WORKING PAPER

MODELING TRAFFIC CONGESTION EFFECTS ON AIR POLLUTANTS

ESCAP 75

MOVING FORWARD TOGETHER

*The shaded areas of the map indicate ESCAP members and associate members.**

Acknowledgement: This policy brief on "Modelling Traffic Congestion Effects on Air Pollutants" has been issued in January 2023. Preparation of this paper is coordinated by the Economic and Social Commission for Asia and the Pacific, through its Environment and Development Division, under the supervision of Sangmin Nam, Director, Environment and Development Division, Curt Garrigan, Chief, Environment and Development Policy Section and authorship by Matthew Perkins, Environmental Affairs Officer, Divya Nair and Vinayak Dixit.

This document also benefitted from contributions by Abigail Smith and Mervin Chin, whose inputs were integral to the production of this report. Financial support by the Republic of Korea is gratefully acknowledged.

For further information on this policy brief, please address your inquiries to: Environment and Development Division United Nations Economic and Social Commission for Asia and the Pacific (ESCAP) Email: ESCAP-EDD@un.org

** The designations employed and the presentation of material on this map do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.*

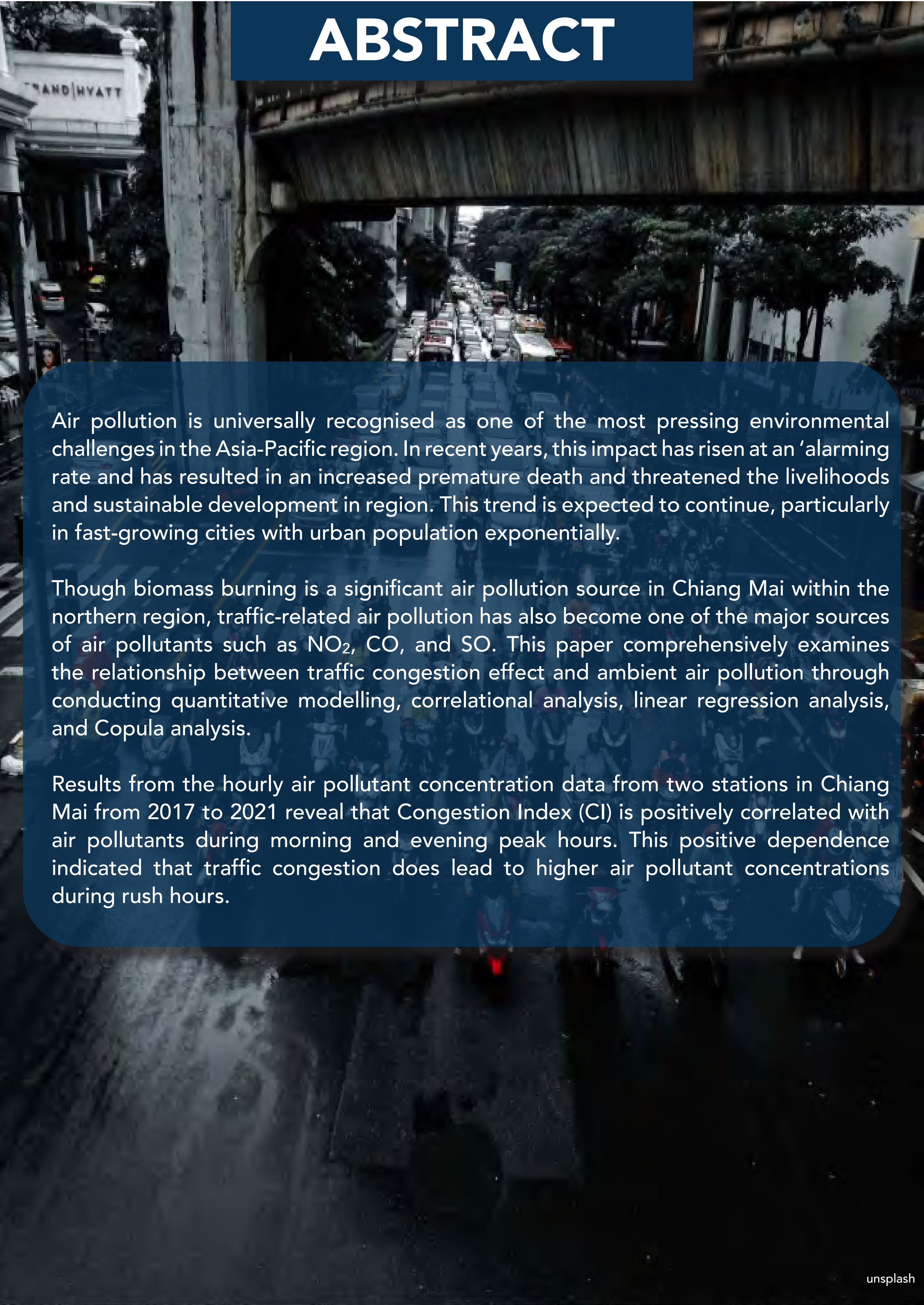
Table of Contents

ABSTRACT

Air pollution is universally recognised as one of the most pressing environmental challenges in the Asia-Pacific region. In recent years, this impact has risen at an ‘alarming rate and has resulted in an increased premature death and threatened the livelihoods and sustainable development in region. This trend is expected to continue, particularly in fast-growing cities with urban population exponentially.

Though biomass burning is a significant air pollution source in Chiang Mai within the northern region, traffic-related air pollution has also become one of the major sources of air pollutants such as NO_2, CO, and SO. This paper comprehensively examines the relationship between traffic congestion effect and ambient air pollution through conducting quantitative modelling, correlational analysis, linear regression analysis, and Copula analysis.

Results from the hourly air pollutant concentration data from two stations in Chiang Mai from 2017 to 2021 reveal that Congestion Index (CI) is positively correlated with air pollutants during morning and evening peak hours. This positive dependence indicated that traffic congestion does lead to higher air pollutant concentrations during rush hours.



unsplash

METHODS

Study Area

Chiang Mai is the second largest city in Thailand. The Thailand Meteorological Department[1] divides the climate into three seasons: hot (summer) season (March to May), rainy season (May to October), and cool (winter) season (November to February). There are two rush hour periods (7:00 – 8:00a.m. and 5:00 – 6:00p.m.). It is found that the traffic contributed approximately 90 per cent of total CO emissions during rush hours[2].

DATA

Road Network

The road network of Chiang Mai including the latitude and longitude coordinates of each link, speed limit, length, name, highway types is obtained from OpenStreetMap (OSM) data sets[3]. There are seven main highway types in OSM, namely motorway, trunk, primary, secondary, tertiary, unclassified, and residential. The process of acquiring road networks data is detailed in the Appendix download_network.py.

Figure 1: The road network of study area (Chiang Mai)

Speed Data

The speed data of Chiang Mai is collected from Intelligent Traffic Information Center (iITC). The probe-data provides vehicle unique ID, latitude and longitude coordinates, time stamp and speed from 2017 to 2021. By using the vehicle latitude and longitude coordinates, we can determine which road the vehicle is on and the speed limit of the road.

The example of speed data:

```
wAwKpIfvv077zeaMM+nbSi09bHg, 1, 13.72878, 100.49035, 2016-12-31 23:59:27, 0, 331, 0, 0
hgzoCSq3DmCzsq7gjDkL1vWmuwc, 1, 13.93083, 100.60743, 2016-12-31 23:59:40, 48, 27, 0, 1
eknoba7CnGxjhTpmq531Im/Rx7s, 1, 13.79902, 100.36575, 2016-12-31 23:58:56, 0, 263, 1, 0

Fields description:

VehicleID, gpsvalid, lat, lon, timestamp, speed, heading, for_hire_light, engine_acc
```

Air Pollutants

We collected hourly air pollutant concentration data from two stations in Chiang Mai from 2017 to 2021 (Figure 2). The atmospheric pollutants are $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, O_3.

Table 1 The example of air pollutant data

(cont'd on next page)

Date	PM2.5	PM10	O3	CO	NO2	SO2	station
1/01/2017 0:00		38		1.1	11	1	36t
1/01/2017 1:00		37		1	12	1	36t
1/01/2017 2:00		44		1.3	12	1	36t
1/01/2017 3:00		39		1.1	9	1	36t
1/01/2017 5:00		35		1	15	1	36t
1/01/2017 6:00		44		1	12	1	36t
1/01/2017 7:00		33		0.8	8	1	36t
1/01/2017 8:00		13		0.9	12	1	36t
1/01/2017 9:00		25		1	15	1	36t
1/01/2017 10:00		30		1.1	16	1	36t
1/01/2017 11:00		18		1	13	1	36t
1/01/2017 12:00		22		1	10	1	36t
1/01/2017 13:00		29		0.8	5	1	36t
1/01/2017 14:00		26		0.8	5	1	36t
1/01/2017 15:00		30		0.8	5	1	36t
1/01/2017 16:00		27		0.8	6	1	36t
1/01/2017 17:00		27		0.9	7	1	36t
1/01/2017 18:00		23		0.9	8	1	36t
1/01/2017 20:00		16		0.9	8	1	36t
1/01/2017 21:00		15		0.8	5	1	36t
1/01/2017 22:00		11		0.8	5	1	36t
1/01/2017 23:00		14		0.8	4	1	36t
2/01/2017 0:00		19		0.8	3	1	36t
2/01/2017 1:00		14		0.8	0	1	36t
2/01/2017 2:00		15		0.8	0	1	36t
2/01/2017 3:00		15		0.8	1	1	36t
2/01/2017 4:00		16		0.8	1	1	36t
2/01/2017 6:00		19		0.9	9	1	36t
2/01/2017 7:00		24		0.9	7	1	36t
[illegible]		[illegible]		[illegible]	[illegible]	[illegible]	[illegible]

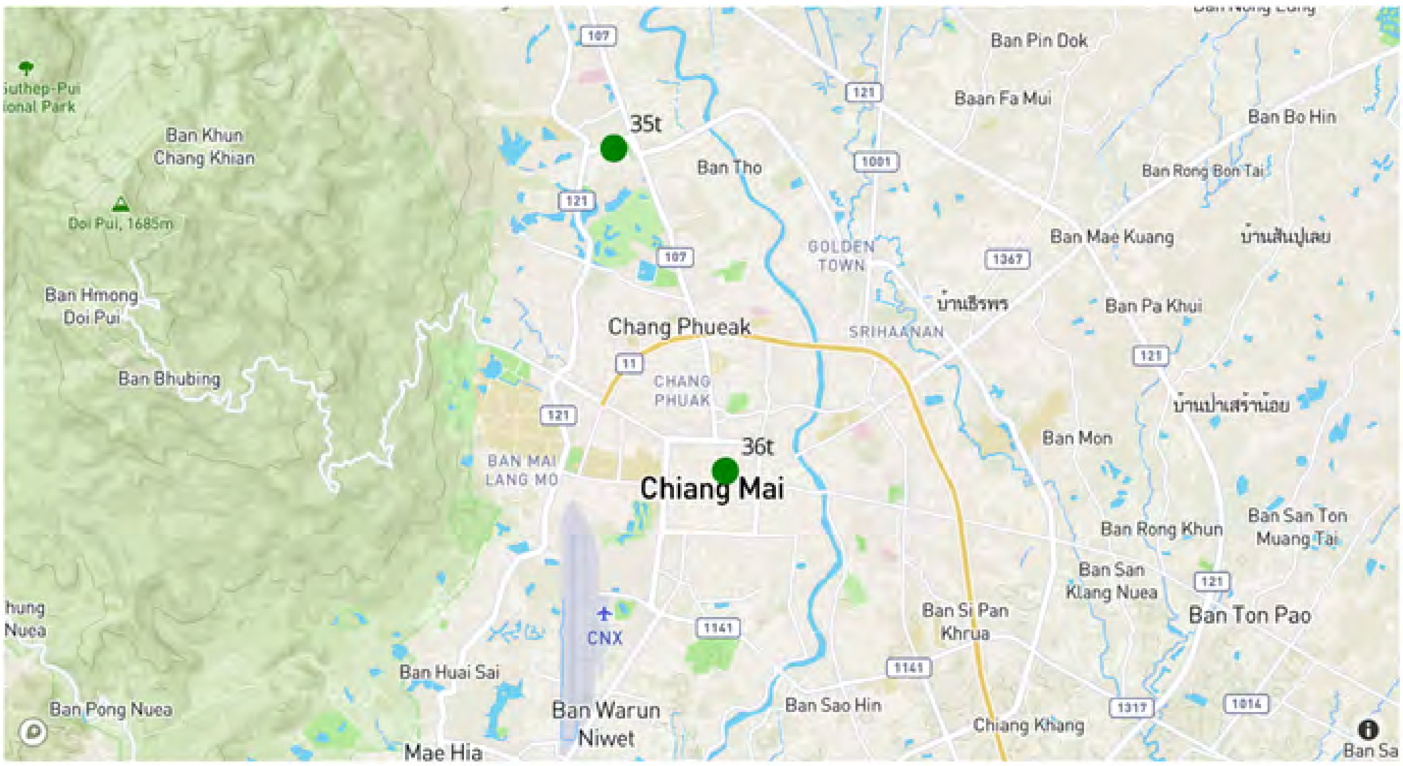


Figure 2: The locations of air quality monitoring stations in Chiang Mai

MODELING

Work Flow

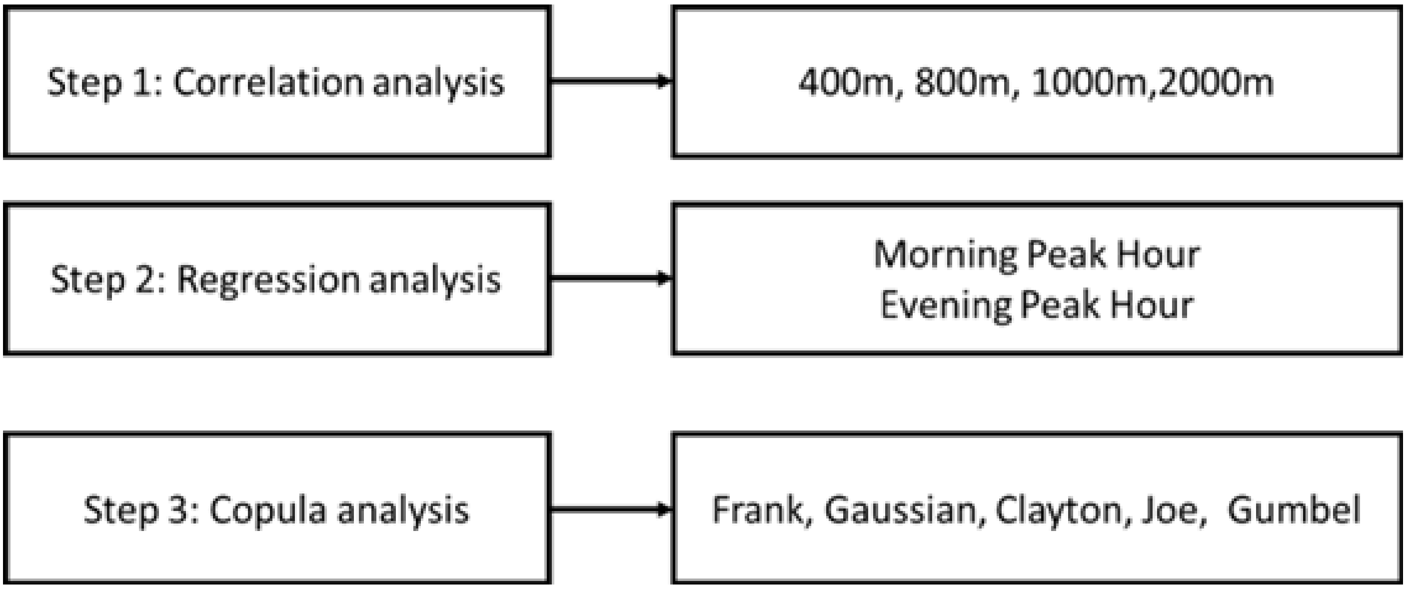


Figure 3: The flow chart of the CI and air pollutants analysis

Firstly, as shown in Figure 3, the correlation analysis is conducted to find the relationship between CI and air pollutants for 24 hours a day of each month. The results indicate that during the morning and evening peak hours, most air pollutants are positively correlated with CI, i.e., the more congested the road, the higher the air pollution concentration. During the off-peak period, most air pollutants and CI become weakly positively or negatively correlated.

Therefore, in the second step, linear regression is applied only to the peak period between air pollutants and CI. We find that the during peak hours, traffic congestion does lead to higher air pollutant concentrations, but the R^2 values are small.

Finally, the Copula functions are used to find the dependency between CI and air pollutants during peak hours for each group each month.

Congestion Index

Congestion Index (CI)[4] is adopted to represent the beginning of the delay by comparing the current traffic speed and the free-flow speed. The free-flow speed (V_f) is defined as the maximum speed during the off-peak period. As shown in Equation (1), the congestion index of a link *i* at time *t* (CI_{it}) is computed as the ratio of its delay $(\delta_{it} = \frac{1}{V_{it}} - \frac{1}{V_{fit}})$ and the free flow travel time/km.

$$CI_{it} = \frac{\delta_{it}}{\frac{1}{V_{fit}}}$$

Table 2 The example of CI data

V_{it}	V_{fit}	CI_{it}
10	100	9
15	100	5.67
20	100	4
30	100	2.33
40	100	1.5
50	100	1
60	100	0.67
70	100	0.43
80	100	0.25
90	100	0.11
100	100	0

Because the station-level CI is calculated by averaging the CI of the links around the station, the distance from the station to the link significantly affects the value of the station-level CI. We set the distance from the station to the link as 400m, 800m, 1000m and 2000m, and the correlation analysis results are used to determine the appropriate distance.

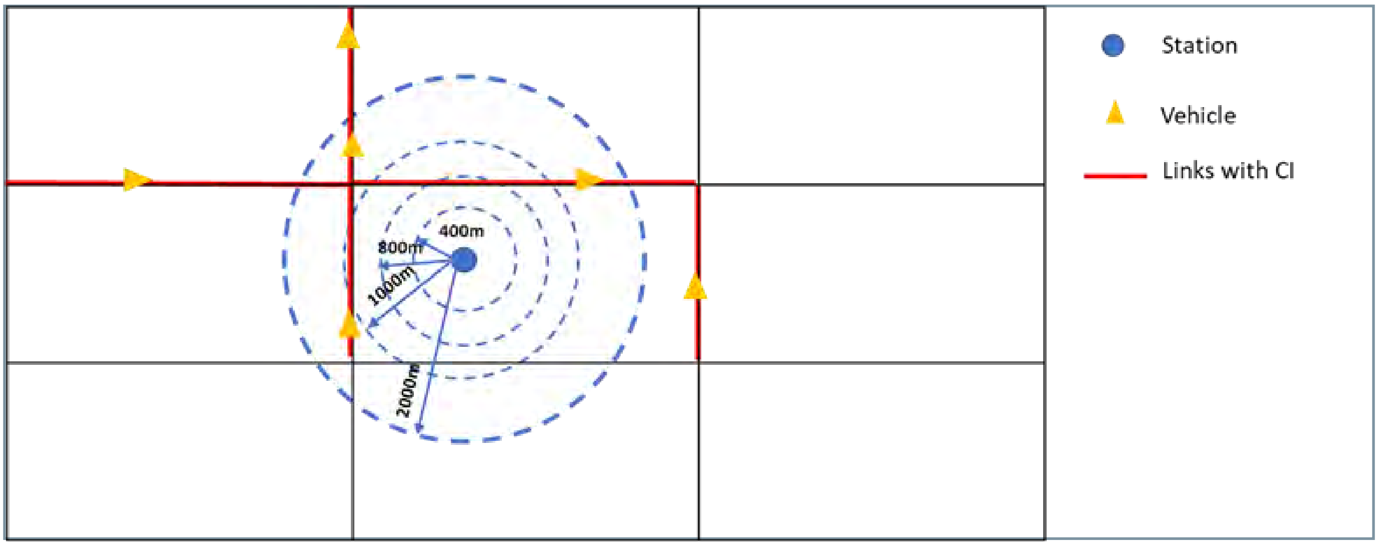


Figure 4: Different distances between stations and links

Congestion Index

The correlations between CI and air pollutants are investigated by using Pearson correlation coefficient (Equation(2)). Correlation analysis is carried out with hourly air pollutant concentrations and CI for each month.

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2}\,(y_i - \bar{y})^2}$$

where $\bar{x}$ and $\bar{y}$ are the means of *x* and *y* variables.

Linear Regression

The linear regression between CI and air pollutants are used in this study. The linear regression is conducted with rush hour air pollutant concentrations and CI for each month, and *R^2* is used to measure the fitness.

Copula Functions

Copulas[5] are functions that associates multivariate distribution functions of random variables with their one-dimensional marginal distribution functions. The dependence structures are measured by correlation coefficients: Pearson's product-moment correlation coefficient, Kendall's tau, and Spearman's rho. Consider a d-dimensional cumulative distribution function (CDF), *F*, with marginals $F_1, \ldots, F_d$. Then there exists a copula, *C*, such that

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d))$$

for all $x_i \epsilon [-\infty, \infty]$ and $i = 1, \ldots, d$.

The marginal distributions of air pollutants and CI are fitted by the empirical cumulative distribution function (CDF), i.e., the data (u_1,u_2) are transformed as the uniformly distributed random variables with support on [0,1].

The proposed study uses Gaussian, Clayton, Frank, Gumbel and Joe copulas functions. Then, the best fitted copula is selected based on the log-likelihood and Akaike information criterion (AIC).

Bivariate Gaussian copula
The bivariate Gaussian copula is one of the most commonly used elliptical copulas. The function with correlation θ function is given by:

$$C(u_1, u_2; \theta) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\theta^2}} \exp\left(\frac{s^2 - 2\theta st + t^2}{2(1-\theta^2)}\right) dsdt$$

where Φ is the standard normal distribution and θ is the Pearson's liner correlation coefficient.

The Frank copula
The Frank copula function[6] is defined as:

$$C_\theta(u_1, u_2) = -\frac{1}{\theta} \ln\left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right)$$

The generator function is expressed as $\varphi(t) = -\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$. The correlation between the random variables can be positive and negative.

The Clayton copula

The Clayton copula function[7] is defined as:

$$C_\theta(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}$$

The generator function is expressed as $\varphi(t) = \frac{1}{\theta}(t^{-\theta} - 1)$. The Clayton copula only accounts for the positive dependence, especially for the lower tail of the random variables.

The Gumbel copula

The Gumbel copula function[8] is expressed as:

$$C_\theta(u_1, u_2) = \exp\left(-[(-\ln u_1)^\theta + (-\ln u_2)^\theta]^{1/\theta}\right)$$

The generator function is expressed as $\varphi(t) = (-\ln t)^\theta$. Compared with the Clayton copula, The Gumbel copula only account for the positive dependence, but especially for the upper tail of the random variables.

The Joe copula

The Joe copula function is proposed by Joe[9,10] with the generator function $\varphi(t) = -\ln[1 - (1-t)^\theta]$ and the function is expressed as:

$$C_\theta(u_1, u_2) = 1 - [(1-u_1)^\theta + (1-v_1)^\theta - (1-u_1)^\theta(1-v_1)^\theta]^{1/\theta}$$

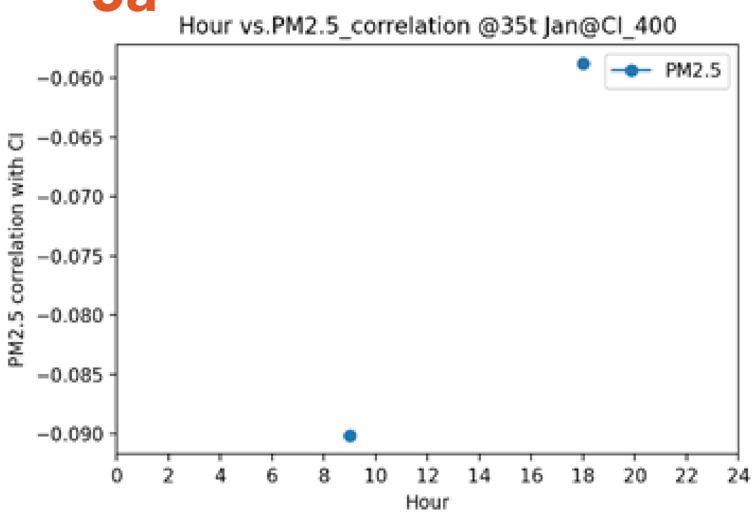
Similar with Gumbel copula, the Joe copula only accommodates positive dependence especially for the upper tail of the random variables.

RESULTS

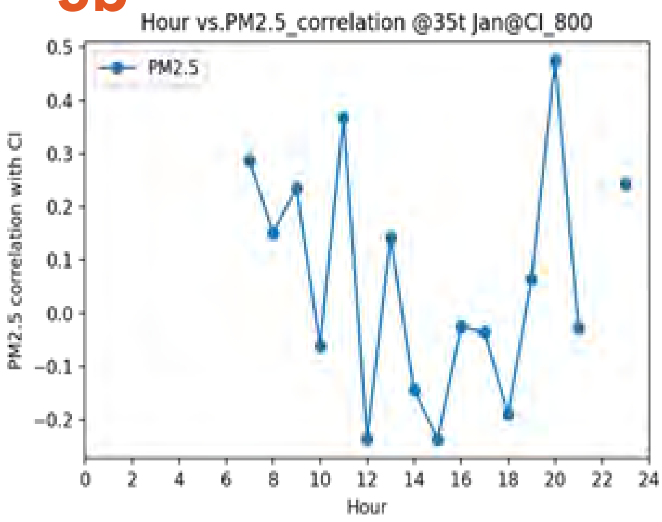
Correlation Analysis Results

The correlation coefficient is used to test the correlation between individual air pollutants ($PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, O_3, CO) and CI on an hourly basis per month. By comparing the different distances from the station to the links (400m, 800m, 1000m, 2000m), we found that 1000m is a suitable threshold.

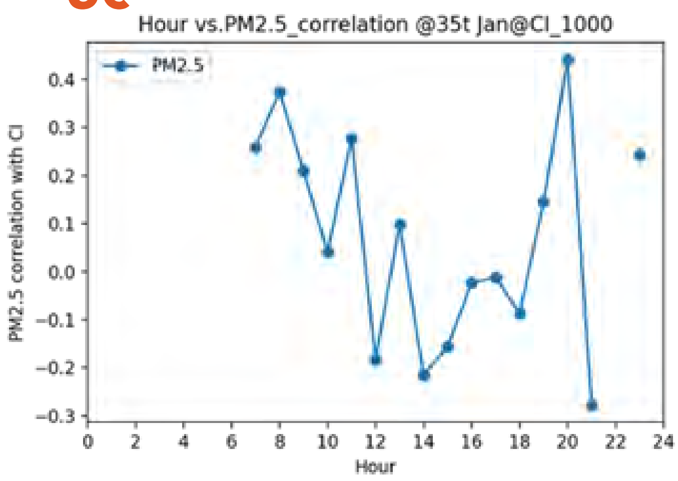
5a



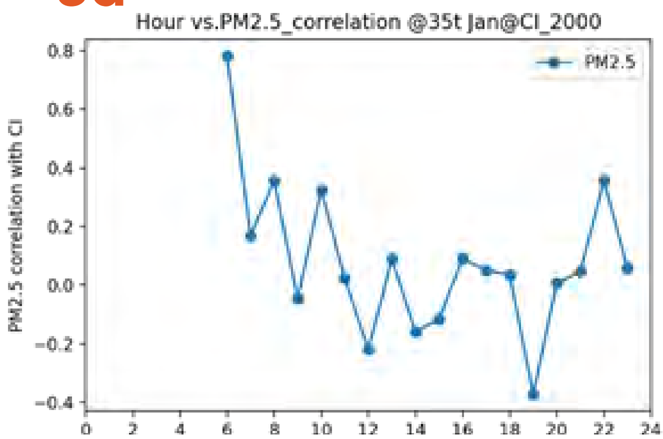
5b



5c



5d



Figures 5a, 5b, 5c, 5d: Correlation results

Notes: The correlation results with different distances between station and links (a) 400m, (b) 800m, (c) 1000m, (d) 2000m

As shown in Figure 5, except for Figure 5a, all plots show roughly two peaks of correlation coefficient between $PM_{2.5}$ and CI, (b) at 10am and 20pm, (c) at 8am and 20pm, (d) at 6am and 22pm. Since the morning and evening peak time periods are 7am - 8am and 17pm - 18pm, the correlation between CI and $PM_{2.5}$ is better represented in Figure 5c.

In Figure 5d, there is an unreasonable fluctuation at 18 pm - 20pm (0.11 at 18pm, -0.4 at 19pm, 0.09 at 20pm), probably due to the long setting distance, the CI of other links affects the value of 35t station setting. The distance of 1000m between station and links is selected in the following analysis.

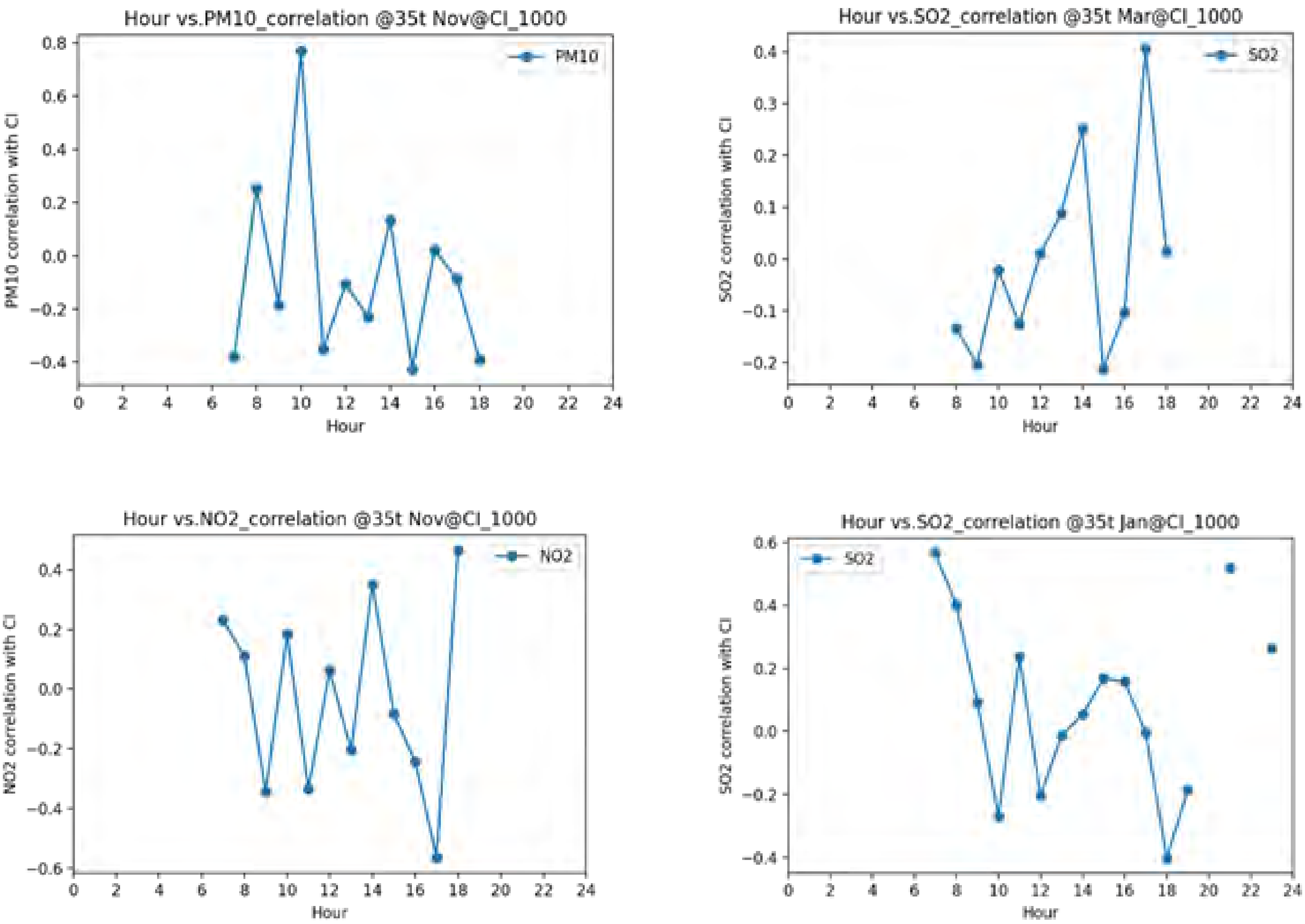


Figure 6 The correlation analysis results for different months at different stations

As shown in the Figure 6, the result show that CI and air pollutants are positively correlated during peak hours and are not correlated or even negatively correlated during off-peak hours. More correlation analysis results can be found in the Appendix.

The results confirm our intuition that the more congested the traffic is, the more air pollutants are emitted by vehicles. Therefore, we did a linear regression analysis of CI and air pollutants in the morning and evening peak periods.

Linear Regression Results

Morning Peak Hours

The morning peak hour is set to 7am to 8am. As shown in Figure 7, the air pollutants such as PM_{10}, O_3, $PM_{2.5}$ are showing the positive correlation with CI. The range of R-squared is 0.02 to 0.06, indicating that the linear relationship between air pollutants and CI is insignificant, which may be mainly due to the lack of data from other regional stations and the influence of weather.

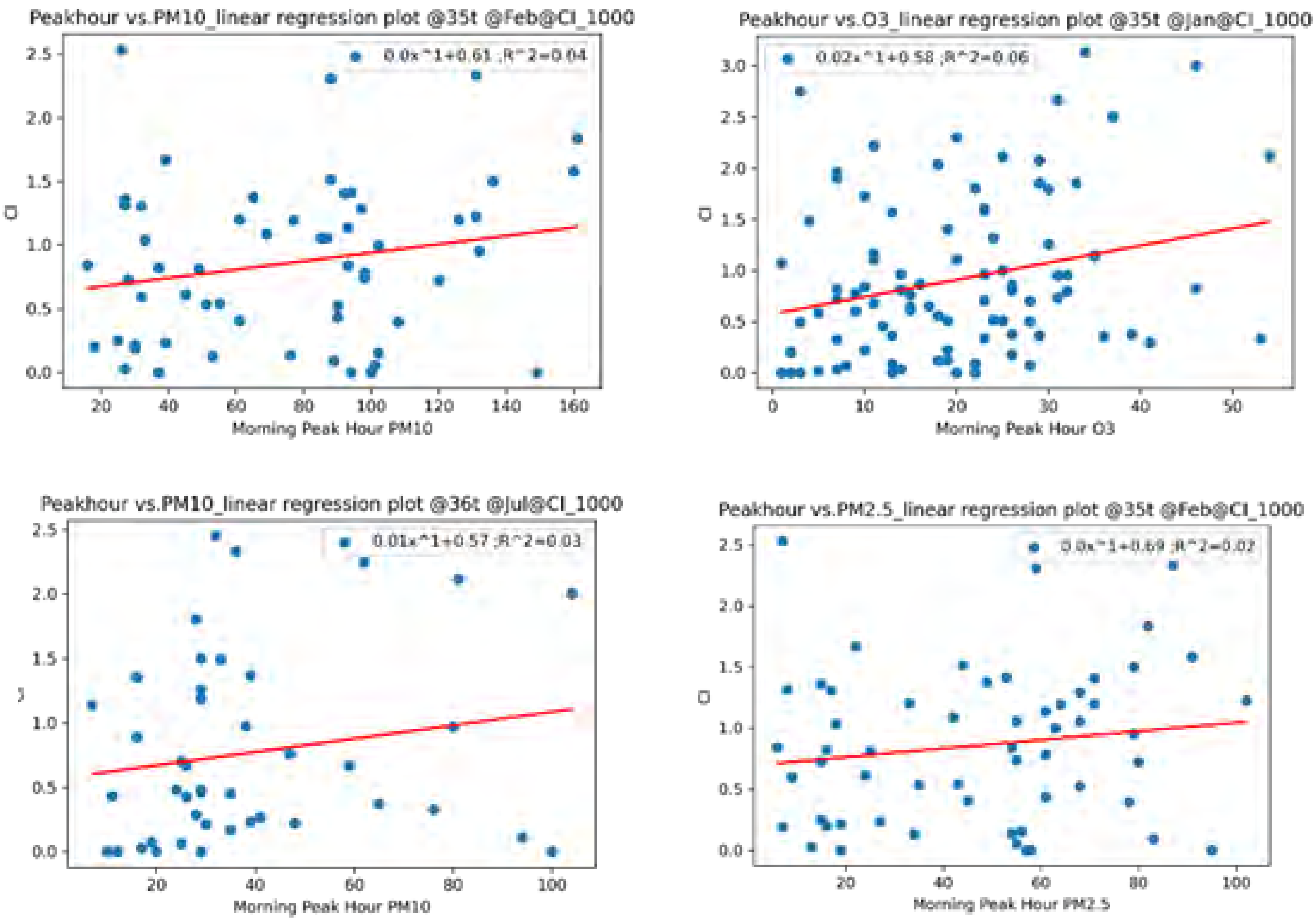


Figure 7: The morning peak hour linear regression results

Evening Peak Hours

The evening peak hour is set to 17pm to 18pm. As shown in Figure 8, the air pollutants such as $PM_{2.5}$, PM_{10} and NO_2 are showing the positive correlation with CI.

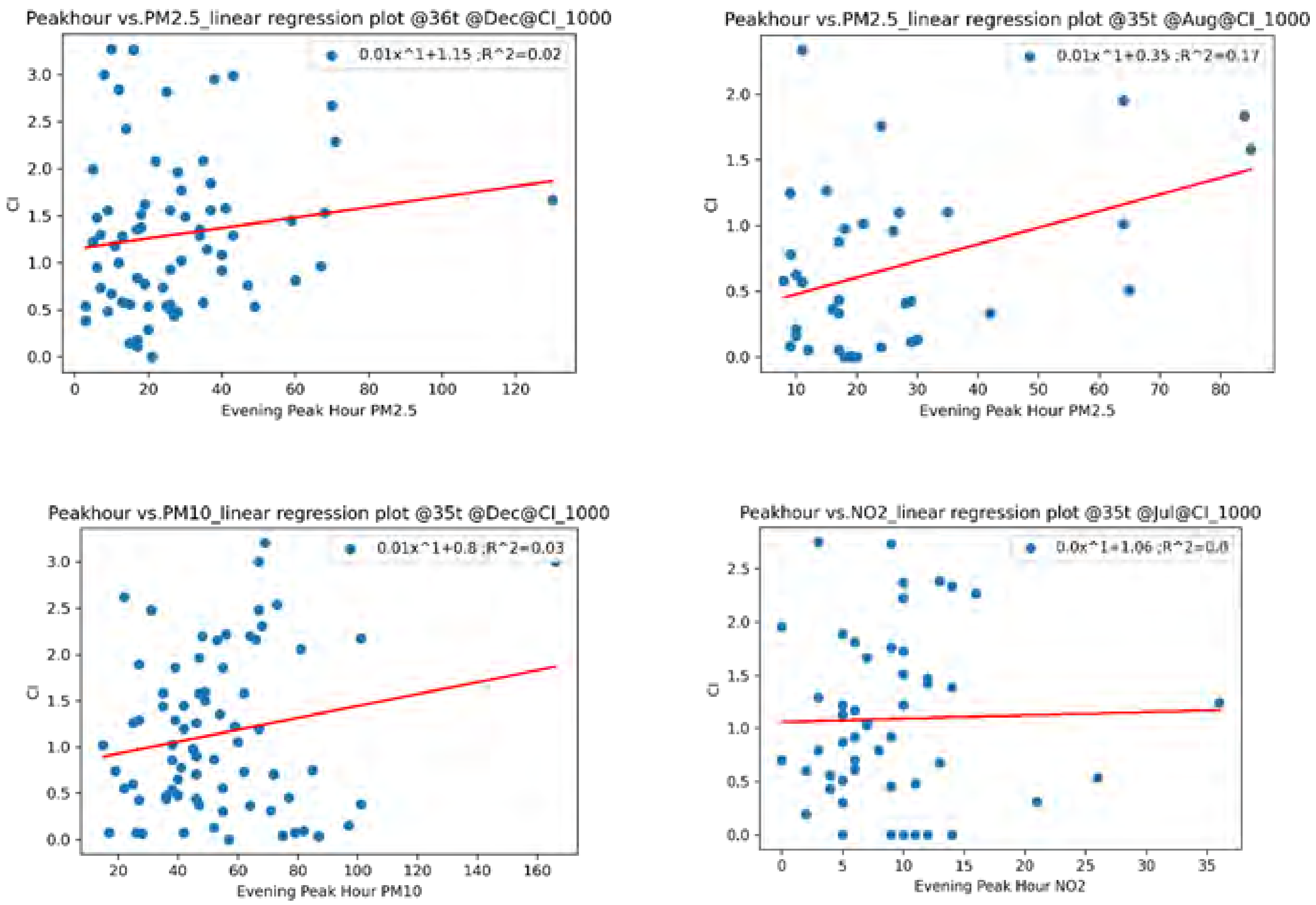


Figure 8: The evening peak hour linear regression results

Copula Results

Copulas are functions that associates multivariate distribution functions of random variables with their one-dimensional marginal distribution functions. The marginal distributions of air pollutants and CI are fitted by the empirical cumulative distribution function (CDF), i.e., the data (u_1,u_2) are transformed as the uniformly distributed random variables with support on [0,1].

As shown in Figure 9, the first two plots are scatterplots of the original data being transformed into a uniform distribution and the fitted scatterplot, respectively, and the third is the isolines of probability density of the empirical copula functions and the estimated ones. Taking CI and $PM_{2.5}$ as an example, the results indicate that there is no significant relationship between CI and air pollutant.

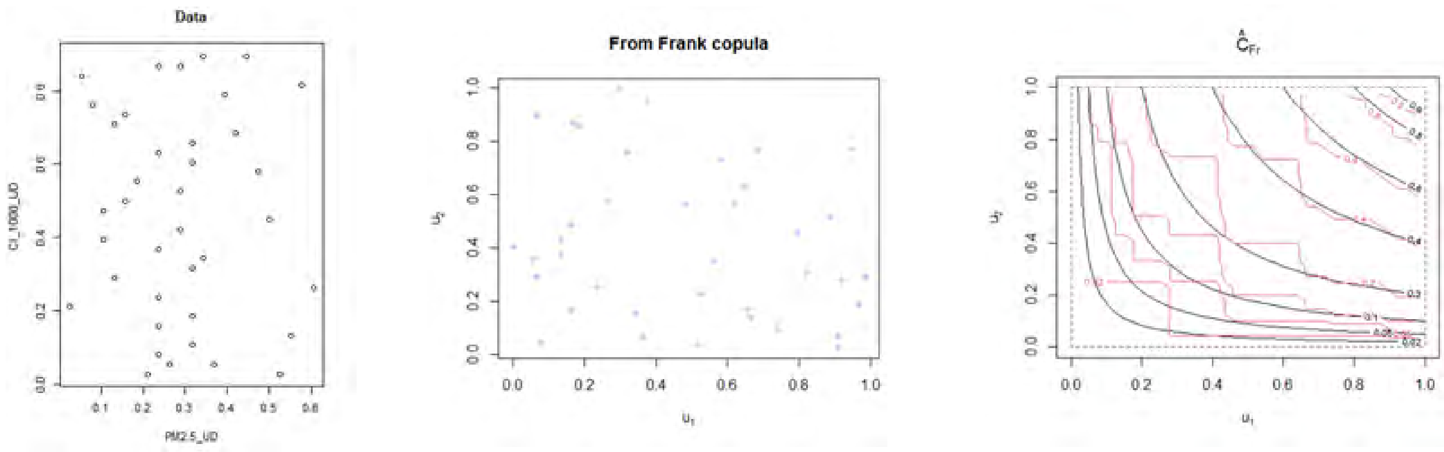


Figure 9: Copula results

References

[1] TMD, “The Climate of Thailand.” 2015. [Online]. Available: https://www.tmd.go.th/en/arch ive/thailand_climate.pdf

[2] S. Sathitkunarat, P. Wongwises, R. Pan-Aram, and M. Zhang, “Carbon monoxide emission and concentration models for Chiang Mai urban area,” Advances in Atmospheric Sciences, vol. 23, no. 6, pp. 901–908, 2006.

[3] O. Contributors, “OpenStreetMap. 2019,” URL http://planet. openstreetmap. org.[Online, 2019.

[4] D. J. Nair, F. Gilles, S. Chand, N. Saxena, and V. Dixit, “Characterizing multicity urban traffic conditions using crowdsourced data,” PLoS ONE, vol. 14, no. 3, p. e0212845, Mar. 2019, doi: 10.1371/journal.pone.0212845.

[5] M. Sklar, “Fonctions de repartition an dimensions et leurs marges,” Publ. inst. statist. univ. Paris, vol. 8, pp. 229–231, 1959.

[6] M. J. Frank, “On the simultaneous associativity ofF (x, y) andx+y− F (x, y),” Aequationes mathematicae, vol. 19, no. 1, pp. 194–226, 1979.

[7] D. G. Clayton, “A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence,” Biometrika, vol. 65, no. 1, pp. 141–151, 1978.

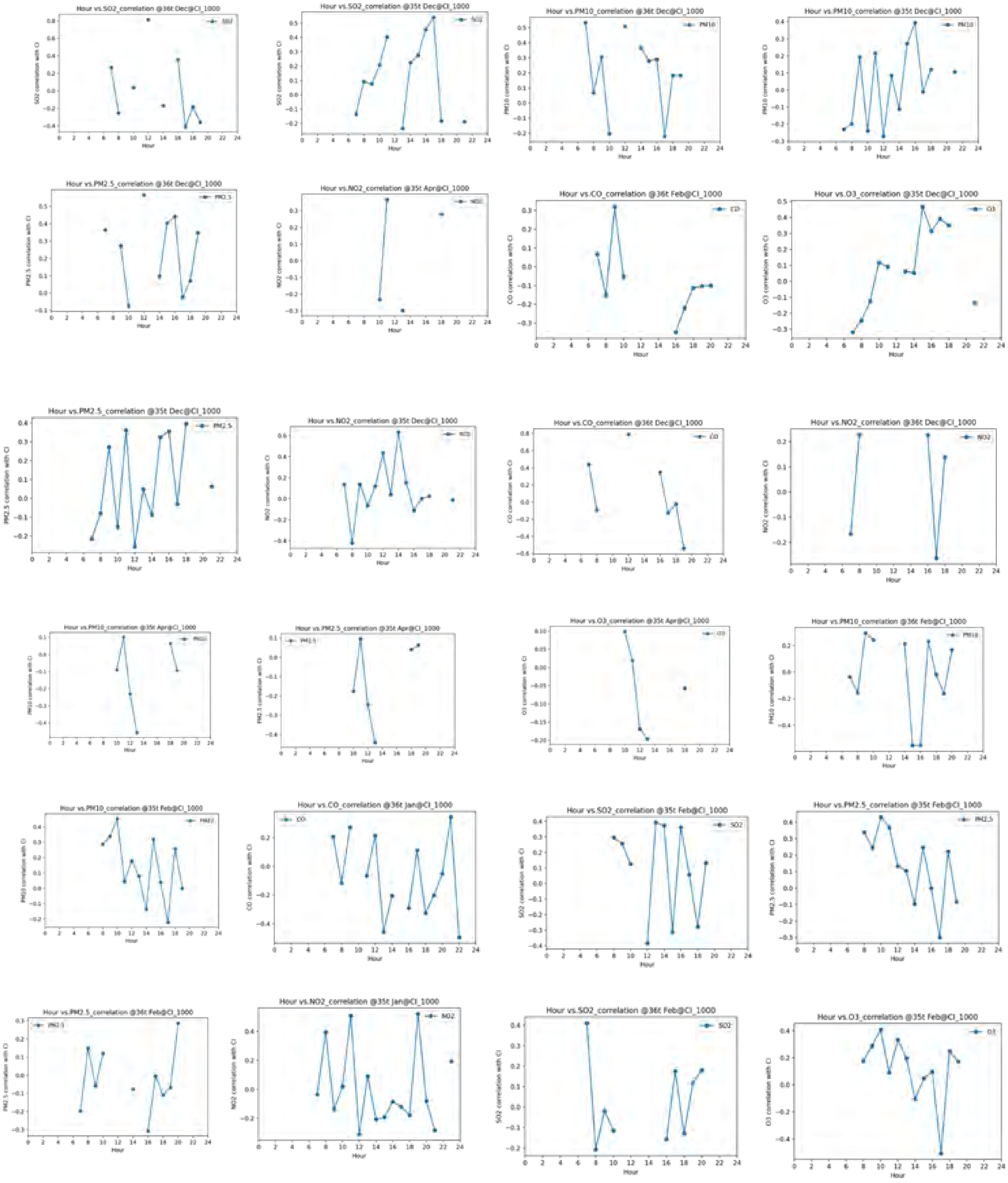
[8] E. J. Gumbel, “Bivariate exponential distributions,” Journal of the American Statistical Association, vol. 55, no. 292, pp. 698–707, 1960.

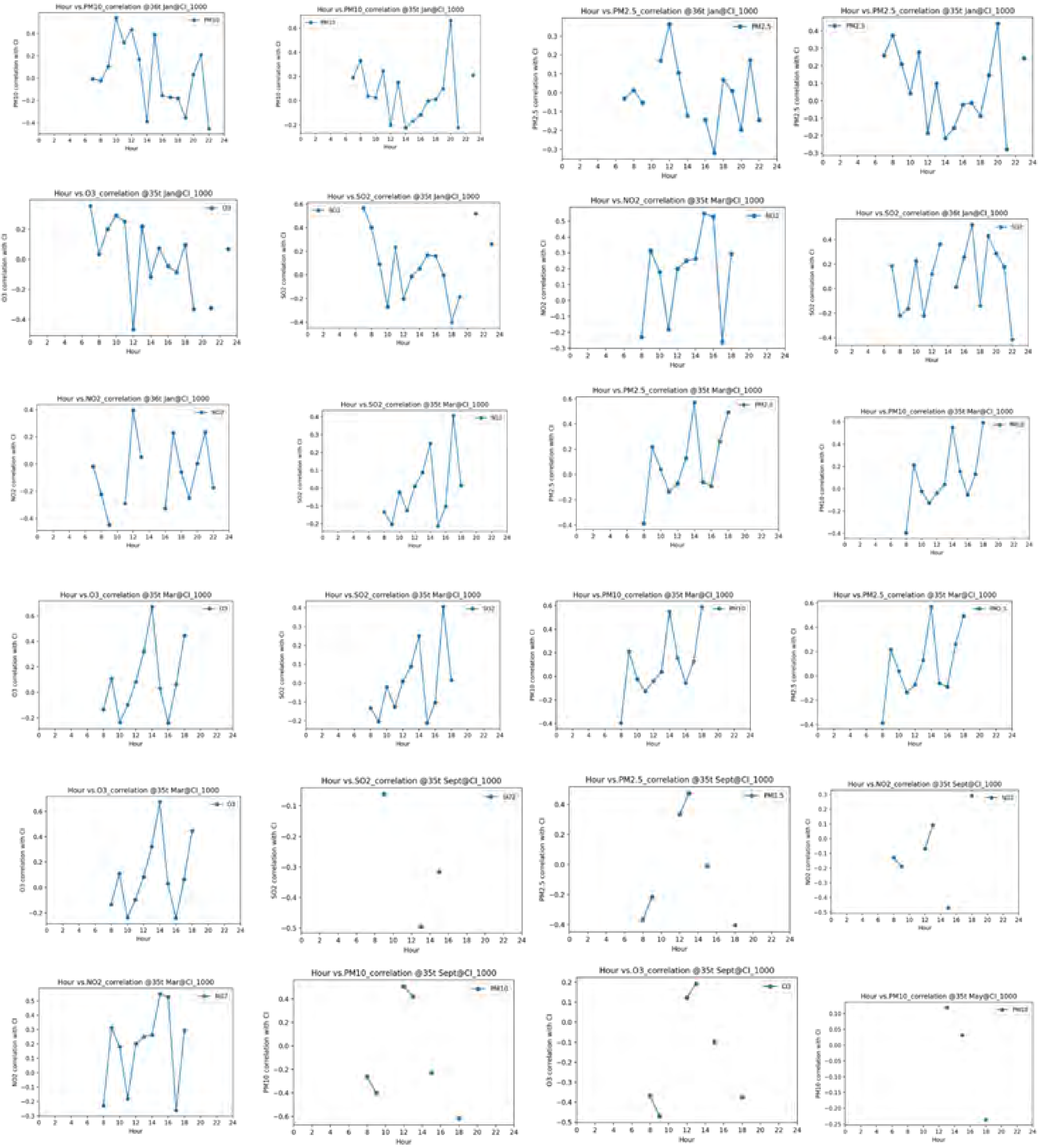
[9] H. Joe, “Parametric families of multivariate distributions with given margins,” Journal of multivariate analysis, vol. 46, no. 2, pp. 262–282, 1993.

[10] H. Joe, Multivariate models and multivariate dependence concepts. CRC Press, 1997.

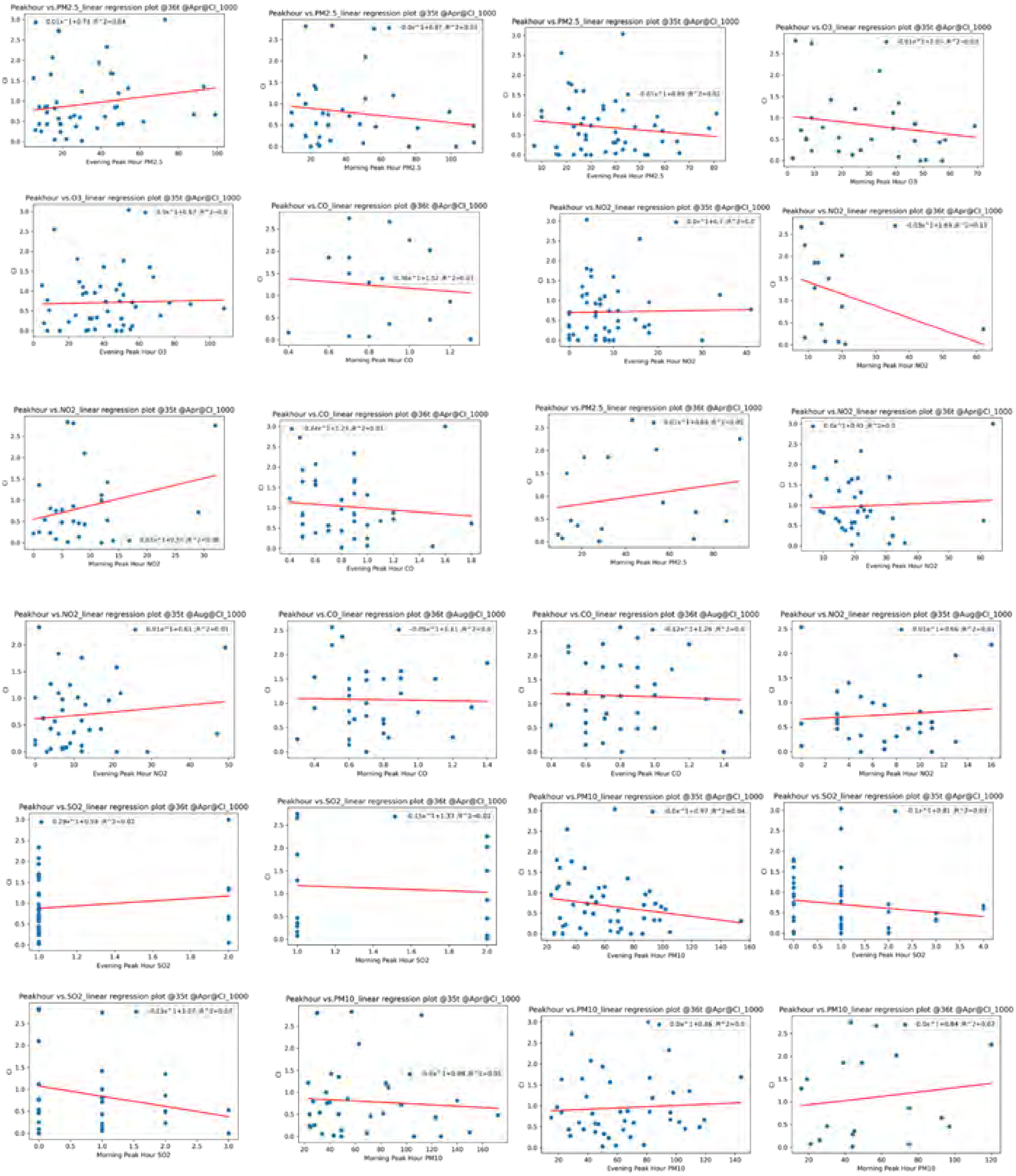
Appendix

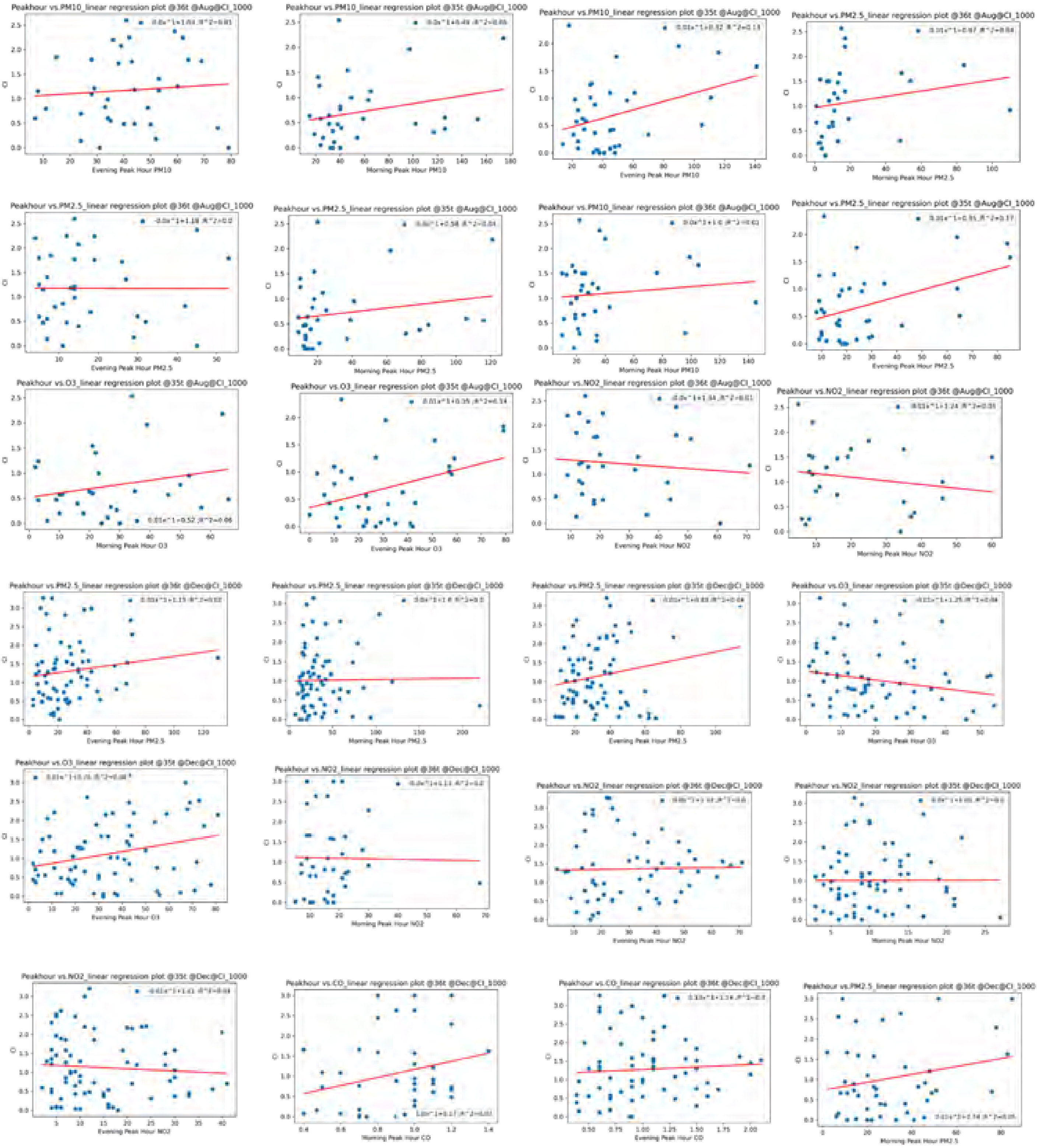
Correlation analysis results

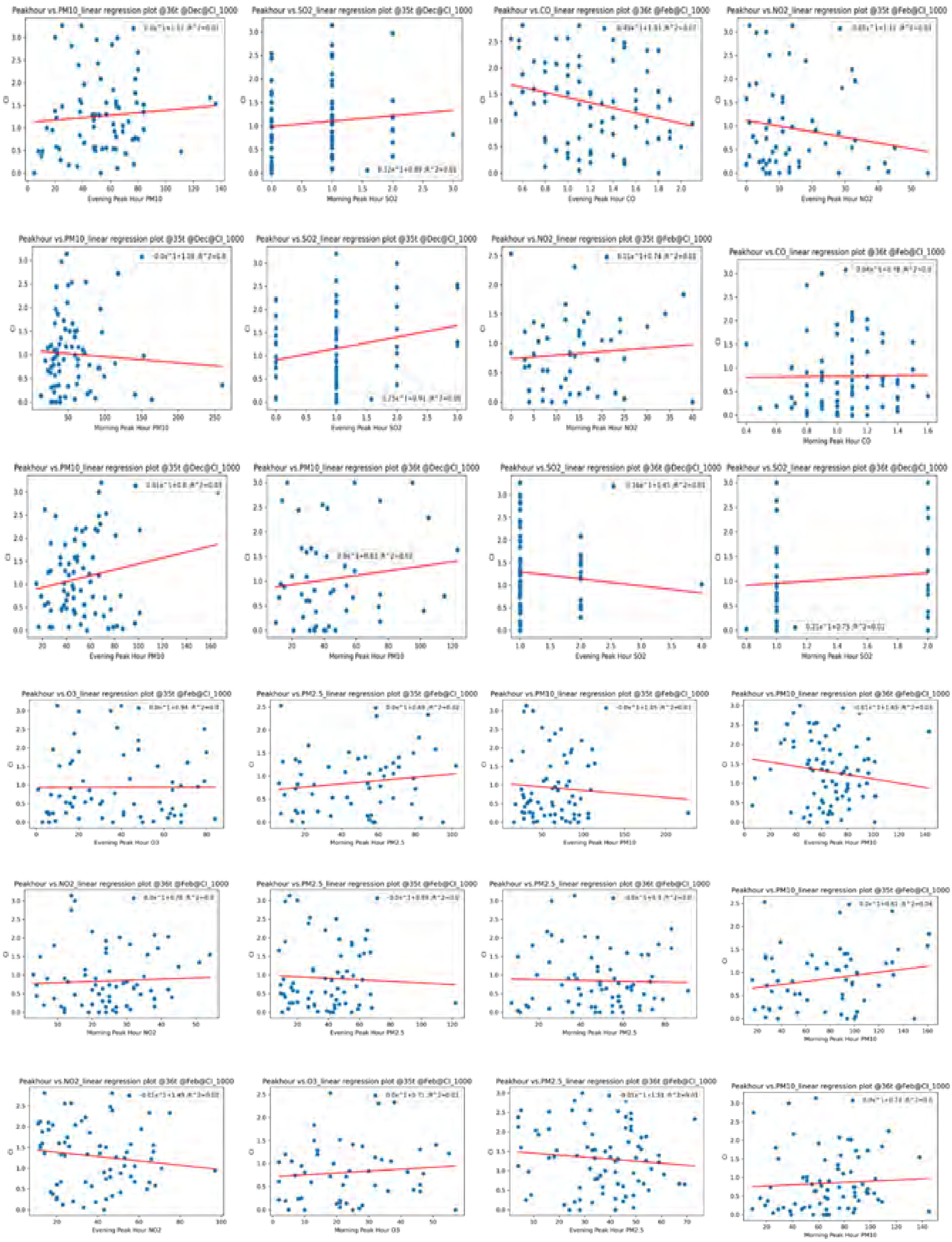


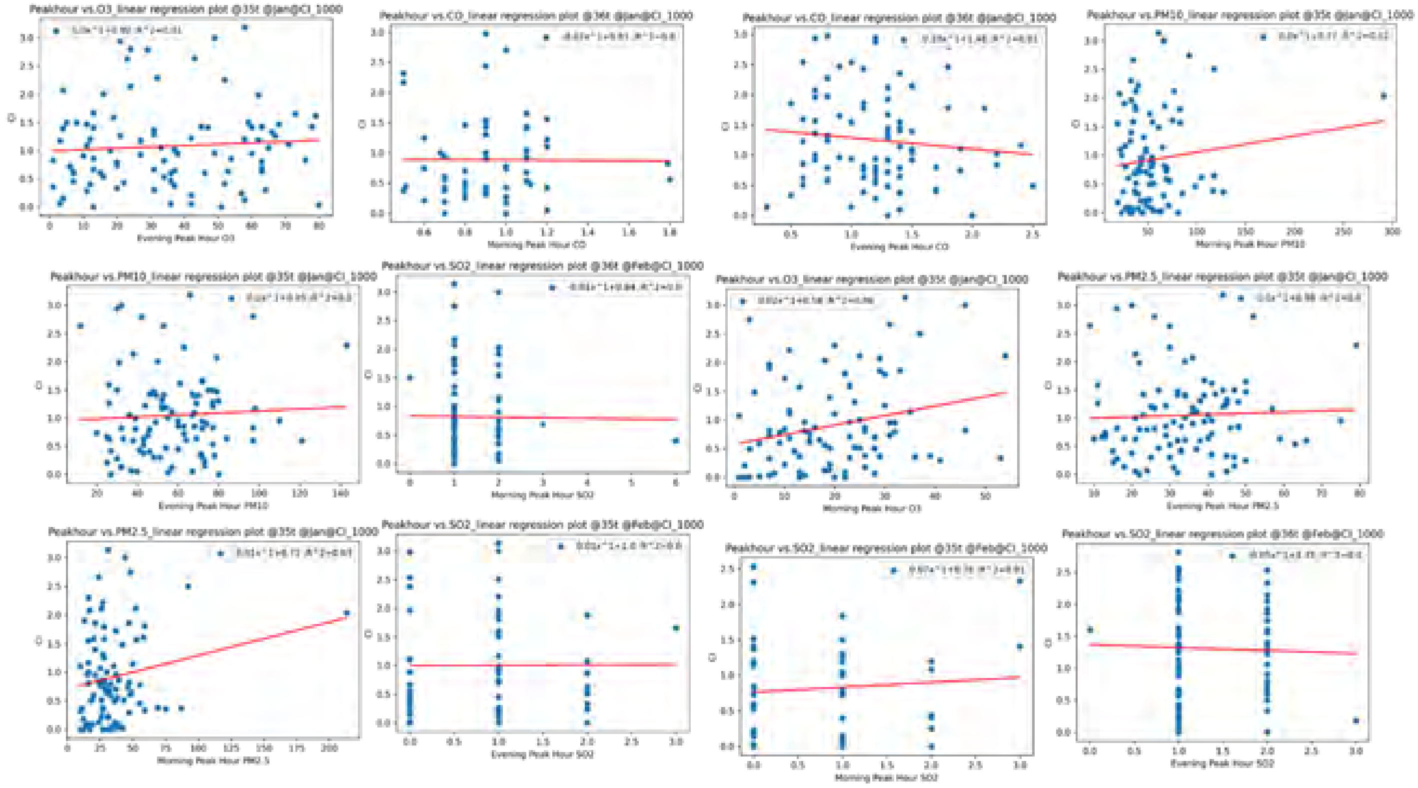

Hour vs.PM10_correlation @36t Jan@CI_1000
Hour vs.PM10_correlation @35t Jan@CI_1000
Hour vs.PM2.5_correlation @36t Jan@CI_1000
Hour vs.PM2.5_correlation @35t Jan@CI_1000
Hour vs.O3_correlation @35t Jan@CI_1000
Hour vs.SO2_correlation @35t Jan@CI_1000
Hour vs.NO2_correlation @35t Mar@CI_1000
Hour vs.SO2_correlation @36t Jan@CI_1000
Hour vs.NO2_correlation @36t Jan@CI_1000
Hour vs.SO2_correlation @35t Mar@CI_1000
Hour vs.PM2.5_correlation @35t Mar@CI_1000
Hour vs.PM10_correlation @35t Mar@CI_1000
Hour vs.O3_correlation @35t Mar@CI_1000
Hour vs.SO2_correlation @35t Mar@CI_1000
Hour vs.PM10_correlation @35t Mar@CI_1000
Hour vs.PM2.5_correlation @35t Mar@CI_1000
Hour vs.O3_correlation @35t Mar@CI_1000
Hour vs.SO2_correlation @35t Sept@CI_1000
Hour vs.PM2.5_correlation @35t Sept@CI_1000
Hour vs.NO2_correlation @35t Sept@CI_1000
Hour vs.NO2_correlation @35t Mar@CI_1000
Hour vs.PM10_correlation @35t Sept@CI_1000
Hour vs.O3_correlation @35t Sept@CI_1000
Hour vs.PM10_correlation @35t May@CI_1000
Hour

Linear analysis results

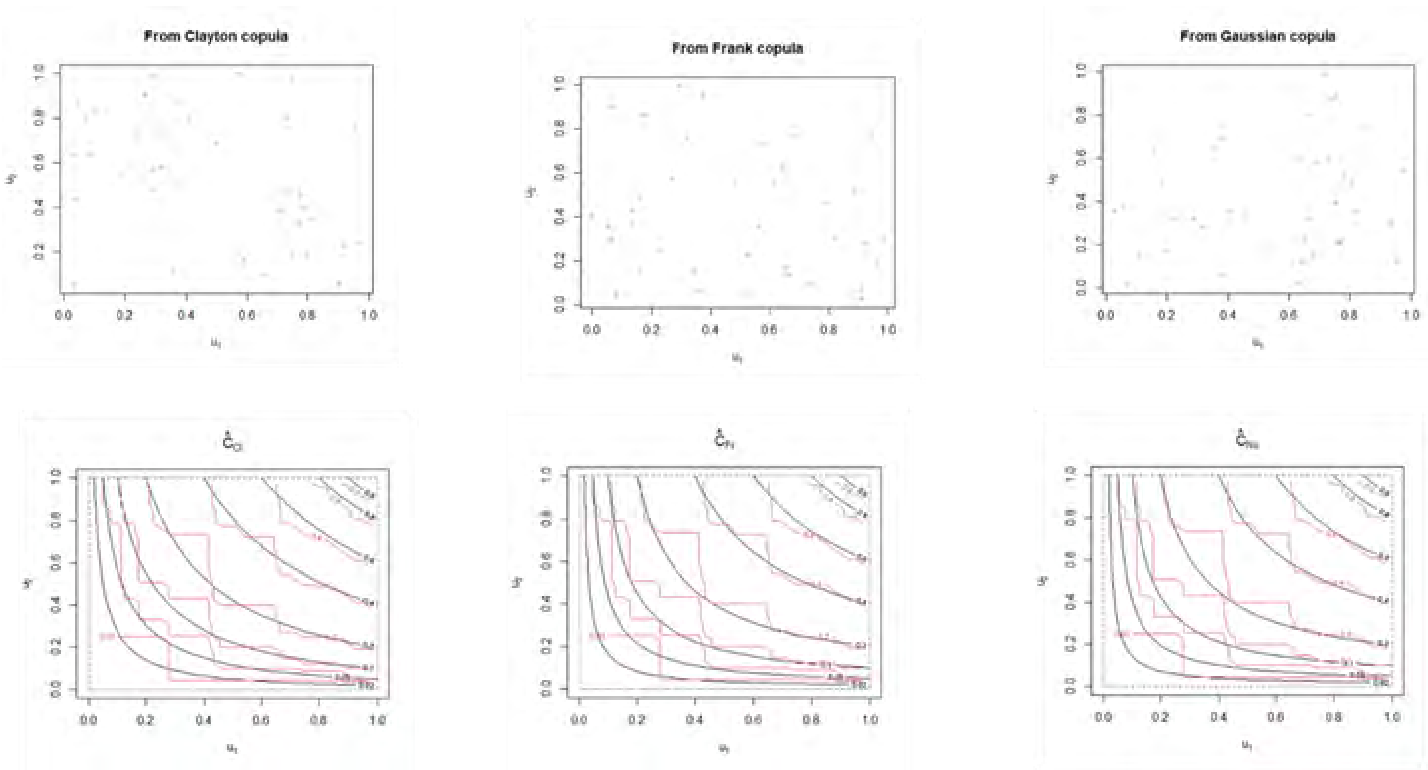



Peakhour vs.PM10_linear regression plot @36t @Aug@CI_1000
Evening Peak Hour PM10
Peakhour vs.PM10_linear regression plot @35t @Aug@CI_1000
Morning Peak Hour PM10
Peakhour vs.PM10_linear regression plot @35t @Aug@CI_1000
Evening Peak Hour PM10
Peakhour vs.PM2.5_linear regression plot @36t @Aug@CI_1000
Morning Peak Hour PM2.5
Peakhour vs.PM2.5_linear regression plot @36t @Aug@CI_1000
Evening Peak Hour PM2.5
Peakhour vs.PM2.5_linear regression plot @35t @Aug@CI_1000
Morning Peak Hour PM2.5
Peakhour vs.PM10_linear regression plot @36t @Aug@CI_1000
Morning Peak Hour PM10
Peakhour vs.PM2.5_linear regression plot @35t @Aug@CI_1000
Evening Peak Hour PM2.5
Peakhour vs.O3_linear regression plot @35t @Aug@CI_1000
Morning Peak Hour O3
Peakhour vs.O3_linear regression plot @35t @Aug@CI_1000
Evening Peak Hour O3
Peakhour vs.NO2_linear regression plot @36t @Aug@CI_1000
Evening Peak Hour NO2
Peakhour vs.NO2_linear regression plot @36t @Aug@CI_1000
Morning Peak Hour NO2
Peakhour vs.PM2.5_linear regression plot @36t @Dec@CI_1000
Evening Peak Hour PM2.5
Peakhour vs.PM2.5_linear regression plot @35t @Dec@CI_1000
Morning Peak Hour PM2.5
Peakhour vs.PM2.5_linear regression plot @35t @Dec@CI_1000
Evening Peak Hour PM2.5
Peakhour vs.O3_linear regression plot @35t @Dec@CI_1000
Morning Peak Hour O3
Peakhour vs.O3_linear regression plot @35t @Dec@CI_1000
Evening Peak Hour O3
Peakhour vs.NO2_linear regression plot @36t @Dec@CI_1000
Morning Peak Hour NO2
Peakhour vs.NO2_linear regression plot @36t @Dec@CI_1000
Evening Peak Hour NO2
Peakhour vs.NO2_linear regression plot @35t @Dec@CI_1000
Morning Peak Hour NO2
Peakhour vs.NO2_linear regression plot @35t @Dec@CI_1000
Evening Peak Hour NO2
Peakhour vs.CO_linear regression plot @36t @Dec@CI_1000
Morning Peak Hour CO
Peakhour vs.CO_linear regression plot @36t @Dec@CI_1000
Evening Peak Hour CO
Peakhour vs.PM2.5_linear regression plot @36t @Dec@CI_1000
Morning Peak Hour PM2.5


Peakhour vs.PM10_linear regression plot @36t @Dec@CI_1000
Evening Peak Hour PM10
Peakhour vs.SO2_linear regression plot @35t @Dec@CI_1000
Morning Peak Hour SO2
Peakhour vs.CO_linear regression plot @36t @Feb@CI_1000
Evening Peak Hour CO
Peakhour vs.NO2_linear regression plot @35t @Feb@CI_1000
Evening Peak Hour NO2
Peakhour vs.PM10_linear regression plot @35t @Dec@CI_1000
Morning Peak Hour PM10
Peakhour vs.SO2_linear regression plot @35t @Dec@CI_1000
Evening Peak Hour SO2
Peakhour vs.NO2_linear regression plot @35t @Feb@CI_1000
Morning Peak Hour NO2
Peakhour vs.CO_linear regression plot @36t @Feb@CI_1000
Morning Peak Hour CO
Peakhour vs.PM10_linear regression plot @35t @Dec@CI_1000
Evening Peak Hour PM10
Peakhour vs.PM10_linear regression plot @36t @Dec@CI_1000
Morning Peak Hour PM10
Peakhour vs.SO2_linear regression plot @36t @Dec@CI_1000
Evening Peak Hour SO2
Peakhour vs.SO2_linear regression plot @36t @Dec@CI_1000
Morning Peak Hour SO2
Peakhour vs.O3_linear regression plot @35t @Feb@CI_1000
Evening Peak Hour O3
Peakhour vs.PM2.5_linear regression plot @35t @Feb@CI_1000
Morning Peak Hour PM2.5
Peakhour vs.PM10_linear regression plot @35t @Feb@CI_1000
Evening Peak Hour PM10
Peakhour vs.PM10_linear regression plot @36t @Feb@CI_1000
Evening Peak Hour PM10
Peakhour vs.NO2_linear regression plot @36t @Feb@CI_1000
Morning Peak Hour NO2
Peakhour vs.PM2.5_linear regression plot @35t @Feb@CI_1000
Evening Peak Hour PM2.5
Peakhour vs.PM2.5_linear regression plot @36t @Feb@CI_1000
Morning Peak Hour PM2.5
Peakhour vs.PM10_linear regression plot @35t @Feb@CI_1000
Morning Peak Hour PM10
Peakhour vs.NO2_linear regression plot @36t @Feb@CI_1000
Evening Peak Hour NO2
Peakhour vs.O3_linear regression plot @35t @Feb@CI_1000
Morning Peak Hour O3
Peakhour vs.PM2.5_linear regression plot @36t @Feb@CI_1000
Evening Peak Hour PM2.5
Peakhour vs.PM10_linear regression plot @36t @Feb@CI_1000
Morning Peak Hour PM10



Copula function results

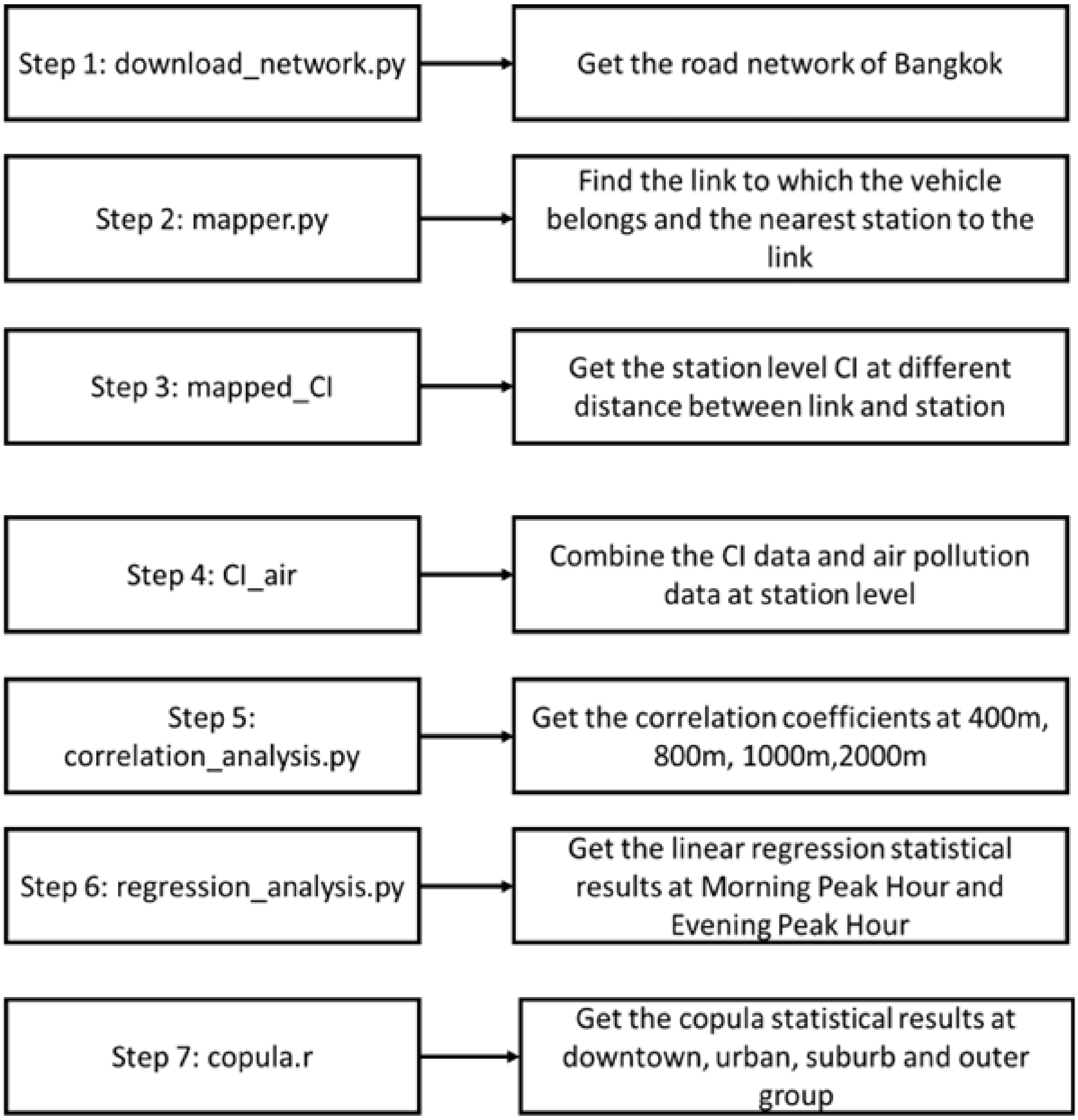


Air quality monitoring stations

Table 3 The locations of air quality monitoring stations

Station ID	Latitude	Longitude
35t	18.84063	98.96966
36t	18.79092	98.98811

Scripts



download_network.py

Expected Arguments

Argument	Type	Description
-north_lat	float	North latitude
-south_lat	float	South latitude
-east_lon	float	East longitude
-west_lon	float	West longitude
-save_path	str	Path to a save folder

Output
Directory output

Name	Date modified	Type	Size
Temp	21/06/2021 10:29 AM	File folder	
edges.cpg	21/06/2021 10:29 AM	CPG File	1 KB
edges	21/06/2021 10:29 AM	Microsoft Excel C...	83 KB
edges.dbf	21/06/2021 10:29 AM	DBF File	102 KB
edges.prj	21/06/2021 10:29 AM	PRJ File	1 KB
edges.shp	21/06/2021 10:29 AM	SHP File	76 KB
edges.shx	21/06/2021 10:29 AM	SHX File	2 KB
nodes.cpg	21/06/2021 10:29 AM	CPG File	1 KB
nodes.dbf	21/06/2021 10:29 AM	DBF File	21 KB
nodes.prj	21/06/2021 10:29 AM	PRJ File	1 KB
nodes.shp	21/06/2021 10:29 AM	SHP File	3 KB
nodes.shx	21/06/2021 10:29 AM	SHX File	1 KB

Example Usage
./download_network.py -north_lat 14.17
-west_lon 99.96
-south_lat 13.46
-east_lon 101.033
-save_path network_files/

mapper.py

Expected Arguments

Argument	Type	Description
-nodes_shapefile	str	Path to the nodes shapefile
-edges_shapefile	str	Path to the edges shapefile Note: This shapefile should contain station data. If it doesn't, see optional arguments
-data_file OR -data_directory	str	Path to data file or directory depending on argument
-save_path	str	Path to a save folder

Optional arguments

Argument	Type	Description
-station_list	str	Path to the station list csv
-edge_with_station_save_path	str	Save path for the generated edge shape file which contains station information

Output
Directory output

20170315_mapped	11/02/2021 10:08 AM	Microsoft Excel C...	535,747 KB
20170319_mapped	11/02/2021 9:52 AM	Microsoft Excel C...	525,964 KB
20170328_mapped	11/02/2021 12:25 PM	Microsoft Excel C...	533,171 KB

File output
Optional argument output
Directory output

edges_w_stations.cpg	21/06/2021 11:06 AM	CPG File	1 KB
edges_w_stations.dbf	21/06/2021 11:06 AM	DBF File	143 KB
edges_w_stations.prj	21/06/2021 11:06 AM	PRJ File	1 KB
edges_w_stations.shp	21/06/2021 11:06 AM	SHP File	76 KB
edges_w_stations.shx	21/06/2021 11:06 AM	SHX File	2 KB

Example Usage
./mapper.py -nodes_shapefile ./network_files/nodes.shp
-edges_shapefile ./network_files/edges_w_stations.shp
-data_directory ./data/

Generating edge shapefiles with station information
./mapper.py -nodes_shapefile ./network_files/nodes.shp
-edges_shapefile ./network_files/edges.shp
-station_list ./station_location.csv
-edge_with_station_save_path ./network_files/edges_w_stations.
shp

mapped_CI.py

Expected Arguments

Argument	Type	Description
-data_directory	str	Path to data directory

Output

CI_20170101_mapped.csv
CI_20170102_mapped.csv
CI_20170103_mapped.csv
CI_20170105_mapped.csv
CI_20170106_mapped.csv
CI_20170107_mapped.csv
CI_20170108_mapped.csv
CI_20170109_mapped.csv
CI_20170110_mapped.csv
CI_20170111_mapped.csv
CI_20170112_mapped.csv
CI_20170113_mapped.csv
CI_20170114_mapped.csv
CI_20170115_mapped.csv
CI_20170116_mapped.csv
CI_20170117_mapped.csv
CI_20170118_mapped.csv
CI_20170119_mapped.csv
CI_20170120_mapped.csv

CI_air.py

Expected Arguments

Argument	Type	Description
-data_directory	str	Path to data directory

Output

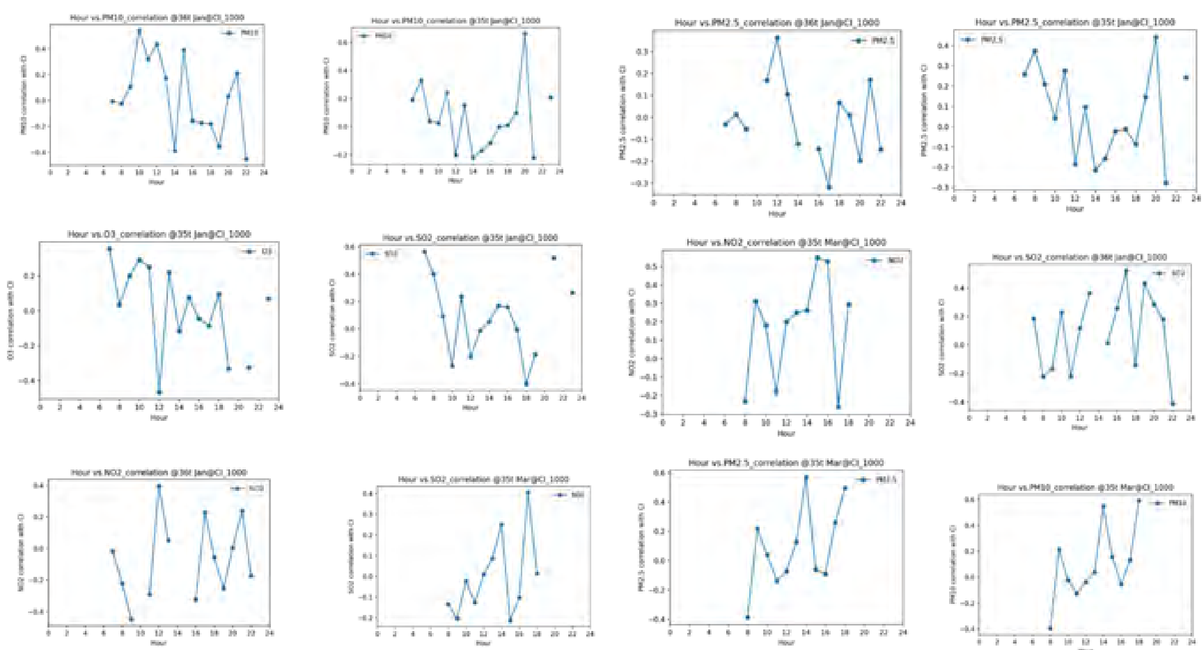
year	Month	day	hour	Station	CI_inf	CI_400	CI_500	CI_600	CI_700	CI_800	CI_900	CI_1000	CI_1200	CI_1500	CI_1800	CI_2000	Date	PM2.5	PM10	O3	CO	NO2	SO2
2017	1	1	0	36t	0.486625										3	3	1/01/2017 0:00		38		1.1	11	1
2017	1	1	5	36t	2.066196			0	0	0	0	0	0	0	1.5	1.5	1/01/2017 5:00		35		1	15	1
2017	1	1	6	35t	0.289084				0.152174	0.152174	0.152174	0.152174	0.160273	0.160273	0.120205	0.120205	1/01/2017 6:00	29	34	11		0	1
2017	1	1	6	36t	1.949526				5.666667	5.666667	5.666667	5.666667	3.125259	3.125259	2.244	2.244	1/01/2017 6:00		44		1	12	1
2017	1	1	7	36t	0.627064			0.212121	0.295716	0.295716	0.295716	0.259067	0.236093	0.188875	0.296284	0.296284	1/01/2017 7:00		33		0.8	8	1
2017	1	1	8	36t	1.347144				2.973974	2.973974	2.973974	2.973974	1.288556	1.288556	1.159612	1.159612	1/01/2017 8:00		13		0.9	12	1
2017	1	1	9	35t	0.197017								0.213235	0.213235	0.213235	0.142157	1/01/2017 9:00	25	25	18		0	0
2017	1	1	9	36t	3.02442												1/01/2017 9:00		25		1	15	1
2017	1	1	10	36t	1.360815		0.37931	0.928369	1.225002	1.333367	1.433364	1.433364	1.22091	1.169251	1.133673	1.133673	1/01/2017 10:00		30		1.1	16	1
2017	1	1	11	35t	0.959785			0.25	0.25	0.25	0.25	0.125	0.127493	0.09562	0.076496	0.076496	1/01/2017 11:00	36	40	21		0	0
2017	1	1	11	36t	3.583847				5.666667	5.666667	5.666667	5.666667	2.743096	2.743096	2.394477	2.131761	1/01/2017 11:00		18		1	13	1
2017	1	1	12	35t	1.168214		4	4	4	4	4	2.008475	2.008475	2.008475	1.431182	1.431182	1/01/2017 12:00	35	40	23		0	1
2017	1	1	12	36t	1.012248									0.269231	0.179487	0.179487	1/01/2017 12:00		22		1	10	1
2017	1	1	13	35t	2.12912												1/01/2017 13:00	36	40	29		0	1
2017	1	1	13	36t	0.838371			0.651786	0.651786	0.651786	0.651786	0.461551	0.304849	0.526606	0.478637	0.47596	1/01/2017 13:00		29		0.8	5	1
2017	1	1	14	35t	1.608806												1/01/2017 14:00	35	35	30		0	0
2017	1	1	14	36t	0.546361			1.352941	1.352941	1.352941	1.352941	1.352941	1.428673	1.148067	0.827566	0.827566	1/01/2017 14:00		26		0.8	5	1
2017	1	1	15	35t	0.301745				0.132075	0.132075	0.132075	0.132075	0.191038	0.191038	0.334566	0.334566	1/01/2017 15:00	34	35	28		0	0
2017	1	1	15	36t	1.336056				0.428571	0.428571	0.428571	0.690476	1.473893	1.577623	1.577623	1.577623	1/01/2017 15:00		30		0.8	5	1
2017	1	1	16	35t	0.887292												1/01/2017 16:00	29	31				
2017	1	1	16	36t	1.0929								0	0	0	0	1/01/2017 16:00		27		0.8	6	1
2017	1	1	17	35t	1.162162												1/01/2017 17:00	30	31	13		0	0
2017	1	1	17	36t	0.184211												1/01/2017 17:00		27		0.9	7	1
2017	1	1	18	35t	0												1/01/2017 18:00	30	34	10		0	1
2017	1	1	18	36t	0.835569												1/01/2017 18:00		23		0.9	8	1

correlation_analysis.py

Expected Arguments

Argument	Type	Description
-data_directory	str	Path to data directory
-CI_distance	str	The distance between links and station

Output

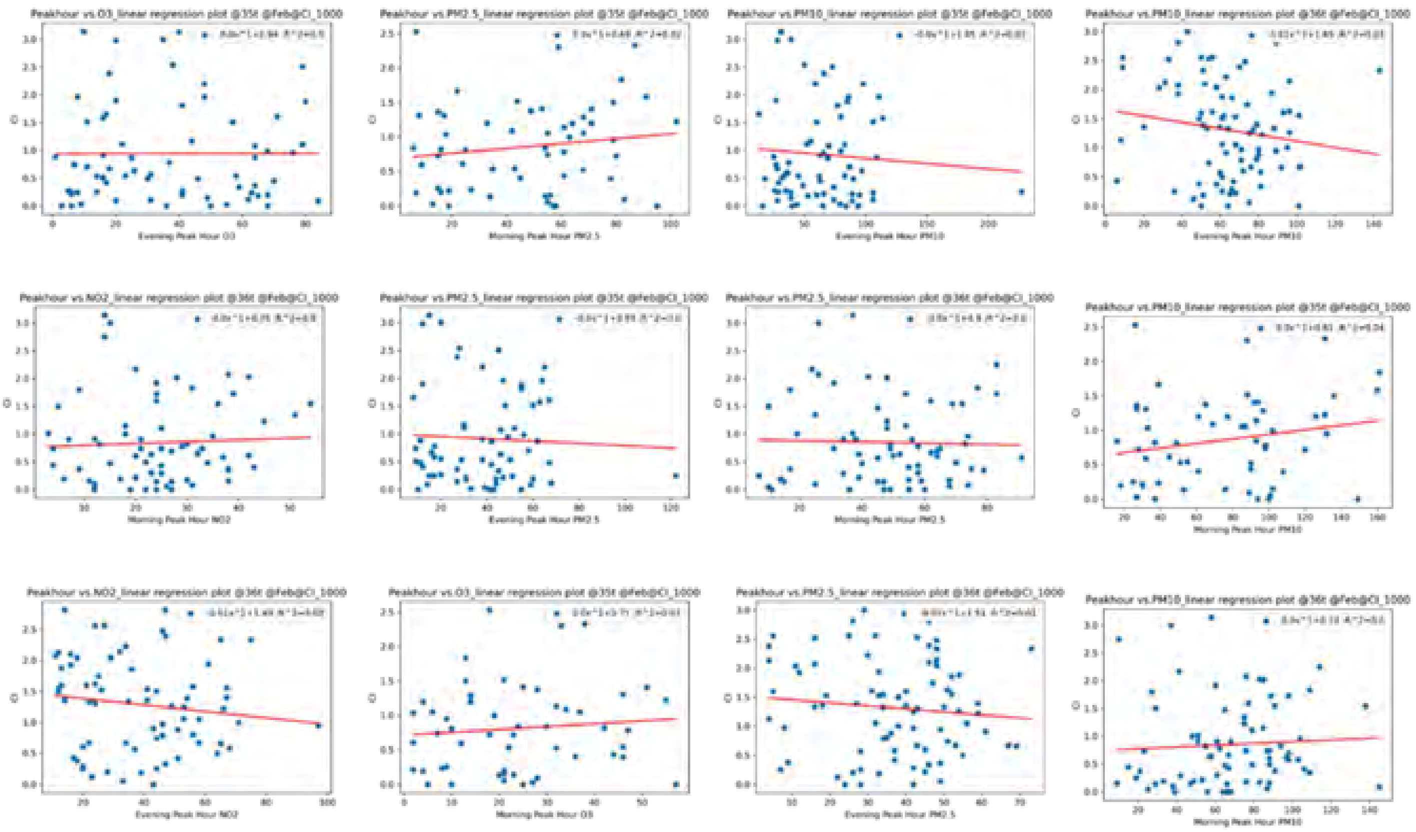


regression_analysis.py

Expected Arguments

Argument	Type	Description
-data_directory	str	Path to data directory
-CI_distance	str	The distance between links and station

Output



copula.r

Expected Arguments

Argument	Type	Description
-data_directory	str	Path to data directory

Output