# Temporal Peak Time Frame of Traffic-Related Air Pollutants in Surabaya

Arie Dipareza Syafei[1], Akimasa Fujiwara[2], Junyi Zhang[3] and Makoto Chikaraishi[4]

[1] Member of JSCE, Lecturer, Dept. of Env. Eng., Institut Teknologi Sepuluh Nopember (Kampus ITS, Surabaya, Indonesia, 60111); Email: asyafei@gmail.com

[2] Member of JSCE, Professor, Graduate School for Int'l Dev. and Coop., Hiroshima Univ. (1-5-1 Kagamiyama, Higashi-Hiroshima, 739-8529, Japan); E-mail: afujiw@hiroshima-u.ac.jp

[3] Member of JSCE, Professor, Graduate School for Int'l Dev. and Coop., Hiroshima Univ. (1-5-1 Kagamiyama, Higashi-Hiroshima, 739-8529, Japan); E-mail: zjy@hiroshima-u.ac.jp

[4] Member of JSCE, Associate Professor, Graduate School for Int'l Dev. and Coop., Hiroshima Univ. (1-5-1 Kagamiyama, Higashi-Hiroshima, 739-8529, Japan); E-mail: chikaraishim@hiroshima-u.ac.jp

Air quality management is crucial for preventing air pollution from exceeding threshold ambient (permissible) levels, which adversely affect human health. This study investigates temporal peak time frames and events, in which nitrogen dioxide ($NO_2$) concentrations exceed ambient levels, using daily diurnal analysis and an independent component analysis (ICA) technique. Based on the daily diurnal analysis, the intervals for peak concentrations of $NO_2$ occur between 6 am and 9 am, which indicates a correlation with traffic during work and school hours. The evening interval for peak $NO_2$ concentrations in the range between 4:30 pm and 9 pm. The ICA generates detailed profiles for each zone, e.g., the $NO_2$ profiles for the morning and evening periods for the city center, suburban1 (suburban west of Surabaya) and near-highway zones are similar, which suggests similar travel behavior patterns. Based on the ICA, we are able to determine time frames from which we can apply traffic-related policies for locations near the monitoring stations, e.g., in the trading zone. The local government may adjust the time distribution of large vehicles that pass the road near the station. These results imply the importance of obtaining the temporal peak time frame as an initial step in integrated transportation management by the government to improve environmental conditions.

***Keywords:*** *diurnal analysis, air quality management, nitrogen dioxide, independent component analysis, monitoring stations, and traffic emissions*

## 1. Introduction

Air pollution is a crucial issue in the development of cities because it is related to human health and activities. For example, visibility may be interrupted if particulate matter exceeds ambient criteria. Adverse health effects may occur if some nitrogen dioxide ($NO_2$) and particulate matter size≤10μm $PM_{10}$ concentrations exceed threshold levels. The exposure to fine or coarse ambient particular matter or $NO_2$ can negatively affect human health (Han et al. 2012), other gas pollutant e.g., $PM_{10}$, also affects visibility and ecosystems (Minguillón et al. 2013).

In Surabaya City, a significant annual increase in vehicles—20% for motorcycles and 10% for private cars (Department of Transportation, Surabaya City Government)—has been observed. Because these vehicles emit pollutants, the background concentrations in the atmosphere are affected. $NO_2$ and particulate matter with particle sizes less than 10 μm are emitted from vehicles. However, $PM_{10}$ is produced not only by transportation sources but also by industrial and other activities, e.g., gardening and construction, whereas $NO_2$ is emitted largely from vehicles. Because the concentration of $NO_2$ has increased over the years, the local government has considered this issue in its air quality management program. The government proposed the installation of sensors at five monitoring stations to monitor ambient air quality, particularly traffic-related air quality and the concentrations of $NO_2$. The monitoring data that was continuously obtained from these sites was expected to be effective for improving in air quality because the monitoring data exhibit multiple distributions, which are caused by various factors such as meteorology, topography and human activities. However, traffic produced by urban economic activity is likely to be the most influential factor of pollution.

A monitoring station is useful for identifying events in which the pollutants exceed threshold values (permissible level). Table 1 shows the number of times that pollutants breach these levels from 2001 to 2002. The permissible levels for $NO_2$ based on regulations in East Java is 92.5 $\mu g/m^3$. We were forced to use old data because the missing data for the majority of stations after 2002 were considered to be extreme (>50%) due to a sensor disorder. However, the data adequately represent the poor pollution levels, which are caused by a rapid increase of 50% for the vehicles in Surabaya in 2010 (Revenue Agency of East Java Province). As a result, the objective of this study will not be compromised.

**Table 1** Number of times in which pollutants breach their permissible levels for each station

| Monitoring Station | $NO_2$ ($\mu g/m^3$) |
|---|---|
| Station 1 (city center) | 284 |
| Station 2 (trading zone) | 110 |
| Station 3 (suburban1, west of Surabaya) | 85 |
| Station 4 (near-highway zone) | 390 |
| Station 5 (suburban2, east side of Surabaya) | 88 |

Due to the occurrence of numerous breaches and the substantial increase in vehicles in Surabaya, policies must be formulated and issued. Currently, some policies that have been implemented to reduce ambient levels are related to land use, e.g., numerous green spaces were constructed in Surabaya. No revolutionary policies were established to control vehicle ownership. Therefore, policy makers frequently lack options, with the exception of transportation management. In transportation management, this objective should be included as part of the ultimate goal to reduce the risk of exposure to pollutants. An aspect of transport management is to assess the temporal patterns of pollutant concentrations. After the temporal pattern is observed, we can extract and identify the time frame for which the concentrations are likely to be high and breach permissible levels during the day. This time frame is referred to as the temporal peak time frame. We can determine suitable policies, specifically transportation-related policies, for a particular zone.

To obtain the temporal peak time frame, we analyze and employ original data to determine the time frame and the day in which the events are breached. When performing the analysis, our basic hypothesis is that the temporal peak time frame for each location may be unique because the monitoring stations are located in areas with different land use patterns and different concentration profiles. To determine the pattern and identify the obscure and distributed concentration for each zone, we also employ an independent component analysis (ICA). The ICA is the first analysis of temporal peak patterns of pollutant concentrations in environmental and urban management. Ordinary component analyses are discussed in the literature in terms of air quality. Abdul-Wahab et al. (2005) employed a principal component analysis (PCA) to model and investigate factors of ground-level ozone, and Al-Alawi et al. (2008) combined principal component regression with an artificial neural network to forecast ground-level ozone. PCA can also be used to identify air quality factors from ceramic tile clusters (Minguillón et al. 2013). Pires et al. (2008a), Pires et al., (2008b) employed PCA to optimize air quality monitoring stations in Portugal. However, PCA is limited compared with ICA.

Back and Weigend (1997) explained that ICA provides a better assessment of the underlying structure and information about the stock prices compared with PCA. PCA provides an orthogonal representation of the data and maximizes the variance explained by the components; however, the interpretability of the second and higher components may be limited (Westra et al. 2010). ICA produces output with a clear structure, e.g., clustered, because ICA focuses on non-Gaussian data. It is similar to projection because it finds directions on which the projection of the data set show the most interesting distribution, and it is shown to be when the data has the least Gaussian distribution. Hyvärinen and Oja (2000) demonstrated that PCA produces an indistinct and structured output.

ICA is an alternative approach for detecting components from a mixed distribution. Thus, it is frequently used to analyze time series data (Cheung and Xu 2001) in many research fields, including finance (Lu et al. 2009), climate change (Westra et al., 2010), multimedia (Long et al. 2012), and air quality forecasting (Pires et al. 2008c). The majority of the literature has concluded that ICA is helpful for filtering noises contained in the data. Thus, this study assumes that the independent patterns of $NO_2$ concentrations obtained from an ICA may detect noise-free temporal patterns between zones and maintain independency in the case of Surabaya City. These free-noise concentration profiles can be analyzed to obtain information about the magnitude of the concentrations for a specific time frame, how it differs between days, and how it differs between zones. We also explore the benefit of using ICA for identifying distributions of $NO_2$.

## 2. Data and monitoring stations

Data are collected from five monitoring stations in Surabaya City. Due to the high rate (>50%) of missing data after 2002, which was caused by a sensor disorder, we employ 30-min intervals for $NO_2$ concentration data from February 1, 2001 to September 30, 2002, because the omission rate is less than 16% for the pollutants and other parameters. However, we consider the data to be valid for current applications because the land use patterns surrounding the monitoring stations do not differ significantly. The number of vehicles has significantly increased, which substantially deteriorates air quality. The data are obtained from the Laboratory of Air Quality at the Environmental Agency of Surabaya City Government; all missing data is caused by the expectation-maximization algorithm.

Surabaya City installed five monitoring stations in areas with different types of land use and topography, as shown in **Fig. 1**. Generally, 80.72% of Surabaya consists of lowland with elevations ranging from 3-8m due to a low water spring (LWS); the remainder of the areas consist of hilly land (12.77%) ton the west and east of Surabaya (6.52%). The land slope ranges between 0 and 2% in lowland and 2-15% in shallow hilly lands.

- **Station 1** is located along Ketabang Kali St. in the flat land section of the city center. We refer to this point as the city center.
- **Station 2** is located along Johor Road (Rd.) in the coastal trading zone in north Surabaya. The sensor is located in the flat lowland within approximately 2000 m from Tanjung Perak Port Harbor. We refer to this station as the trading zone.
- **Station 3** is located in the Sukomanunggal region on the west side of Surabaya City. Despite the suburban characteristics of the area, the surrounding land use consists of towers, business offices, houses, and department stores. We refer to this station as suburban1.
- **Station 4** is located along Pagesangan Road in the Gayungsari region, which is located near a congested highway that connects Surabaya City to neighboring cities, e.g., Gresik, Porong, and Sidoarjo. We refer to this point as the highway zone.
- **Station 5** is located along Arief Rahman Hakim Street in the Sukolilo region. This east suburban area is dominated by universities, business offices, and houses. We refer to this station as suburban2.

**Table 2** lists the characteristics of the zones in which each monitoring station is located.

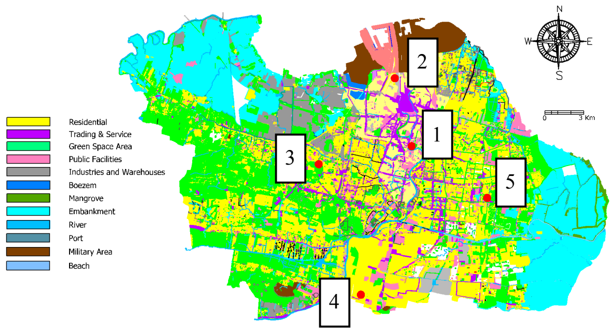


**Fig. 1** Five monitoring stations in Surabaya City

**Table 2** Monitoring stations and their coordinates

| Station | Zone | Land use | Population size (ca) | Topography | Average wind speed (m/s) | Dominant Wind direction |
|---|---|---|---|---|---|---|
| 1 | City center | Green space, parks, hotels, offices, houses, businesses | 64,400-65,296 | Lowland | 1.850 | East |
| 2 | Trading | Green spaces, houses, offices, industries, gas station | 89,348-90,418 | Lowland, near coastal area | 2.161 | Southeast to east |
| 3 | Suburban | Offices, few factories, houses, shopping malls | 85,012-86,421 | Hilly land | 3.373 | East |
| 4 | Near-highway | Religious building, businesses, houses | 83,644-85,012 | Lowland | 2.181 | Between east and southeast |
| 5 | Suburban | Convention halls, offices, universities, offices, houses | 75,021-77,362 | Lowland | 4.144 | East and southeast |

## 3. Methodology

ICA can be considered as an advanced method of PCA. The latter focuses on identifying components based on covariance and second-order statistics, whereas ICA uses higher-order statistics, which enable the algorithm to obtain statistically independent components (Westra et al. 2009).

The form of ICA consists of the observation matrix X, in this case $NO_2$ concentrations (ug/m$^3$), which is derived by mixing the $n$-dimensional source matrix S = $(s_1, \ldots, s_n)^T$ with the temporal dimension of $l$, which is referred to as independent components (ICs) and is assumed to be non-Gaussian, mutually statistically independent and a zero mean, where $n$ represents the independent extracted components. Since there are five monitoring stations, then we would like to extract five ICs. Assuming that the mixing is both linear and stationary, a typical ICA model is expressed as

$$\boldsymbol{X} = \boldsymbol{SA} \quad (1)$$

where $\boldsymbol{A}$ is the mixing matrix of dimension $n*n$ or $m*n$ in which $m \leq n$. The objective of an ICA is to estimate $\boldsymbol{A}$ and $\boldsymbol{S}$ with knowledge of the observations matrix X. A scalar multiple of $\boldsymbol{S}$ is obtained because any constant that is multiplied by an independent component in the equation can be cancelled by dividing the corresponding column of the mixing matrix $\boldsymbol{A}$ by the same constant.

The independent component $\boldsymbol{S}$ in the model is determined by searching the matrix $\boldsymbol{W}$ such that $\boldsymbol{S} = \boldsymbol{WX}$ with some indeterminacies. The *FastICA* algorithm is employed for an independent component analysis using the iterative fixed-point algorithm to obtain one unit as

$$\widetilde{w}_{n+1} = E\{x(w_n x) * g(|w_n x|^2)\} - E\{g(|w_n x|^2) + |w_n x|^2 g'(|w_n x|^2)\}w_n \quad (2)$$

where, $w_{n+1} = \frac{\widetilde{w}_{n+1}}{\|\widetilde{w}_{n+1}\|}$

By estimating $w$, we can obtain an IC by $\boldsymbol{S} = \boldsymbol{WX}$ because part of the process involves decorrelation of the outputs $w_1x$, … $w_nx$ after every iteration. Using the *FastICA* algorithm from R open source platform, we can estimate $\boldsymbol{A}$ and $\boldsymbol{S}$ from the observations $\boldsymbol{X}$, where $\boldsymbol{A} = \boldsymbol{W}^1$.

Note that ICA is sensitive to the selection of $n$; it is a key parameter that must be determined prior to performing ICA for interpretive applications. The process of obtaining components must undergo a preprocessing stage, which consists of centering and whitening. Details are explained in related literature (Hyvärinen and Oja 2000). The process of parameter estimation employs maximum likelihood estimation.

## 4. Results and Discussion

### (1) Hypothesis

Prior to determining the temporal pattern of pollutants, we establish hypotheses for each zone to represent a unique temporal time frame pattern. Because the city center is located at the midpoint of the city and is surrounded not only by houses but also by government offices, we can hypothesize that the pattern in this zone is related to work and school activities, which are indicated by morning and evening peak periods. This zone is characterized by two daily cycle peaks during the week with the possibility of lower value activity on weekends or Sundays.

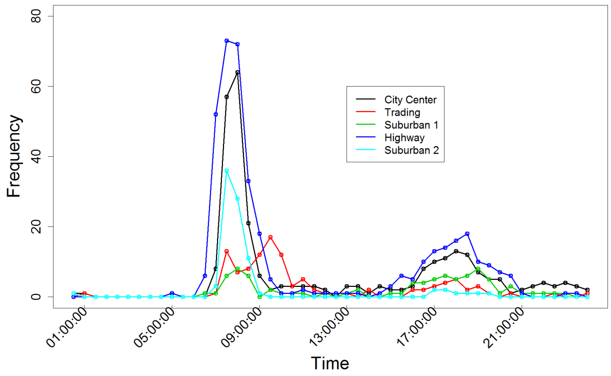
We assume that the pattern in the trading zone differs from the city center because the land use primarily involves fewer urban industries, business offices and this zone is located near the harbor, which is the second largest port harbor in Indonesia. Therefore, our hypotheses are that several short intervals during the day show peak concentrations which indicate that the movements of large vehicles may be concentrated in these intervals. The dominant emission source is a road that services the port harbor. In Suburban 1, which is located west of Surabaya, we assume a pattern that is similar to the pattern of Suburban 2 due to similar suburban locations. However, the monitoring station in Suburban 1 is surrounded by highly dense residential areas. The location in Suburban 2 is surrounded not only by houses but also by universities, businesses, and offices. In these two regions, we hypothesize that an extensive interval (wide interval or longer time frame) with a high concentration of $NO_2$ exists which is related to traffic volume.

In the zone near the highway, we assume earlier temporal(2morning peak time frame and delayed evening temporal peak time. Our hypothesis is based on the situation in which the highway connects Surabaya to neighboring cities. Therefore, the tendency for people to depart to work early in the morning and return from work late in the evening exists. Based on our previous description, our preliminary conclusion is that these zones may have unique pollutant characteristics based on the land use that each monitoring stations represents. In the next section, we analyze the course of pollutants when the permissible level is breached.

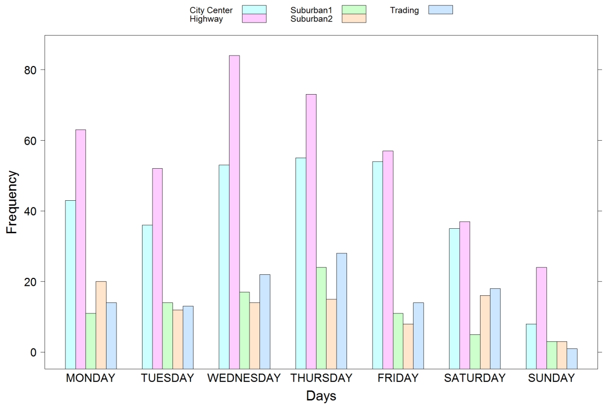
### (2) Temporal description when pollutants breach permissible levels

We observe some patterns for $NO_2$. For example, the frequency of events for $NO_2$ for all zones is high between 8am and 10am, with the exception of the trading zone (**Fig. 2(a)**). The time frame for observation during the evening session is 5pm to 9pm along the highway and city center, whereas an ambiguous pattern is observed for the remaining zones. **Fig. 2(a)** reveals limited information. For example, the concentration profile for $NO_2$ is similar among the four zones, with the exception of the trading zone. This result violates our hypothesis and is subject to additional investigation. A distinct temporal pattern for each location is also difficult to obtain. In the next discussion we attempt to answer the following question: when are these events dominant?

**Fig. 2(b)** shows the frequency of events on particular days of the week. For $NO_2$, the most frequent event occurs on Wednesday and Thursday. Lower risks are observed on the weekends. Note that these figures explain few temporal activities because they only depict the days in which the events occur for a two-year period. We also observe that events occur daily for both $NO_2$. In some zones, we observe low frequency, e.g., in the trading zone on Sunday ($NO_2$). This finding may violate commonly held beliefs—that weekends generate lower concentrations.

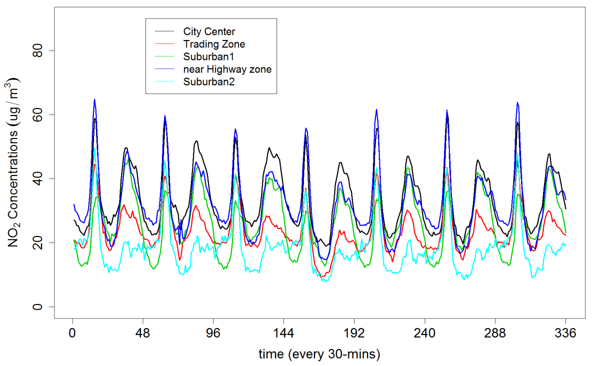



(a) Time of day




(b) Day of week

**Fig. 2** Frequency of pollutants that exceed the threshold based on different time frames (a) and days (b) for $NO_2$




Note: the first segment (48 time step) is Thursday

**Fig. 3** Weekly scale of daily diurnal of $NO_2$.

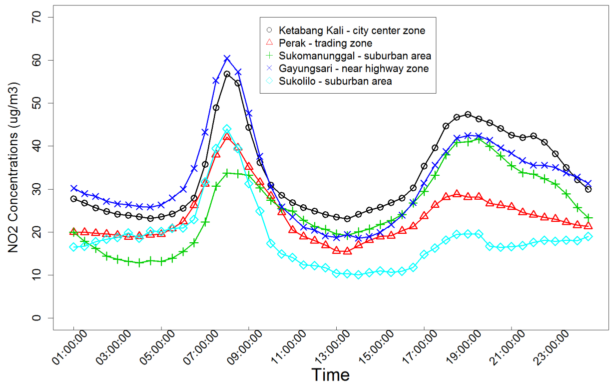



**Fig. 4** Diurnal concentration of $NO_2$

### (3) Diurnal variation in pollutant concentrations

To determine the temporal peak time frames, in which both concentrations may attain peaks and exceed permissible criteria, we averaged 30-min interval concentrations and constructed a weekly-scale schematic. **Fig. 3** shows the weekly diurnal variation for each pollutant from Thursday to Wednesday. As previously mentioned, the 30-min interval provided the most accurate description and business and industry activities varied by week. The diurnal cycle for $NO_2$ exhibits a double-wave shape, in which the morning peak exhibits a higher magnitude than the evening peak. The reduction in $NO_2$ and NO correlate with an increase in $O_3$ (Han, 2011). Because the cycle is similar for all days, we depict the daily diurnal on a daily-scale, as shown in **Fig. 4**.

**Fig. 4** shows the daily diurnal variations for $NO_2$ concentrations for the five monitoring stations. The two-peak distribution is similar to the distribution observed by Bigi and Harrison (2012) with a site in London and by Han (2011) with a site in Tian Jin, China. In our case, the $NO_2$ concentrations begin to increase at approximately 6am, attain a peak at 7:30 am, and subsequently decline. The intervals for the peak concentrations for $NO_2$ range between 6am and9am. These patterns may be explained by travel behavior because the intervals closely represent the morning peak traffic during the commute to school. In the evening, the concentrations begin to increase again at approximately 4:30pm, which parallels work hours. The peak concentration occurs at 6:30pm. Because workdays in the city usually end at 4pm, the commute traffic is likely to be the primary cause of the evening concentrations.

Observation of the magnitudes of the $NO_2$ concentrations for five stations reveals that the highest $NO_2$ concentration was observed the near-highway zone during the morning period, followed by the city center, suburban 2, trading zone, and suburban 1. For the evening session, a different order of the $NO_2$ concentration is observed, i.e., the highest $NO_2$ concentration was observed in the city center followed by the highway zone, suburban1, trading zone, and suburban 2. The concentrations of $NO_2$ in the trading zone are lower than the concentration in the city center, highway zone and suburban. A higher influence of traffic volume exists at the city center compared with the highway zone in the case of $NO_2$. The concentration profile for suburban 2 shows consistency over time, which is lower in both cases.

Although a temporal pattern is observed, similar patterns are observed for regions, the only difference is in terms of magnitude. Considering that the pollutants sources are unique for each location, for which we expect a different peak interval, therefore we should identify and confirm the temporal pattern using another method, which is described in the following section. We employ an ICA method for weekly average distributions to identify the different concentration distribution and temporal peak time frame between days (workdays and weekends).

### (4) Temporal peak time frame of pollutants using independent component analysis

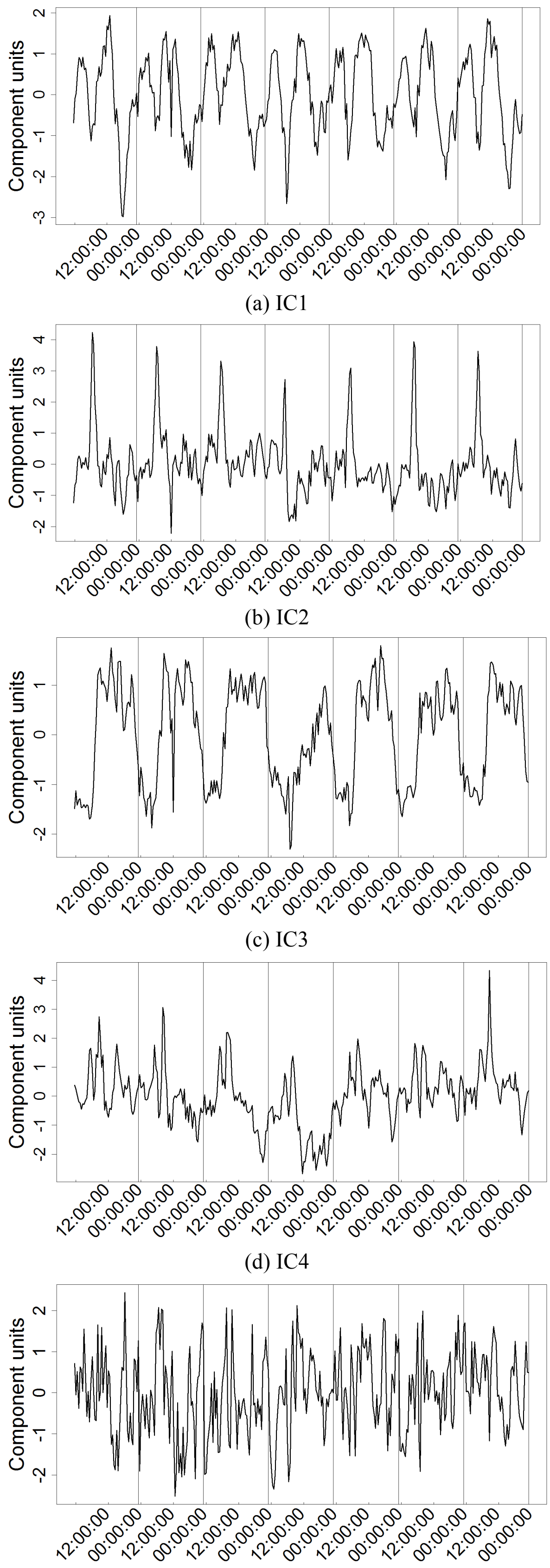
In this section, we display the ICA results, which consist of weight loadings for each component and figures that depict component units. We identified temporal patterns from these figures. We also determined the time frame in which the component units are high. If the weight loading has a negative value, the negative component units in the figure will reflect positive (higher) air pollutant concentrations.

**Fig. 5** shows the independent components (ICs) of the $NO_2$ concentration obtained by ICA, and **Table 3** summarizes the weight loadings for every station that comprises a component. IC1 represents the city center, near-highway, and suburban1, IC2 represents the city center, near-highway and suburban 2, IC3 represents suburban1, IC4 represents the trading zone, and IC5 represents suburban1 and suburban 2. The following section explains the temporal pattern for each station (zone).

$NO_2$ emissions are derived from traffic-related sources. Therefore, the pattern formed by the ICA (**Fig. 5**) may partially explain the traffic patterns. Based on the weight loadings (**Table 3**) for $NO_2$, the city center pattern can be characterized by IC1 and IC2. The ICs that explain the remaining zones are listed in **Table 3**. To explain the usefulness of an ICA and how we interpret the results from **Fig. 5**, we extract some important information regarding the concentration profile of air pollutant concentrations. A specific feature that we derive from the figure is that we can temporally determine an increase in concentration as shown in **Table 4**. Thus, we can identify the time frame in which the concentrations are high. Therefore, we focus on these particular time frames and propose a suitable policy for a particular zone based on the extracted time frame.

**Table 3** Weight loading for the independent components of $NO_2$ concentration

| Sites | IC1 | IC2 | IC3 | IC4 | IC5 |
|---|---|---|---|---|---|
| City center | **-7.639** | **5.404** | 3.164 | -0.357 | 0.041 |
| Trading zone | -3.741 | 4.546 | 0.949 | **2.912** | 0.044 |
| Suburban 1 | **-6.693** | 1.561 | **5.428** | 0.096 | **1.050** |
| Near highway | **-8.067** | **6.134** | -0.685 | 1.584 | 0.439 |
| Suburban2 | -2.817 | **6.237** | -2.527 | 1.844 | **1.293** |




(a) IC1
(b) IC2
(c) IC3
(d) IC4
(e) IC5

**Fig. 5** Five independent component profiles of the $NO_2$ concentration for five monitoring stations

**Table 4** Summary of $NO_2$ temporal peak time frame extracted from Independent Components

| Sites | IC1 | IC2 | IC3 | IC4 |
|---|---|---|---|---|
| City Center | Concentration grows up between 530am to 10am, and between 330pm to 12pm | Concentration grows up between 6am to 10am | | |
| Trading | | | | Concentration grows up between 9am-10am, anf its peak is in around 5pm except weekends |
| Suburban 1 | Concentration grows up between 530am to 10am, and between 330pm to 12pm | | Concentration grows up steadily between 8am to 11pm except Sunday | |
| Near Highway | Concentration grows up between 530am to 10am, and between 330pm to 12pm | Concentration grows up between 6am to 10am | | |
| Suburban 2 | | Concentration grows up between 6am to10am | | |

**Table 5** Cycle of temporal peak time frame extracted from the ICA

| Zones | Every day | Workdays | Monday to Saturday |
|---|---|---|---|
| City center | 5:30am to 10am<br>3:30pm to 12pm | | |
| Trading zone | 9am to 10am | approximately 5pm excluding weekends | |
| Suburban1 | 5:30am to 10am<br>3:30pm to 12pm | | 8am to 11pm |
| Near highway | 5:30am to 10am<br>3:30pm to 12pm | | |
| Suburban2 | 6am to10am | | |

Based on the derived information as shown on **Table 4**, we can estimate the profile and character for each IC. IC1 pattern forms double waves on morning and evening session which occurring every day. We estimate this component corresponds to traffic that is related to working and business activities. We observe this pattern is associated with three zones, they are city center, suburban1, and near highway zone. In those zones, the concentration is high between 5:30am and 10am (morning session), whereas high concentrations are detected between 3:30pm and 12pm, in the evening. In the city center, traffic on Gubernur Suryo Road is responsible for high concentrations because on this road, the traffic volume is high because it is protocol road where the Office of the Governor resides. Furthermore, monitoring station is situated near the business and shopping malls. Therefore, arranging traffic between these peak times will positively affect air quality. One possible suggestion is to limit distribution of vehicles between sessions e.g., motorbikes or large vehicles (trucks) are not allowed to pass the road. Additional suggestion is to change business hours e.g., change open hours for shopping mall.

In the near highway zone, as the highway connects between Surabaya and buffer cities as well as connecting to the east region of Province of Java, we can highly express that IC1 is dominated by working related traffic. During peak time intervals, there should be traffic management to reduce the risk of pollution exposure. Such preventive measures should be proposed. One suggestion that may be applied is to display pollution warning. In some points in the highway, there are displays that show congestion level reflecting several areas within the highway. If additional information e.g., prediction of pollutant level ($NO_2$) is added, this will be very beneficial for car users, they may act promptly either by choosing alternative exit route or wear a mask to reduce health risk. These proposals are stored and mentioned on **Table 6** below.

In the suburban 2, this pattern is not seen instead the dominant pattern was built by IC5. In suburban2 zone, the land use is occupied by mixed activity types. The residential area in the suburban2 is not dominant as in suburban1 zone where the land use is dominated by residential. In this region, there are few offices with several universities, and mixed business activities, therefore, the peak time frame is not tied strictly on two peak patterns as found in morning and evening. This character strengthens our assumption that IC1 dominantly represents pattern for working activities.

The pattern obtained from IC2 is similar with IC1 during morning session. IC2 gives temporal peak time frame between 6am to 10am every day and there is no peak observed during evening session which is different with IC1. This component pattern belongs to several zones they are city center, near highway, and suburban2 zone and it is not observed on trading zone and suburban1 zones. We estimate this pattern might come from traffic that is related to school commuting activities. The explanation of why there is no peak observed in the afternoon is that during the end school session, the distribution of traffic related to this activity is dispersed to other time frame. The dispersion is caused that some schools have different return time, usually in privately-owned schools. This pattern occurs in the highway can be explained as follows. The pattern found in this zone capture school commuting of people who make use of highway to transport within city.

Temporal peak time frame, which is given by IC3, is between 8am to 11pm. The pattern occurs on all days except Sunday. Therefore based on **Table 3**, this pattern reflects transportation pattern in Suburban 1. From this result, we are able to capture three patterns on Suburban1. First pattern (IC1) is working-related transport activities. We estimate the second pattern (IC3) come from routine activities of transport from residential inhabitants on the surrounding of monitoring station. This is because the station is inside an area that is dominated by houses, dwellings as well as some commercial buildings and apartments occupying a small portion of land use. The commuting and high volume of traffic flow from inhabitants caused the tendency of high concentrations from morning until night. The third pattern is explained by IC5.

The peak time frame as observed and reflected on IC4 exclusively occurs on trading zone. The monitoring station is located very near with Tanjung Perak Harbor, the second largest port in Indonesia. The road connecting center of city to the port is Tanjung Perak Barat and Tanjung Perak Timur. The pattern found in this independent component may dominantly reflect the activities that are related to the port, whereas other portion may be due to commercial and business activities. In the morning the peak is found to be within 9 to 10am indicating morning peak session and in the evening, the peak is observed on 5pm. Since this IC belongs only to trading zone, we assume that the time frame come from port activities. To make use of this information in order to improve air quality and based on information that dominant pollution sources come from the roads connecting the city to the port, we suggest similar proposal with city center that is to manage traffic flow either by restricting certain types of vehicles at those time frames. This also means time shifting of transport behavior related to port activities. One possibility is to apply restriction for large vehicles and forcing them moving to other time slots. The suggestion is shown and compiled in **Table 6**.

In addition to those mentioned ICs, we observe scrambled pattern as reflected on IC5. This IC cannot be easily interpreted and may lead us to a conclusion that there are four traffic related-air quality patterns extracted from five monitoring station in Surabaya. From the discussion, we may produce seasonal traffic pattern on weekly scale as shown by **Table 5**.

**Table 6** Possible policies expected to reduce $NO_2$ by considering peak time frame from each zone

| | |
|---|---|
| City Center | - Traffic flow control on Gubernur Suryo Road between 6am to 10am, e.g., no large vehicles are allowed<br>- Staggered business hours<br>- Traffic route management |
| Trading | - Traffic flow control by limiting vehicles between 7am and 10am daily on the road near the station<br>- Traffic flow control at 5pm excluding weekends |
| Suburban 1 | - Staggered business hours at some shopping malls |
| Near-highway | - Air quality level display (e.g., $NO_2$) during peak periods to promote alternative route choice |
| Suburban 2 | - Air quality display to avoid this area |

## 5. Conclusions

This study investigates the events in which $NO_2$ concentrations exceed the permissible level in Surabaya and the importance of obtaining temporal peak time frames using a general statistical method and an ICA. Generally, the concentrations of $NO_2$ increase during peak hours in the morning and evening. If we use a general statistical method that employs a daily diurnal analysis, the temporal pattern is similar in all five monitoring sites. The only difference is the magnitude. Because the source of emissions differs among stations, the unique temporal peak time frame for each station should be reflected, which cannot be extracted from the diurnal pattern; therefore, we employ an ICA.

The use of ICA has been useful for the creation of independent $NO_2$ temporal patterns in each zone. The profile is characterized by similar concentration profiles in the morning for the city center, suburban1, and near highway zones, which suggests similar traffic activities e.g., working activities. In the region in which the near highway zone is located, an increase in $NO_2$ concentrations occurs due to transportation to/from and within Surabaya, which indicates that the regions in the vicinity of a highway may experience high concentrations of $NO_2$.

In this study, we have demonstrated how we can derive seasonal and daily temporal peak time frames for each zone. This information may be used to formulate policies to reduce human health risks due to exposure to pollutants. An important step from this information is that we can formulate traffic-related policies for these particular time frames for each zone. Traffic-related policies may involve vehicle restrictions during peak hours (direct impact) or indirect policies, e.g., installation of a warning system to facilitate the adaptation of user behavior.

ICA can be expanded to an air quality management system, such as the evaluation of the number of monitoring stations and the prediction method.

## References


Abdul-wahab SA, Bakheit CS and Al-Alawi SM (2005) Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations, *Environmental Modelling and Software* 20, 1263–1271, doi:10.1016/j.envsoft.2004.09.001

Al-Alawi SM, Abdul-Wahab SA and Bakheit CS (2008) Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone, *Environmental Modelling and Software* 23, 396–403, doi:10.1016/j.envsoft.2006.08.007

Back A and Weigend A (1997) Discovering structure in finance using independent component analysis, *Proceedings of 5th International Conference on Neural Networks in Capital Market*, Kluwer Academic, 15–17.

Bigi A and Harrison RM (2012) Analysis of the air pollution climate at a central urban background site, *Atmospheric Environment* 44(16), 2004–2012, doi:10.1016/j.atmosenv.2010.02.028

Cheung Y and Xu L (2001) Independent component ordering in ICA time series analysis, *Neurocomputing* 41, 145–152.

Han S (2011) Analysis of the Relationship between $O_3$, NO and $NO_2$ in Tianjin, China, *Aerosol and Air Qual Research* 2, 128–139, doi:10.4209/aaqr.2010.07.0055

Hyvärinen A and Oja E (2000) Independent component analysis: algorithms and applications, *Neural networks : the official journal of the International Neural Network Society,* 13(4-5), 411–30.

Long F, Wu T, Movellan, JR, Bartlett MS and Littlewort G (2012) Learning spatiotemporal features by using independent component analysis with application to facial expression recognition, *Neurocomputing* 93, 126–132, doi:10.1016/j.neucom.2012.04.017

Lu C-J, Lee T-S and Chiu C-C (2009) Financial time series forecasting using independent component analysis and support vector regression, *Decision Support Systems* 47(2), 115–125, doi:10.1016/j.dss.2009.02.001

Minguillón MC, Monfort E, Escrig A, Celades I, Guerra L, Busani G, Sterni A and Querol X (2013). Air quality comparison between two European ceramic tile clusters, *Atmospheric Environment* 74, 311–319, doi:10.1016/j.atmosenv.2013.04.010

Pires JCM, Sousa SIV, Pereira MC, Alvim-Ferraz MCM and Martins FG (2008a) Management of air quality monitoring using principal component and cluster analysis—Part I: SO2 and PM10, *Atmospheric Environment* 42(6), 1249–1260, doi:10.1016/j.atmosenv.2007.10.044

Pires JCM, Sousa SIV, Pereira MC, Alvim-Ferraz MCM and Martins FG (2008b). Management of air quality monitoring using principal component and cluster analysis—Part II: CO, $NO_2$ and $O_3$, *Atmospheric Environment* 42(6), 1261–1274, doi:10.1016/j.atmosenv.2007.10.041

Pires JCM, Martins FG, Sousa SIV, Alvim-Ferraz MCM and Pereira MC (2008c) Prediction of the Daily Mean $PM_{10}$ Concentrations Using Linear Models, *American Journal of Environmental Sciences* 4(5), 445–453.

Syafei AD, Fujiwara A and Zhang J (2013) A comparative study on NO concentration interpolation in Surabaya City, *Proceedings of the Eastern Asia Society for Transportation Studies*, Vol. 9

Westra S, Brown C, Lall U, Koch I and Sharma A (2009) Interpreting variability in global SST data using independent component analysis and principal component analysis, *International Journal of Climatology* 30, 333-346, doi:10.1002/joc.1888

Wu EHC and Yu PLH (2005) Independent Component Analysis for Clustering Multivariate Time Series Data. As Book Chapter of Advanced Data Mining and Applications, *First International Conference of ADMA* 2005, Wuhan, China, July 22-24, 474-482.



**(Received April 24, 2015)**