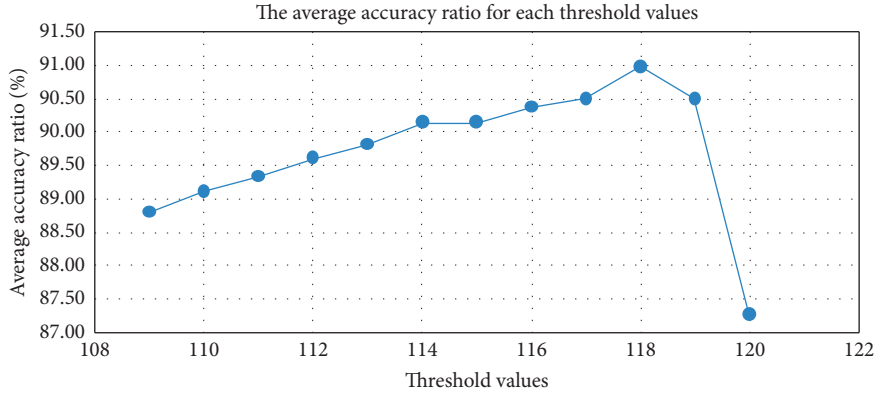
FIGURE 3: The plot between the average accuracy ratio of the eight selected locations and the threshold values.

*Algorithm 3.* Haze warning system based on weighted choice values.

(1) At the warning time, input the atmospheric data of each location: PM10 density, air pressure, relative humidity, wind speed, rain, and temperature. The inputted data are the average of hourly data of the components in the preceding 4 hours. Additionally, input the weight information $w = (w_1, w_2, \ldots, w_7)$.

(2) Calculate the membership values of the parameters of the fuzzy soft set:

(i) For PM10 density parameter, the membership values are calculated from

$$\mu_F(x) = \begin{cases} \dfrac{x}{120}, & x < 120, \\ \\ 1, & x \geq 120, \end{cases} \tag{15}$$

where $x$ is the inputted PM10 density data.

(ii) For the other positive atmospheric parameters, the membership values are calculated from

$$\mu_F(x) = \frac{x - m}{M - m}, \tag{16}$$

where $x$ is the inputted atmospheric component data and $m$ and $M$ are the minimum value and the maximum value of the atmospheric component during January to May 2016, respectively.

(iii) For negative atmospheric parameters, the membership values are calculated from

$$\mu_F(x) = \frac{M - x}{M - m}, \tag{17}$$

where $x$, $m$, and $M$ are defined in (ii)

(iv) For the topographic parameters, the membership values are 0, 0.25, 0.5, 0.75, and 1 when the airflow difficulty levels are 0, 1, 2, 3, and 4, respectively.

(3) Calculate the weighted choice values according to the weight information $w$ of each location.

(4) A warning is signaled if the choice values of the location exceed the threshold value $\alpha$ where $0 \leq \alpha \leq 1$.

The flowchart of Algorithm 3 is given in Figure 4.

Clearly, the accuracy ratio of the model depends on the weight information and the threshold value. The calculation examples when weighted information is $w_1 = (1, 1, 1, 1, 1, 1, 1)$, $w_2 = (5, 2, 1, 2, 2, 1, 2)$, $w_3 = (10, 2, 1, 2, 2, 1, 2)$, $w_4 = (15, 2, 1, 2, 2, 1, 2)$, and $w_5 = (20, 2, 1, 2, 2, 1, 2)$ are shown in Table 10. In these examples, the threshold value is set to be 90% of the possible maximum values, which depend on the weight information.

Since the aim of this problem is to find the weight information and the threshold value that give the best accuracy ratio, this problem coincides with the optimization problem:

$$\begin{aligned} \text{max} \quad & \text{Average accuracy ratio,} \\ & w_1, w_2, \ldots, w_7 \text{ are integers} \\ \text{subject to} \quad & 0 \leq w_1 \leq 30 \\ & 0 \leq w_2, w_3, \ldots, w_7 \leq 10 \\ & \alpha = 0, 0.01, 0.02, \ldots, 0.99, 1. \end{aligned} \tag{18}$$

By employing the particle swarm optimization method in Matlab programme, the optimum average accuracy ratio is 92.12% with the optimum weight (17, 0, 0, 2, 3, 0, 0) and the optimum threshold $\alpha = 0.98$. This optimum result is shown in Table 11.

*4.2. Identification of the Most Hazardous Location.* The second aim of this research is to identify the location with the most serious haze pollution problem based on real-time atmospheric data. The location is identified at the same time as the warning. The effective prediction will benefit the community in the affected area and assist the authority to provide safety aids and prepare helping devices such as mobile air purifier.

*4.2.1. Identification of the Most Hazardous Location Based on PM10 Density.* Similar to Section 4.1.1, the simple decision

TABLE 10: The accuracy ratio of the haze warning system by Algorithm 3 with weight information. The threshold value of each weight is set to be 90% of the possible maximum.

| Location | Weight information | | | | |
|---|---|---|---|---|---|
| | $w_1$ (%) | $w_2$ (%) | $w_3$ (%) | $w_4$ (%) | $w_5$ (%) |
| $L_1$ | 83.75 | 89.13 | 90.34 | 90.45 | 90.34 |
| $L_2$ | 91.33 | 91.33 | 91.33 | 91.33 | 92.97 |
| $L_3$ | 91.22 | 91.22 | 91.22 | 93.85 | 92.97 |
| $L_4$ | 85.18 | 85.18 | 85.51 | 88.58 | 88.80 |
| $L_5$ | 84.52 | 84.52 | 84.52 | 84.52 | 85.40 |
| $L_6$ | 85.73 | 85.73 | 85.73 | 85.73 | 85.73 |
| $L_7$ | 87.49 | 87.49 | 87.49 | 87.49 | 89.79 |
| $L_8$ | 89.24 | 89.24 | 89.24 | 89.24 | 89.13 |
| *Average* | *87.31* | *87.98* | *88.17* | *88.90* | *89.39* |

TABLE 11: The optimum average accuracy ratio of the haze warning system based on Algorithm 3.

| Location | Threshold value $\alpha = 0.98$ $w = (17, 0, 0, 2, 3, 0, 0)$ (%) |
|---|---|
| $L_1$ | 92.11 |
| $L_2$ | 94.08 |
| $L_3$ | 93.97 |
| $L_4$ | 93.65 |
| $L_5$ | 91.56 |
| $L_6$ | 88.16 |
| $L_7$ | 91.23 |
| $L_8$ | 92.22 |
| Average | 92.12 |

(1) At the warning time, input PM10 density data of each location. The inputted data are the average of hourly data of the components in the preceding 4 hours.

(2) The decision is $L_k$, the location with the maximum value of PM10 density at current time. Optimal choices may have more than one if there are more than one element corresponding to the maximum.

The efficiency of the algorithm is evaluated by the accuracy ratio compared to the real data. The prediction is counted as accurate if the most severe location in the next 4 hours is correctly identified. By making decision based on Algorithm 4, the average accuracy ratio from eight locations is 51.15% and Cohen's kappa index of agreement is 0.4312.

*4.2.2. Identification of the Most Hazardous Location Based on Fuzzy Soft Set with Weighted Information.* The fuzzy soft set with weighted information can be comprised in order to improve the efficiency of the decision makings. With a similar reason to Section 4.1.2, the fuzzy soft set with weight is more suitable. Note that the membership values of the atmospheric parameters change in every decision making based on the real-time data, while the topographic parameter remains the same throughout the time period. Additionally, when the weighted information is $w = (1, 0, \ldots, 0)$, this decision making turns out to be the warning system based on PM10 density defined in Section 4.2.1. It should be emphasized that the membership calculation of PM10 density parameter is different from Algorithm 3. This is because we need to make a comparison of location.

Finally, the evaluation of decision making must be chosen. Note that it can be evaluated based on choice values, score values, or grey relation grade. In our result, we will use all three evaluations in order to choose which evaluation gives the best result.

Our proposed algorithm for decision making of the most hazardous location based on weighted choice values is as follows.

*Algorithm 5.* Identification of the most hazardous location based on weighted choice values.

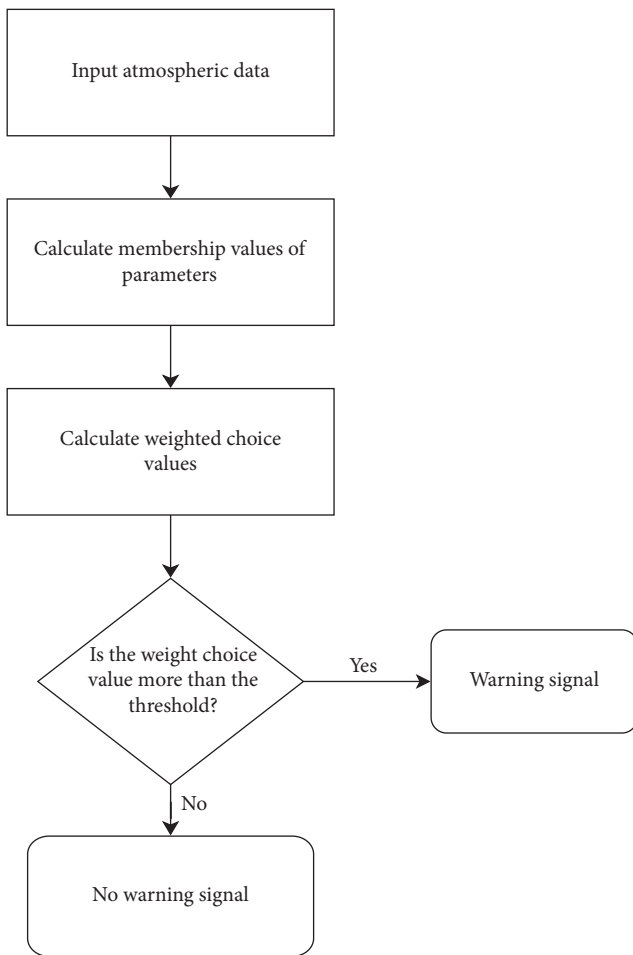(1) At the warning time, input the atmospheric data of each location: PM10 density, air pressure, relative



FIGURE 4: Flowchart of Algorithm 3.

making is to choose a location based on the information of PM10 density only. That is, the location with the highest value of PM10 density at current time is chosen as the most hazardous location in the following 4 hours.

The algorithm of the decision making is as follows.

*Algorithm 4.* Identification of the most hazardous location based on PM10 density

humidity, wind speed, rain, and temperature. The inputted data are the average of hourly data of the components in the preceding 4 hours. Additionally, input the weight information $w = (w_1, w_2, \ldots, w_7)$.

(2) Calculate the membership values of the parameters of the fuzzy soft set:

(i) For positive atmospheric parameters, the membership values are calculated from

$$\mu_F(x) = \frac{x - m}{M - m}, \tag{19}$$

where $x$ is the inputted atmospheric component data and $m$ and $M$ are the minimum value and the maximum value of the atmospheric component during January to May 2016, respectively.

(ii) For negative atmospheric parameters, the membership values are calculated from

$$\mu_F(x) = \frac{M - x}{M - m}, \tag{20}$$

where $x$, $m$, and $M$ are defined in (i).

(iii) For the topographic parameters, the membership values are 0, 0.25, 0.5, 0.75, and 1 when the airflow difficulty levels are 0, 1, 2, 3, and 4, respectively.

(3) Calculate the weighted choice values $c_w$.

(4) The decision is $L_k$ if $c_w(L_k) = \max_i\{c_w(L_i)\}$. Optimal choices may have more than one if there are more than one element corresponding to the maximum.

The flowchart of Algorithm 5 is given in Figure 5.

*Remark 3.* If the decision making is based on weighted score values $s_w$ or grey relational grades $\gamma_w$, then Step 3 and Step 4 of Algorithm 5 will be changed accordingly.

Clearly, the accuracy ratio of the model depends on the weight information. Table 12 displays the accuracy ratio of the location identification by Algorithm 5 where the decision makings are based on choice values, score values, and grey relational grades. The weighted information are $w_1 = (1, 1, 1, 1, 1, 1, 1), w_2 = (5, 2, 1, 2, 2, 1, 2), w_3 = (10, 2, 1, 2, 2, 1, 2), w_4 = (15, 2, 1, 2, 2, 1, 2)$  and  $w_5 = (20, 2, 1, 2, 2, 1, 2)$.

Since our desire of this problem is to find the weight information that gives the best accuracy ratio, this is similar to the optimization problem:

$$
\begin{aligned}
\text{max} \quad & \text{Average accuracy ratio,} \\
& w_1, w_2, \ldots, w_7 \text{ are integers} \\
\text{subject to} \quad & 0 \leq w_1 \leq 30 \\
& 0 \leq w_2, w_3, \ldots, w_7 \leq 10.
\end{aligned}
\tag{21}
$$

By employing the particle swarm optimization method in Matlab programme, the optimum average accuracy ratio based on weighted choice values, weighted score values, and



Input atmospheric data

Calculate membership values of parameters

Calculate weighted choice values or weighted score values or grey relational grade

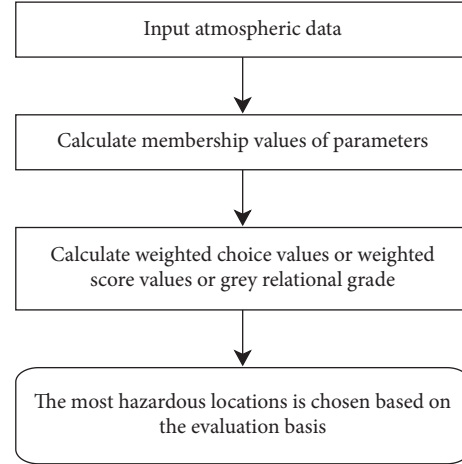The most hazardous locations is chosen based on the evaluation basis

Figure 5: Flowchart of Algorithm 5.

grey relational grades is 56.58%, 57.13%, and 57.02%, respectively. The summary of the optimum result is shown in Table 13. Cohen's kappa of the decision making based on weighted choice values, weighted score values, and grey relational grades is 0.4457, 0.4521, and 0.4489, respectively.

### 4.3. Discussion

*4.3.1. Haze Warning System.* By introducing the fuzzy soft model with weighted information, the prediction accuracy ratio of the warning system is improved slightly from 90.99% to 92.12% compared to the simple warning system that only considers the PM10 density. Moreover, it is clear that the fuzzy soft models with weighted information provide better prediction than the original (equal weight) fuzzy soft model. Table 14 shows the parameters' weights that provide the best accuracy ratio. Note that the principal parameters are PM10 density, rainfall, and wind speed, respectively, while the other parameters have no weight. This suggests that a simple judgment on the warning can be done by observing only PM10 density, wind speed, and rainfall. The problem is expected to be severe if PM10 density is high, with no wind and no rain. This agrees with the principle study in environmental science research.

*4.3.2. Identification of the Most Hazardous Location.* By selecting the most severe location based on the information from PM10 density only, the accuracy ratio is 51.12%. However, this ratio is improved to 57.13% when the locations are chosen by the fuzzy weight model. The decision making is decided by weighted score values. Table 15 shows the parameters' weights that provide the best accuracy ratio.

Based on the optimal parameters' weights, this would imply the following:

(1) *PM10 density* is clearly the main factor in the decision making.

(2) This result shows that *topography* plays a role in the haze pollution problem for this region of study.

TABLE 12: The accuracy ratio of the location identification by Algorithm 5 with weight information.

| Decision based | | Weight information | | | | |
|---|---|---|---|---|---|---|
| | | $w_1$ (%) | $w_2$ (%) | $w_3$ (%) | $w_4$ (%) | $w_5$ (%) |
| Average accuracy ratio | Choice values | 21.18 | 25.80 | 33.37 | 40.40 | 45.55 |
| | Score values | 26.56 | 42.15 | 48.96 | 50.49 | 51.81 |
| | Grey relational grades | 26.13 | 37.43 | 45.99 | 49.84 | 51.04 |

TABLE 13: The best accuracy ratio.

| Decision based | Accuracy ratio (%) | Optimal weight |
|---|---|---|
| Choice values | 56.58 | (20, 0, 0, 2, 4, 1, 1) |
| Score values | 57.13 | (18, 1, 0, 2, 2, 3, 3) |
| Grey relational grades | 57.02 | (23, 1, 0, 2, 2, 3, 2) |

TABLE 14: Parameters and optimum weights for the haze warning system.

| Label | Parameters | Weight | Effects (for atmospheric parameters) |
|---|---|---|---|
| $P_1$ | PM10 density | 17 | Positive |
| $P_2$ | Air pressure | 0 | Positive |
| $P_3$ | Relative humidity | 0 | Positive |
| $P_4$ | Wind speed | 2 | Negative |
| $P_5$ | Rainfall | 3 | Negative |
| $P_6$ | Temperature | 0 | Negative |
| $P_7$ | Airflow difficulty level | 0 | N/A |

TABLE 15: Parameters and optimum weights for the identification of the most hazardous location.

| Label | Parameters | Weight | Effects (for atmospheric parameters) |
|---|---|---|---|
| $P_1$ | PM10 density | 18 | Positive |
| $P_2$ | Air pressure | 1 | Positive |
| $P_3$ | Relative humidity | 0 | Positive |
| $P_4$ | Wind speed | 2 | Negative |
| $P_5$ | Rainfall | 2 | Negative |
| $P_6$ | Temperature | 3 | Negative |
| $P_7$ | Airflow difficulty level | 3 | N/A |

(3) Temperature, wind speed, and rainfall are factors in the model. Unfortunately, these atmospheric parameters are uncontrollable.

(4) Air pressure and relative humidity have less or no impact for the prediction model.

This study analysis agrees with principle study in environmental science research. It should be emphasized that the only parameter that can be controlled is PM10 density. The activities that contribute to PM10 such as outdoor burn or car emissions should be disregarded.

*4.3.3. Other Discussion.* By the results from Sections 4.1 and 4.2, it should be pointed out that a simple warning system and location identification based on the information of PM10 density is reasonable enough. By adding the parameters, the efficiency of the model is improved very slightly. This emphasizes the fact that environmental modeling is complicated. However, since the calculation of our algorithm is not expensive, Algorithms 3 and 5 should still be in use to improve the decision-making problem.

For further works, our suggestions are to add the following parameters:

*Atmospheric parameters:* PM2.5 density, $SO_2$, ozone, and wind direction.

*Topographic parameters:* height above sea level of the location, location of surrounded mountains, and height of surrounded mountains.

*Others parameters:* population.

## 5. Conclusions

In this article, we propose a fuzzy soft model to benefit in the haze pollution management in northern Thailand. The main aims of this research are to provide a haze warning system based on real-time atmospheric data and to identify the most hazardous location of the study area. The study area covers eight provinces in the northern Thailand, where the problem severely occurs every year. The parameters of the fuzzy soft set include both atmospheric parameters and topographic parameter. The membership values of atmospheric parameters are calculated based on the real-time data. The

efficiency of the model is tested with the real data from 1st January 2016 to 31st May 2016. The results show that our fuzzy models improve the prediction accuracy ratio compared to the prediction based on PM10 density only. The optimum results and optimum weights are chosen based on particle swarm optimization. The meaning of optimum weights also agrees with the principle study in environmental science research. Another benefit of our model is that the topographic parameter, which is normally being disregarded from many models, is included. Moreover, our model would offer an alternative prediction model for the haze pollution problem in northern Thailand.

The fuzzy soft set approach in the application to haze pollution management furnishes very promising prospect and possibilities. We strongly believe that the efficiency of the model can be improved when appropriate parameters are added. The calculation formula for the membership values and the severity index can also be adjusted. The efficient model will clearly improve the health safety and raise the life quality of the sufferers.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The author declares that there are no conflicts of interest.

## Acknowledgments

## References

[1] Pollution Control Department (PCD) Website, http://www.pcd.go.th.

[2] Climate Change Data Centre of Chiang Mai University (CMU CCDC) Website, http://www.cmuccdc.org.

[3] Smoke Haze Integrated Research Unit (SHIRU) Website, http://www.shiru-cmu.org/.

[4] S. Phoothiwut and S. Junyapoon, "Size distribution of atmospheric particulates and particulate-bound polycyclic aromatic hydrocarbons and characteristics of PAHs during haze period in Lampang Province, Northern Thailand," *Air Quality, Atmosphere & Health*, vol. 6, no. 2, pp. 397–405, 2013.

[5] S. Chantara, S. Sillapapiromsuk, and W. Wiriya, "Atmospheric pollutants in Chiang Mai (Thailand) over a five-year period (2005–2009), their possible sources and relation to air mass movement," *Atmospheric Environment*, vol. 60, pp. 88–98, 2012.

[6] W. Wiriya, T. Prapamontol, and S. Chantara, "PM10-bound polycyclic aromatic hydrocarbons in Chiang Mai (Thailand): seasonal variations, source identification, health risk assessment and their relationship to air-mass movement," *Atmospheric Research*, vol. 124, pp. 109–122, 2013.

[7] P. Pengchai, S. Chantara, K. Sopajaree, S. Wangkarn, U. Tengcharoenkul, and M. Rayanakorn, "Seasonal variation, risk assessment and source estimation of PM10 and PM10-bound PAHs in the ambient air of Chiang Mai and Lamphun, Thailand," *Environmental Monitoring and Assessment*, vol. 154, no. 1–4, pp. 197–218, 2009.

[8] D. Molodtsov, "Soft set theory-first results," *Computers & Mathematics with Applications*, vol. 37, no. 4-5, pp. 19–31, 1999.

[9] P. K. Maji, R. Biswas, and A. R. Roy, "Fuzzy soft sets," *The Journal of Fuzzy Mathematics*, vol. 9, no. 3, pp. 589–602, 2001.

[10] N. Cagman, S. Enginoglu, and F. Citak, "Fuzzy soft set theory and its applications," *Iranian Journal of Fuzzy System*, vol. 8, pp. 137–147, 2011.

[11] B. Fisher, "Fuzzy environmental decision-making: applications to air pollution," *Atmospheric Environment*, vol. 37, no. 14, pp. 1865–1877, 2003.

[12] A. K. Gorai, K. Kanchan, A. Upadhyay, and P. Goyal, "Design of fuzzy synthetic evaluation model for air quality assessment," *Environment Systems and Decisions*, vol. 34, no. 3, pp. 456–469, 2014.

[13] A. Khan, "Use of fuzzy set theory in environmental engineering applications: a review," *International Journal of Engineering Research and Applications*, vol. 7, no. 6, pp. 1–6, 2017.

[14] H. Yang, Z. Zhu, C. Li, and R. Li, "A novel combined forecasting system for air pollutants concentration based on fuzzy theory and optimization of aggregation weight," *Applied Soft Computing*, vol. 87, Article ID 105972, 2020.

[15] S. L. Chen, "The application of comprehensive fuzzy judgement in the interpretation of water-flooded reservoirs," *The Journal of Fuzzy Mathematics*, vol. 9, no. 3, pp. 739–743, 2001.

[16] S. J. Kalayathankal and G. S. Singh, "A fuzzy soft flood alarm model," *Mathematics and Computers in Simulation*, vol. 80, Article ID 887893, 2010.

[17] S. J. Kalayathankal, G. S. Singh, and P. B. Vinodkumar, "MADM models using ordered ideal intuitionistic fuzzy sets," *Advances in Fuzzy Mathematics*, vol. 4, no. 2, pp. 101–106, 2009.

[18] P. C. Nayak, K. P. Sudheer, and K. S. Ramasastri, "Fuzzy computing based rainfall-runoff model for real time flood forecasting," *Hydrological Processes*, vol. 19, no. 4, pp. 955–968, 2005.

[19] E. Toth, A. Brath, and A. Montanari, "Comparison of short-term rainfall prediction models for real-time flood forecasting," *Journal of Hydrology*, vol. 239, no. 1–4, pp. 132–147, 2000.

[20] S. J. Kalayathankal, G. S. Singh, S. Joseph et al., "Ordered ideal intuitionistic fuzzy model of flood alarm," *Iranian Journal of Fuzzy Systems*, vol. 9, no. 3, pp. 47–60, 2012.

[21] J. Tan, J. Fu, G. Carmichael et al., "Why models perform differently on particulate matter over East Asia? A multi-model intercomparison study for MICS-Asia III," *Atmospheric Chemistry and Physics Discussions*, 2019, in review.

[22] R. Bhakta, P. S. Khillare, and D. S. Jyethi, "Atmospheric particulate matter variations and comparison of two forecasting models for two Indian megacities," *Aerosol Science and Engineering*, vol. 3, no. 2, pp. 54–62, 2019.

[23] B. Pimpunchat and S. Junyapoon, "Modeling haze problems in the north of Thailand using logistic regression," *Journal of Mathematical and Fundamental Sciences*, vol. 46, no. 2, pp. 183–193, 2014.

[24] B. Mitmark and W. Jinsart, "A GIS model for PM10 exposure from biomass burning in the north of Thailand," *Applied Environmental Research*, vol. 39, no. 2, pp. 77–87, 2017.

[25] A. R. Roy and P. K. Maji, "A fuzzy soft set theoretic approach to decision making problems," *Journal of Computational and Applied Mathematics*, vol. 203, no. 2, pp. 412–418, 2007.

[26] Z. Kong, L. Wang, and Z. Wu, "Application of fuzzy soft set in decision making problems based on grey theory," *Journal of Computational and Applied Mathematics*, vol. 236, no. 6, pp. 1521–1530, 2011.

[27] J. Kennedy and R. Eberhart, "Particle swarm optimization," *Proceedings of the International Conference on Neural Networks*, vol. 4, 1995.

[28] A. Meneses, M. Machado, and R. Schirru, "Particle swarm optimization applied to the nuclear reload problem of a pressurized water reactor," *Progress in Nuclear Energy*, vol. 51, no. 3, pp. 319–326, 2009.

[29] R. Samphutthanon, N. K. Tripathi, S. Ninsawat et al., "Spatio-temporal distribution and hotspots of hand, foot and mouth disease (HFMD) in northern Thailand," *International Journal of Environmental Research and Public Health*, vol. 11, no. 1, pp. 312–336, 2013.

[30] L. Wang, N. Zhang, Z. Liu et al., "The influence of climate factors, meteorological conditions, and boundary-layer structure on severe haze pollution in the beijing-tianjin-hebei region during January 2013," *Advances in Meteorology*, vol. 2014, Article ID 685971, 14 pages, 2014.

[31] S. Zhou, S. Peng, M. Wang et al., "The characteristics and contributing factors of air pollution in nanjing: a case study based on an unmanned aerial vehicle experiment and multiple datasets," *Atmosphere*, vol. 9, no. 9, p. 343, 2018.

[32] N. Pasukphun, "Environmental health burden of open burning in northern Thailand: a review," *PSRU Journal of Science and Technology*, vol. 3, no. 3, pp. 11–28, 2018.

[33] W. Mankan, "The causing factors of the smog phenomenon in Lampang basin," *Burapha Science Journal*, vol. 22, no. 3, pp. 226–239, 2017.

[34] L. Pardthaisong, P. Sin-ampol, C. Suwanprasit, and A. Charoenpanyanet, "Haze pollution in Chiang Mai, Thailand: a road to resilience," *Procedia Engineering*, vol. 212, pp. 85–92, 2018.