

International Journal of Environmental Research and Public Health



Article Haze Influencing Factors: A Data Envelopment Analysis Approach

Yi Zhou¹, Lianshui Li², Ruiling Sun¹, Zaiwu Gong^{2,*}, Mingguo Bai³ and Guo Wei⁴

- ¹ School of Applied Meteorology, Nanjing University of Information Science and Technology, Nanjing 210044, China; 20171103093@nuist.edu.cn (Y.Z.); 20171103091@nuist.edu.cn (R.S.)
- ² School of Management Science and Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; Ilsh@nuist.edu.cn
- ³ School of Business, Anhui University of Technology, Maanshan 243032, China; baimingguo@163.com
- ⁴ Department of Mathematics and Computer Science, University of North Carolina at Pembroke, Pembroke, NC 28372, USA; guo.wei@uncp.edu
- * Correspondence: zwgong@nuist.edu.cn; Tel.: +86-25-58731142

Received: 4 December 2018; Accepted: 11 March 2019; Published: 14 March 2019



Abstract: This paper investigates the meteorological factors and human activities that influence PM_{2.5} pollution by employing the data envelopment analysis (DEA) approach to a chance constrained stochastic optimization problem. This approach has the two advantages of admitting random input and output, and allowing the evaluation unit to exceed the front edge under the given probability constraint. Furthermore, by utilizing the meteorological observation data incorporated with the economic and social data for Jiangsu Province, the chance constrained stochastic DEA model was solved to explore the relationship between the meteorological elements and human activities and PM_{2.5} pollution. The results are summarized by the following: (1) Among all five primary indexes, social progress, energy use and transportation are the most significant for PM_{2.5} pollution. (2) Among our selected 14 secondary indexes, coal consumption, population density and civil car ownership account for a major portion of PM_{2.5} pollution. (3) Human activities are the main factor producing PM_{2.5} pollution. While some meteorological elements generate PM_{2.5} pollution, some act as influencing factors on the migration of PM_{2.5} pollution. These findings can provide a reference for the government to formulate appropriate policies to reduce PM_{2.5} emissions and for the communities to develop effective strategies to eliminate PM_{2.5} pollution.

Keywords: PM_{2.5}; human activities; meteorological factors; chance constrained stochastic DEA

1. Introduction

Recognizing the key influencing factors of haze is critically important for haze control. The purpose of this study is to explore contributing factors related to human activities or meteorological events.

Haze not only pollutes the ecological environment, but also seriously affects the physical and mental health of humans. Low visibility in haze weather could easily lead to traffic accidents, endangering human life and property [1]. The particulate pollutants in haze can cause great harm to humans and the toxic effect is mainly ascribed to $PM_{2.5}$, defined as fine particulate matter with a diameter smaller than 2.5 µm [2]. Due to their size, $PM_{2.5}$ particles can penetrate into the respiratory tract and reach the alveoli [3], and can even enter the systemic circulation and affect other organs [4]. $PM_{2.5}$ may also directly affect the cardiovascular system by entering into the systemic circulation and causing cardiovascular dysfunction [5]. They can also cause depression and other mental illnesses [6]. As $PM_{2.5}$ is the main component of haze and causes great harm to humans, the key to control haze

pollution is to decrease PM_{2.5} pollution. Therefore, this paper focuses on investigating the influencing factors of PM_{2.5} pollution.

In recent years, the haze condition in China has become so significant that the problem has triggered a warning to the government, the communities, and all citizens. The annual average concentration of $PM_{2.5}$ from 2013 to 2016 is as high as 57.75 µg/m³ in China [7], which is over six times higher than that of the United States (8.47 µg/m³ [8]) or 5 times higher than that of the WHO standard (10 µg/m³).

1.1. Human Activities

Many scholars have conducted various studies on the causes and mechanisms of haze, hoping to provide scientific advice for the government to prevent and control haze [9–11]. It is known that human activities, such as industrial pollution and fuel combustion, automobile exhaust, construction dust, road dust and industrial dust, contribute substantially to haze pollution. Taking Beijing as an example, in 2017, vehicle exhaust, coal combustion, industrial production and dust emission were the main sources of PM_{2.5}, accounting for 31.1%, 22.4%, 18.1% and 14.3%, respectively [12].

Direct emissions mainly come from various types of combustion, including fossil fuel (coal, oil, natural gas) combustion and biomass burning (straw, firewood). Coal accounts for a large proportion of China's energy consumption structure. In 2017, coal consumption accounted for 60.4% of total energy consumption [13].

Haze pollution caused by coal combustion in China can be divided into two types by scale: one is large-scale high-density coal combustion such as power plants and iron and steel enterprises. You and Xu [14] along with Tang et al. [15] found that coal combustion is the main source of particulate pollution in China.

The other is small-scale coal combustion, such as household coal combustion, which is widely distributed and easily overlooked. Cheng et al. [16] pointed out that the actual coal consumption by residents is 4.6 times that of the amount reported in the statistical yearbook for that year, and hence household coal combustion should not be ignored. The studies by Xue et al. [17] and Zhang et al. [18] indicated that PM_{2.5} emissions from the residents of the Beijing-Tianjin-Hebei region accounted for nearly 50% of the total PM_{2.5} emissions and is exhibiting an upward trend.

Pollutants that can be turned to $PM_{2.5}$ pollution include nitrogen oxides, sulfur dioxide and volatile organic compounds. Vehicle exhaust is a main source of this kind of $PM_{2.5}$ pollution. Ming et al. [19], Gao et al. [20] and Yang et al. [21] pointed out that the increase in the number of cars has caused a significant increase in haze pollution. Gao et al. [22] assessed the health damage caused by vehicle exhaust. Goel and Kumar [23] studied the exposure to vehicle exhaust under different traffic conditions, and their findings show that $PM_{2.5}$ pollution caused by vehicle exhaust when traffic jams is more serious than when traffic flow is smooth.

Construction dust, road dust and industrial dust are also important sources of $PM_{2.5}$. Kinsey et al. [24] and Hassan et al. [25] found that the dissipation of dust during construction is one major source of PM emissions. Zhao et al. [26] pointed out that high-intensity economic activities and large-scale infrastructure construction are reasons for the increase in $PM_{2.5}$ concentration in China. Lin and Zhu [27] and Wu et al. [28] found that the concentration of $PM_{2.5}$ is closely related to the urbanization process in China. Zhou et al. [29] showed that there was a significant positive relation between the emission of industrial dust and the concentration of $PM_{2.5}$.

1.2. Meteorological Factors

Some scholars suggest that meteorological events are natural causes of haze formation. Studies found that meteorological factors related to haze include relative humidity, wind speed, and precipitation. The studies by Fu et al. [30], Chambers et al. [31], Yang et al. [32] and Li et al. [33] indicated that relative humidity and wind speed have a great relationship to haze pollution. Guo et al. [34] found that the washing effect of precipitation can significantly reduce haze pollution.

While meteorological factors are natural causes or objective factors for the formation of haze, human activities account for a significant part of the pollution, especially for $PM_{2.5}$ pollution. $PM_{2.5}$ pollution conditions are the combined effects of meteorological and human factors. However, in the past, only a few studies have considered the combined effects of meteorological conditions and human activities on haze formation.

1.3. Our Approach

Jiangsu is an important industrial province in China, with a GDP (Gross Domestic Product) of 8.59 trillion RMB in 2017, which is about 10.03% of China's total GDP or the national GDP of Russia (8.51 trillion RMB) [35]. With the fast development of industrialization and urbanization, people there are also facing heavy pollutant emissions, and haze and fog have become a serious concern for the residents and a burden on the public. Under China's new standard economic system, people urgently demand air quality improvement. As the key task of ecological environment governance, PM_{2.5} pollution prevention and treatment have become a critical task to win "the battle for blue skies". Differentiating the impact of different factors on PM_{2.5} pollution in Jiangsu Province can provide more precise information for the government and communities to design suitable policies and care strategies.

Meteorological factors and human reasons will be taken as the influencing factors of $PM_{2.5}$ pollution (Section 3). Through related literature analyses, human factors are divided into four major categories: (1) Social progress: including the impact of urbanization, housing construction and demolition, and population growth. (2) Transportation: vehicle exhaust emissions. (3) Energy use: fossil fuel combustion. (4) Environmental protection: human efforts to reduce pollution. Because the social and economic data are described by years, but the meteorological data are provided by units of hours or days, we face the problem of the time scale not being uniform when the relevant data is processed.

DEA is a linear programming methodology to evaluate the efficiency of multiple decision-making units (DMUs) when the production process presents a structure of multiple inputs and outputs [36]. This method is not required to determine the explicit expression of the relationship between input and output variables, which eliminates many subjective factors, and hence, has strong objectivity. Therefore, the DEA method is widely used in efficiency evaluation [37]. However, traditional DEA does not consider the random errors of input and output variables, such as measurement error and data noise [38,39], which are unavoidable in data collection [40].

In this paper, the stochastic data envelopment analysis model is applied to resolve this problem. When compared with traditional DEA, chance constrained stochastic DEA has two advantages: admitting random input and output, and allowing the evaluation unit to exceed the front edge under the given probability constraint [41,42]. We study the coupling effect of multiple factors in the effort to identify the influencing factors of PM_{2.5} pollution in Jiangsu province.

The remainder of the article is organized as follows: Section 2 presents the traditional DEA method and chance constrained stochastic data envelopment analysis method. Section 3 introduces (1) the sources of meteorological data, as well as economic and social data, for the 13 prefecture-level cities in Jiangsu Province, (2) the construction of primary and secondary level indicators affecting $PM_{2.5}$ pollution are provided, (3) three sets of experiments are designed to identify key factors of $PM_{2.5}$ pollution based upon different levels of indicators. Section 4 contains an empirical analysis. Through corresponding analyses and comparisons of the results from the three sets of experiments, the key factors affecting $PM_{2.5}$ pollution are analyzed, and the comprehensive effects of various factors for $PM_{2.5}$ pollution at different risk levels are evaluated. Section 5 is the summary and conclusion.

2. Model and Method

Using the chance constrained stochastic DEA method, different experiments were carried out using different levels of indexes. If the result changed substantially after a certain index was deleted,

the deleted index had a strong relation with $PM_{2.5}$. Therefore, the following three experiments were designed:

- I. All 14 secondary indexes were used as input indicators, GDP and PM_{2.5} as output indicators, and chance constrained stochastic DEA was used for the calculation.
- II. Based on the settings of experiment I, a secondary index was deleted and the remaining 13 input indicators and 2 output indicators were used for the calculation. The purpose of this experiment was to identify which secondary input index has the greatest impact on output indicators.
- III. As indicators may interact with each other, considering the interrelatedness and mutual influence between indicators, and treating the same type of indexes as one group, we obtained five primary indexes. Based on the settings of experiment I, a primary index was deleted and the remaining four input indicators and two output indicators were used for the calculation. The purpose of this experiment was to identify which primary index has the greatest impact on output indicators.

Data envelopment analysis (DEA) was proposed by the famous operational research scholars, Charnes, Copper and Rhodes, in 1978 [43], and is a non-parametric evaluation method that is relatively efficient when compared to the regression analysis method and least squares method. It is suitable for the evaluation of several decision units with the same type of multiple inputs and outputs. This method does not require determination of the explicit expression of the relationship between input and output variables, which eliminates many subjective factors, and hence, has increased objectivity. Therefore, the DEA method is widely applied in efficiency evaluation.

In this paper, we employ the BCC model: a model for evaluating technical efficiency, named after proposers Banker, Charnes and Cooper in 1984 [44], which is the following optimization problem:

Objective function:

$$\operatorname{Min}\left[\theta - \varepsilon \left(e_1^T S^- + e_2^T S^+\right)\right] \tag{1}$$

Constraints:

$$\begin{cases} \sum_{j=1}^{n} X_{j}\lambda_{j} + S^{-} = \theta X_{o} \\ \sum_{j=1}^{n} Y_{j}\lambda_{j} - S^{+} = Y_{o} \\ \sum_{j=1}^{n} \lambda_{j} = 1 \\ \lambda_{j} \ge 0, j = 1, 2, \cdots, n. \\ S^{+} \ge 0 \\ S^{-} \ge 0 \end{cases}$$

$$(2)$$

In (2), the input variable of a decision making unit (DMU) in an economic (production) activity is $X = (x_1, \dots, x_i, \dots, x_m)$, where x_i is the *i*-th input;

The output variable is $Y = (y_1, \dots, y_r, \dots, y_s)$, where y_r is the *r*-th output;

 X_j is the input vector of the j-th DMU, Y_j is the output vector of the j-th DMU;

 X_o and Y_o are the corresponding indices of the evaluation DMU;

 S^- is the slack variable of the input variable;

 S^+ is the slack variable of the output variable;

- e_1^T is the unit vector in the *m*-th dimension;
- e_2^T is the unit vector in the *s*-th dimension;
- ε is the Archimedes infinitesimal;

 λ_i is the nonnegative weight of the *j*-th DMU;

Thus, the entire production activity of the DMU can be represented by (X, Y), and the input set of n DMUs can form an $n \times m$ matrix, while the output set can form an $n \times s$ order output matrix. When the optimum output is 1, i.e., $\theta = 1$, the *j*-th DMU is technically effective. When $\theta < 1$, the *j*-th DMU is not technically effective.

When dealing with practical problems, input and output variables are often disturbed by random factors such as measurement error and data noise. Because of the random nature of the phenomena and the laws themselves, the input and output variables of DMU may not be accurately determined, and these variables are subject to certain random distribution. There will be great errors when using the traditional DEA model to deal with such problems. In addition, the traditional DEA model has some difficulty in dealing with data with different time scales. Therefore, the chance constrained stochastic DEA method becomes important in practice and can be used to calculate data with different time scales [45].

The chance constrained stochastic DEA model based on the BCC output orientation is the following [46]:

Objective function:

$$\operatorname{Max} \theta_{0}$$
 (3)

Constraints:

$$\begin{aligned}
& \Pr\left(\sum_{j=1}^{n} \lambda_{j} \widetilde{x}_{ij} \leq \widetilde{x}_{io}\right) \geq 1 - \alpha, i = 1, 2, \dots, m \\
& \Pr\left(\sum_{j=1}^{n} \lambda_{j} \widetilde{y}_{rj} \geq \theta_{o} \widetilde{y}_{ro}\right) \geq 1 - \alpha, r = 1, 2, \dots, s \\
& \sum_{j=1}^{n} \lambda_{j} = 1 \\
& \lambda_{j} \geq 0, j = 1, 2, \dots, n
\end{aligned} \tag{4}$$

Pr(.) represents the probability;

 $\alpha \in [0, 1]$ is the risk level;

 λ_i is the parameter of DMU_i ;

 \tilde{x}_{ij} is the input variables of DMU_j ;

 \tilde{y}_{rj} is the output variable of DMU_j ;

In our study, n = 13, representing 13 prefecture-level cities in Jiangsu province. We have *m* different input variables \tilde{x}_{ij} , with three different schemes:

- (1) m = 14, representing the 14 secondary indexes used in Experiment I;
- (2) m = 13, representing the 13 secondary indexes used in Experiment II;

(3) m = 5, representing the 5 primary indexes used in Experiment III.

For output variables \tilde{y}_{ri} , s = 2, representing the 2 output variables.

As Model (2) is an output-oriented model, the reciprocal of the optimal solution of the target value obtained from Model (4) is the stochastic chance constraint efficiency of the evaluated DMU. Model (4) can be rewritten as (without the slack variable):

Objective function:

$$\operatorname{Max} \theta_o \tag{5}$$

Constraints:

$$\begin{cases} \sum_{j=1}^{n} \lambda_{j} \overline{x}_{ij} - \Phi^{-1}(\alpha) \sqrt{\sum_{j \neq o} \lambda_{j}^{2} \left(\sigma_{ij}^{I}\right)^{2} + (\lambda_{o} - 1)^{2} \left(\sigma_{io}^{I}\right)^{2}} \leq \overline{x}_{io}, i = 1, 2, \dots, m \\ \sum_{j=1}^{n} \lambda_{j} \overline{y}_{rj} + \Phi^{-1}(\alpha) \sqrt{\sum_{j \neq o} \lambda_{j}^{2} \left(\sigma_{rj}^{o}\right)^{2} + (\lambda_{o} - \theta_{o})^{2} \left(\sigma_{ro}^{o}\right)^{2}} \geq \theta_{o} \overline{y}_{ro}, r = 1, 2, \dots, s \\ \sum_{j=1}^{n} \lambda_{j} = 1 \\ \lambda_{j} \geq 0, j = 1, 2, \dots, n \end{cases}$$

$$(6)$$

 $\Phi^{-1}(\alpha)$ is the value of the inverse distribution function of the standard normal distribution function at α ;

 $\sigma_i^I(\lambda)$ is the standard deviation of the linear combination distribution of all i-th input variables; σ_{ij}^I and σ_{rj}^o are the standard deviations of \tilde{x}_{ij} and \tilde{y}_{rj} , respectively;

 \overline{x}_{ij} and \overline{y}_{rj} are corresponding mean values of the DMU_j input and output.

For the unit to be evaluated DMUo, θ_0 in the objective function is the target optimal solution, $o \in \{1, 2, ..., n\}$. Let its stochastic efficiency be φ_0 , then it holds that $\varphi_0 = \frac{1}{4}$.

3. Data Sources and Indexes Explanation

In this section, we describe the data collection and data characteristics, as well as indexes, we have constructed.

3.1. Data Sources

The meteorological data comes from the China Meteorological Science Data Sharing Service Network (CMSDSSN) ranging from 2013 to 2016 [47]. The CMSDSSN is a meteorological data sharing network system, and a service platform for all domestic and global users.

Social and economic data are derived from the Jiangsu Statistical Yearbook ranging from 2014 to 2017 [48], and statistical yearbook of the 13 prefectural cities (Figure 1) in Jiangsu Province (Xuzhou, Suqian, Lianyungang, Huaian, Yancheng, Yangzhou, Taizhou, Nantong, Zhenjiang, Changzhou, Wuxi, Suzhou, and Nanjing).



Figure 1. Administrative division map of Jiangsu province.

Air pollution data comes from the daily monitoring data of atmospheric pollutants from the Environmental Protection Department of Jiangsu ranging from 2013 to 2016. The Jiangsu Environmental Protection Department is responsible for establishing and improving the basic system of environmental protection, monitoring the emission of pollutants.

3.2. Indexes Explanation

In this section, five indexes; meteorological conditions, social progress, transportation, energy use and environmental protection, were constructed as primary indexes. Each primary index contained one or more secondary indexes. The setting of the indexes is shown in Table 1. The following are further explanations for the five primary indexes:

Weather conditions refers to different meteorological elements, including six secondary indexes: wind speed, precipitation, temperature, atmospheric pressure, sunshine hours and relative humidity.

Social progress includes the progress and development of material civilization. It is mainly reflected in the progress of material production methods and economic life. It is the achievement of humans in reforming and conquering nature, including improvement of the living standard, improvement of food and clothing, change of life style and many other things. Based on the existing data, three secondary indexes were selected: urbanization rate, population density, and building construction area.

Transportation: Automobile exhaust emissions are one of the major sources of PM_{2.5} and a significant part of air pollutants. While the automobile exhaust monitoring data is not available, two relative indexes from the statistical yearbook were selected: civil car ownership and number of public transportation vehicles in operation.

Energy use: Coal accounts for the majority of all energy use in China, and this particular energy structure makes pollution by coal burning especially serious. It is necessary to study the impact of coal consumption and energy efficiency on air pollution. Two secondary indexes were selected: per 10,000 yuan gross output value of industry energy consumption and total coal consumption.

Environmental protection: The Chinese central government has adopted various policies to reduce PM_{2.5} pollution and damage. Considering the efforts made by people in order to reduce PM_{2.5} pollution, this index was set up. One secondary index was selected: green coverage rate of built-up areas.

Output indexes were divided into two categories: one is the desirable output and the other is the undesirable output. The desirable output is the output that is beneficial to the overall goal and meets expectations, and the undesirable output is a byproduct of the beneficial output.

For economic applications in our setting, larger values mean better economic performance. At the same time, the pollutant discharge should be as low as possible. Therefore, we choose GDP as the desirable output and $PM_{2.5}$ as the undesirable output.

Because the weather data and $PM_{2.5}$ data are collected daily, and the social and economic data from the statistical yearbook is annual data, they need to be pre-processed before the calculation based on the requirements of chance constrained stochastic DEA. For weather data and $PM_{2.5}$ data, the annual mean and variance of data were calculated. For social and economic data, the variance of a sequence of 4-year data was calculated. When the calculation was with the chance constrained stochastic DEA, the variance was the same for 4 years.

Variable	Primary Index	Secondary Index	Definition	
	Weather Condition	Wind Speed Precipitation Temperature Atmospheric Pressure Sunshine Hours Relative Humidity	The velocity of air relative to a fixed place on the earth. The amount of precipitation in a region. The degree of air cooling and heating. Force per unit area exerted by an atmospheric column. Duration of sunshine in a day. The percentage of water vapor pressure to the saturated vapor pressure in the air.	
Input	Social Progress	Urbanization Rate Population Density Building Construction Area	The percentage of the total population living in urban areas. Number of people per square kilometer. Total construction area of the buildings constructed during the reporting period.	
	Transportation	Civil Car Ownership Number of Public Transportation Vehicles under Operation	Vehicles registered under civil vehicle licenses. Public transportation vehicles that serve residents, different vehicles are converted to the same standard.	
-	Energy Use	Per 10,000 Yuan Gross Output Value of Industry Energy Consumption Total Coal Consumption	The percentage of energy consumption by enterprises to the total industrial output value. The amount of coal consumed converted into the amount of standard coal.	
-	Environmental Protection	Green Coverage Rate of Built-up Area	The percentage of green coverage area to built-up area.	
Output	Expected Output Undesirable Output	Gross Regional Product PM _{2.5}	The market value of all final goods and services produced in a region. Concentration of particles in air with diameters less than or equal to 2.5 microns.	

Table 1. Description of indexes.

4. Result Analysis

In this section, we carried out the three designed sets of experiments, analyzed and compared the results from them.

4.1. Chance Constrained Stochastic DEA with All 14 Secondary Indexes as Input Indicators Subsection

Experiment I was carried out with all 14 secondary indexes as input indicators. In Formula (3), α is the risk level. In this section, we attempt to analyze the difference in the results obtained by selecting different risk levels α , and compare the results obtained in different years with the same risk level α . We took the risk levels $\alpha = 0.05$, $\alpha = 0.1$, $\alpha = 0.2$, $\alpha = 0.5$, $\alpha = 0.8$, $\alpha = 0.9$, and $\alpha = 0.95$ for the calculation. By using the chance constrained stochastic DEA model, we calculated the target value and stochastic efficiency for the thirteen cities in Jiangsu. The results of all 4 years are shown in Figure 2.



Figure 2. When α is greater than 0.5, the stochastic efficiency of different cities that changed with risk level in (**a**): 2013. (**b**): 2014. (**c**): 2015. (**d**): 2016.

For all 4 years, when the risk level was less than or equal to 0.5, the random efficiency of all cities was 1. When the risk level was greater than 0.5, most cities with changed stochastic efficiency exhibited a decrease in stochastic efficiency with increased risk level.

Table 2 below shows summary results of stochastic efficiency for all 13 cities and the years 2013–2016, with the same risk level in different years. Since the results obtained with different risk levels were highly similar, we used risk level $\alpha = 0.95$ as an example to illustrate.

From Table 2, we can see that from 2013 to 2016, when the risk level was $\alpha = 0.95$, the stochastic efficiency of 4 cities including Suzhou, Lianyungang, Zhenjiang and Suqian remained unchanged at 1, the other cities almost all showed increased stochastic efficiency over time, except Suzhou in 2015, Yancheng in 2014, Huaian in 2014 and Taizhou in 2015. The stochastic efficiencies of four cities including Wuxi, Nantong, Huai'an, and Yancheng in 2013 were not 1 and the stochastic efficiencies of 2016 were 1.

2013	2014	2015	2016
0.7048	0.763	0.7845	0.8127
0.9102	1	1	1
0.574	0.6234	0.6533	0.6933
0.787	0.7945	0.7542	0.8341
1	1	1	1
0.6958	0.7211	0.7342	1
1	1	1	1
0.4629	0.2572	1	1
1	0.3037	1	1
0.6061	0.7027	0.7398	0.7829
1	1	1	1
0.5685	0.664	0.6111	0.6697
1	1	1	1
	2013 0.7048 0.9102 0.574 0.787 1 0.6958 1 0.4629 1 0.6061 1 0.5685 1	2013 2014 0.7048 0.763 0.9102 1 0.574 0.6234 0.787 0.7945 1 1 0.6958 0.7211 1 1 0.4629 0.2572 1 0.3037 0.6061 0.7027 1 1 0.5685 0.664 1 1	$\begin{array}{c c c c c c c c c c c c c c c c c c c $

Table 2. Stochastic efficiency of 13 cities in Jiangsu Province from 2013 to 2016 when risk level $\alpha = 0.95$.

4.2. Chance Constrained Stochastic DEA with 1 Secondary Index Deleted Each Time

There were 14 input indicators in Experiment I, however, the result of Experiment I cannot distinguish which indicator has the greatest impact on the stochastic efficiency, as seen in Experiment I where all 14 secondary indexes were used. Therefore, experiment II was designed. By deleting 1 secondary index each time, we tried to find which secondary index has the greatest impact on the output indicators.

We took the risk level $\alpha = 0.05$, $\alpha = 0.1$, $\alpha = 0.2$, $\alpha = 0.5$, $\alpha = 0.8$, $\alpha = 0.9$, and $\alpha = 0.95$ for the calculation. After removing 1 secondary index, the remaining 13 secondary indexes were used as input indicators. Using the chance constrained stochastic DEA model, we calculated the target value and stochastic efficiency of 13 cities in Jiangsu. For $\alpha = 0.05$, $\alpha = 0.1$, $\alpha = 0.2$, $\alpha = 0.5$ both target value and stochastic efficiency were 1, therefore, these results were omitted. As the results of risk level $\alpha = 0.95$ was more distinguished than other risk levels ($\alpha = 0.8$, $\alpha = 0.9$), we used $\alpha = 0.95$ as an example to illustrate.

From Figure 3, it can be seen that for all 4 years, when the risk level was $\alpha = 0.95$, the change in stochastic efficiency caused by deleting the total coal consumption was the largest, accounting for 23.81% of the changes in stochastic efficiency. The second largest was population density, accounting for 22.62% of the changes in stochastic efficiency. The third largest was civil car ownership, accounting for 15.48% of the changes in stochastic efficiency. The per 10,000 yuan gross output value of industry energy consumption was the fourth, accounting for 9.52% of the changes in stochastic efficiency.

For all 4 years and at a risk level of α = 0.95, when the secondary indexes; wind speed, precipitation, air temperature and atmospheric pressure were deleted, the stochastic efficiency did not change, indicating that these four secondary indexes had no significant effect on the regional GDP and PM_{2.5}.

For all secondary indexes that lead to stochastic efficiency changes, removal of human factors accounted for 98.81% of the total changes in stochastic efficiency. Removal of natural causes (sunshine hours) accounted for only 1.19% of the total change. This clearly indicates that human activities play a leading role in PM_{2.5} pollution.

For 2013 alone, secondary indexes that changed the stochastic efficiency were (in order from high to low): total coal consumption, population density, civil car ownership and per 10,000 yuan gross output value of industry energy consumption. This order remained unchanged in the subsequent years 2014–2016.

By comparing the stochastic efficiency between the deleted and non-deleted secondary index for all 14 secondary indexes, and observing whether the stochastic efficiency changed or not, we could tell which secondary index had the greatest impact on the stochastic efficiency. Therefore, when dealing with PM_{2.5} pollution, corresponding policies and measures can be introduced.





Figure 3. Secondary indexes that changed the stochastic efficiency and their proportion (α = 0.95): total coal consumption (TCC), population density (PD), civil car ownership (CCO), per 10,000 yuan gross output value of industry energy consumption (WYEC), urbanization rate (UR), building construction area (BCA), number of public transportation vehicles under operation (NPT), green coverage rate of built-up area (GCR), sunshine hours (SH).

4.3. Chance Constrained Stochastic DEA with 1 Primary Index Deleted Each Time

As indicators may interact with each other, considering the interrelatedness and mutual influence between indicators, treating the same type of indexes as one group, we get five primary indexes. It is of great significance to examine which primary index has the largest impact on output indicators. Unlike experiment II, experiment III deleted one primary index at a time.

The risk levels $\alpha = 0.05$, $\alpha = 0.1$, $\alpha = 0.2$, $\alpha = 0.5$, $\alpha = 0.8$, $\alpha = 0.9$, and $\alpha = 0.95$ were used for the calculation. Four primary indexes were used as input indicators after removing the remaining one primary index. The chance constrained stochastic DEA model was used to calculate the target value and stochastic efficiency of 13 cities in Jiangsu. For $\alpha = 0.05$, $\alpha = 0.1$, $\alpha = 0.2$, and $\alpha = 0.5$, both the target value and stochastic efficiency were 1, therefore, these results were omitted. We used $\alpha = 0.95$ as an example to illustrate as we did in Figure 3.

From Figure 4, it can be seen that for all 4 years, when the risk level was $\alpha = 0.95$, the change in stochastic efficiency caused by deleting the social progress index was the largest, accounting for 42.86% of the changes in stochastic efficiency. The second largest was the energy use index, accounting for 30.00% of the changes in stochastic efficiency. The third largest was the transportation index, accounting for 20.00% of the changes in stochastic efficiency.

For all primary indexes that changed stochastic efficiency, the removal of the weather condition index accounted for only 1.43% of the total change in stochastic efficiency. Removal of the human factor index accounted for 98.57% of the total change in stochastic efficiency, meaning that human activities are more associated with $PM_{2.5}$ pollution. Hence, human activities are the main reason for $PM_{2.5}$ pollution.

For 2013 alone, primary indexes that changed the stochastic efficiency were (in order from high to low): social progress, energy use and transportation, which remained unchanged in the subsequent years from 2014 to 2016.

Now we summarize results from all three experiments:

- (1) For the same year, the stochastic efficiency decreased with an increased risk level. For the same risk level, the stochastic efficiency increased with later years.
- (2) By deleting one secondary index at a time, we determined the secondary index (total coal consumption) that had the greatest impact on output indicators. The second largest was population density, the third largest was civil car ownership, and per 10,000 yuan gross output value of industry energy consumption was the fourth. Removal of human factors accounted for 97.94% of the total changes in stochastic efficiency. Removal of natural causes (wind speed and sunshine hours) accounted for only 2.06% of the total change. While some meteorological elements such as wind speed and temperature do have a large impact on PM_{2.5}, we used annual data in this study and did not show this. Daily or hourly data will be more likely to show the importance of meteorological elements.
- (3) By deleting one primary index at a time, we found the primary index (social progress) that had the greatest impact on output indicators. The second largest was energy use, and the third largest was transportation. Removal of the human factor index accounted for 97.51% of the total changes in stochastic efficiency while removal of the weather condition index accounted for only 2.49% of the total change.
- (4) Human activities are deeply associated with PM_{2.5} pollution and the impact of meteorological factors is far less than that caused by human activities.



Figure 4. Primary indexes that changed the stochastic efficiency and their proportion ($\alpha = 0.95$). Social progress (SP), energy use (EU), transportation (T), environmental protection (EP), weather condition (WC).

5. Discussion

One of the main drawbacks of the traditional DEA method is that it cannot handle measurement errors. Traditional DEA is very sensitive to data errors because efficient units determine the efficiency [49]. Therefore, it is very important to incorporate random variations into DEA analysis. Compared with traditional DEA, chance constrained stochastic DEA has the following two advantages: it admits random input and output, and allows the evaluation unit to exceed the front edge under the given probability constraint.

In developed countries, people have experienced severe haze pollution during the industrialization era, e.g., the Los Angeles photochemical smog incident [50], the Donora smoke event [51], the great smog of London [52] and the Meuse valley fog [53]. These pollution incidents caused huge damage to people's health and personal property. These pollution incidents were controlled through targeted measures and persistent governance [54]. In China, the Central

Government took various measures to lower the level of haze pollution and reduce the damage caused by haze. On 29 February 2012, the government issued a newly revised "ambient air quality standard" that established a standard to dynamically monitor the haze pollution and officially release timely real time air quality indexes. For the first time, $PM_{2.5}$ was added to the national environmental quality standard. Since then, China's $PM_{2.5}$ pollution has been significantly reduced. The annual average concentration of $PM_{2.5}$ decreased from 72 µg/m³ in 2013 to 47 µg/m³ in 2016.

6. Conclusions

In this paper, we applied the chance constrained stochastic DEA to explore the relationship between PM_{2.5} pollution and meteorological elements, along with human activities, by analyzing the meteorological data, daily air pollutants monitoring data, and social and economic data derived from the Jiangsu Statistical Yearbook and the Statistical Yearbook of 13 prefectural cities in Jiangsu Province. The results showed that chance constrained stochastic DEA can clearly identify that human activities are the main cause of PM_{2.5} pollution in Jiangsu Province. Chance constrained stochastic DEA is rarely used to assess PM_{2.5} pollution. Our research results show that this method is very suitable for evaluating the PM_{2.5} pollution problem.

Author Contributions: Data curation, Y.Z., R.S. and G.W.; Formal analysis, Y.Z.; Funding acquisition, L.L.; Investigation, L.L. and G.W.; Methodology, R.S. and Z.G.; Project administration, Z.G.; Resources, L.L., Z.G. and M.B.; Software, M.B.; Supervision, Z.G.; Visualization, Y.Z.; Writing–original draft, Y.Z.; Writing–review & editing, Y.Z. and G.W.

Funding: This research was partially funded by the Major Project of the National Social Science Foundation grant number 16ZDA047, the National Natural Science Foundation of China grant number 71673145, the Report Project on the Development of Philosophy and Social Sciences of China's Ministry of Education grant number 13JBG004, the National Natural Science Foundation of China grant number 71171115, 71571104, the Reform Foundation of Postgraduate Education and Teaching in Jiangsu Province grant number JGKT10034, the Six Talent Peaks Project in Jiangsu Province grant number 2014-JY-014, the Postgraduate Research & Practice Innovation Program of Jiangsu Province grant number KYCX17_0904, the Major Project of Humanities and Social Sciences of Anhui Education Department grant number SK2015ZD07 And The APC was funded by the Major Project of the National Social Science Foundation.

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

References

- 1. Wu, D.; Bi, X.; Deng, X.; Li, F.; Tan, H.; Liao, G.; Huang, J. Effect of atmospheric haze on the deterioration of visibility over the Pearl River Delta. *Acta Meteorol. Sin.* **2007**, *21*, 215–223.
- 2. Zhou, Z.; Shao, T.; Qin, M.; Miao, X.; Chang, Y.; Sheng, W.; Wu, F.; Yu, Y. The effects of autophagy on vascular endothelial cells induced by airborne PM_{2.5}. *J. Environ. Sci.* **2018**, *66*, 182–187. [CrossRef]
- 3. Martinelli, N.; Olivieri, O.; Girelli, D. Air particulate matter and cardiovascular disease: A narrative review. *Eur. J. Int. Med.* **2013**, *24*, 295–302. [CrossRef] [PubMed]
- 4. Naota, M.; Shimada, A.; Morita, T.; Inoue, K.; Takano, H. Translocation pathway of the intratracheally instilled C60 fullerene from the lung into the blood circulation in the mouse: Possible association of diffusion and caveolae-mediated pinocytosis. *Toxicol. Pathol.* **2009**, *37*, 456–462. [CrossRef]
- 5. Bai, Y.; Sun, Q. Fine particulate matter air pollution and atherosclerosis: Mechanistic insights. *Biochim. et Biophys. Acta* **2016**, *1860*, 2863–2868. [CrossRef] [PubMed]
- 6. Xu, P.; Chen, Y.; Ye, X. Haze, air pollution, and health in China. Lancet 2013, 382, 2067. [CrossRef]
- 7. U.S. Environmental Protection Agency. Particulate Matter (PM_{2.5}) Trends. Available online: https://www.epa.gov/air-trends/particulate-matter-pm25-trends#pmnat (accessed on 2 July 2018).
- 8. Ministry of Ecology and Environment of the People's Republic of China. Report on the State of Environment in China. Available online: http://www.mee.gov.cn/hjzl/zghjzkgb/lnzghjzkgb/ (accessed on 2 July 2018).
- Maji, K.J.; Dikshit, A.K.; Arora, M.; Deshpande, A. Estimating premature mortality attributable to PM_{2.5} exposure and benefit of air pollution control policies in China for 2020. *Sci. Total Environ.* 2018, 612, 683–693. [CrossRef] [PubMed]

- 10. Che, H.; Zhang, X.; Li, Y.; Zhou, Z.; Qu, J.J.; Hao, X. Haze trends over the capital cities of 31 provinces in China, 1981–2005. *Theor. Appl. Climatol.* **2009**, *97*, 235–242. [CrossRef]
- 11. Ding, Y.H.; Liu, Y.J. Analysis of long-term variations of fog and haze in China in recent 50 years and their relations with atmospheric humidity. *Sci. China Earth Sci.* **2014**, *57*, 36–46. [CrossRef]
- Ministry of Ecological Environment of People's Republic of China. Report on the State of Environment in China. Available online: http://www.mee.gov.cn/hjzl/zghjzkgb/lnzghjzkgb/201805/ P020180531534645032372.pdf (accessed on 10 July 2018).
- Ministry of Ecological Environment of People's Republic of China. Analysis of the Source of PM_{2.5} in Beijing. Available online: http://dqhj.mep.gov.cn/dqhjzl/dqklwyjx/201709/t20170915_421691.shtml (accessed on 14 July 2018).
- 14. You, C.F.; Xu, X.C. Coal combustion and its pollution control in China. Energy 2010, 35, 4467–4472. [CrossRef]
- 15. Tang, X.; Snowden, S.; Mclellan, B.C.; Höök, M. Clean coal use in China: Challenges and policy implications. *Energy Policy* **2015**, *87*, 517–523. [CrossRef]
- Cheng, M.; Zhi, G.; Tang, W.; Liu, S.; Dang, H.; Guo, Z.; Du, J.; Du, X.; Zhang, W.; Zhang, Y.; et al. Air pollutant emission from the underestimated households' coal consumption source in China. *Sci. Total Environ.* 2016, 580, 641–650. [CrossRef] [PubMed]
- 17. Xue, Y.; Zhou, Z.; Nie, T.; Wang, K.; Nie, L.; Pan, T.; Wu, X.; Tian, H.; Zhong, L.; Li, J.; et al. Trends of multiple air pollutants emissions from residential coal combustion in Beijing and its implication on improving air quality for control measures. *Atmos. Environ.* **2016**, *142*, 303–312. [CrossRef]
- Zhang, Z.; Wang, W.; Cheng, M.; Liu, S.; Xu, J.; He, Y.; Meng, F. The contribution of residential coal combustion to PM_{2.5}, pollution over China's Beijing-Tianjin-Hebei region in winter. *Atmos. Environ.* 2017, 159, 147–161. [CrossRef]
- Fang, M.; Yao, C.X. Managing air quality in a rapidly developing nation: China. *Atmos. Environ.* 2009, 43, 79–86. [CrossRef]
- 20. Gao, J.; Tian, H.; Cheng, K.; Lu, L.; Zheng, M.; Wang, S.; Hao, J.; Wang, K.; Hua, S.; Zhu, C.; et al. The variation of chemical characteristics of PM_{2.5}, and PM₁₀, and formation causes during two haze pollution events in urban Beijing, China. *Atmos. Environ.* **2015**, *107*, 1–8. [CrossRef]
- Yang, J.; Ma, S.; Gao, B.; Li, X.; Zhang, Y.; Cai, J.; Li, M.; Yao, L.; Huang, B.; Zheng, M. Single particle mass spectral signatures from vehicle exhaust particles and the source apportionment of on-line PM_{2.5} by single particle aerosol mass spectrometry. *Sci. Total Environ.* 2017, *593*, 310–318. [CrossRef]
- 22. Gao, J.; Woodward, A.; Vardoulakis, S.; Kovats, S.; Wilkinson, P.; Li, L.; Xu, L.; Li, J.; Yang, J.; Li, J.; et al. Haze, public health and mitigation measures in China: A review of the current evidence for further policy response. *Sci. Total Environ.* **2017**, *578*, 148–157. [CrossRef]
- 23. Goel, A.; Kumar, P. Characterisation of nanoparticle emissions and exposure at traffic intersections through fast–response mobile and sequential measurements. *Atmos. Environ.* **2015**, *107*, 374–390. [CrossRef]
- 24. Kinsey, J.S.; Linna, K.J.; Squier, W.C.; Muleski, G.E.; Cowherd, C., Jr. Characterization of the fugitive particulate emissions from construction mud/dirt carryout. *Air Repair* **2004**, *54*, 1394–1404. [CrossRef]
- 25. Hassan, H.A.; Kumar, P.; Kakosimos, K.E. Flux estimation of fugitive particulate matter emissions from loose Calcisols at construction sites. *Atmos. Environ.* **2016**, *141*, 96–105. [CrossRef]
- Zhao, D.; Chen, H.; Li, X.; Ma, X. Air pollution and its influencing factors in China's hot spots. J. Clean. Prod. 2018, 185, 619–627. [CrossRef]
- 27. Lin, B.; Zhu, J. Changes in urban air quality during urbanization in China. J. Clean. Prod. 2018, 188, 312–321. [CrossRef]
- 28. Wu, J.; Zheng, H.; Zhe, F.; Xie, W.; Song, J. Study on the relationship between urbanization and Fine Particulate Matter (PM_{2.5}) concentration and its implication in China. *J. Clean. Prod.* **2018**, *182*, 872–882. [CrossRef]
- Zhou, C.; Chen, J.; Wang, S. Examining the effects of socioeconomic development on fine particulate matter (PM_{2.5}) in China's cities using spatial regression and the geographical detector technique. *Sci. Total Environ.* 2018, 619–620, 436–445. [CrossRef]
- 30. Fu, G.Q.; Xu, W.Y.; Yang, R.F.; Li, J.B.; Li, C.S. The distribution and trends of fog and haze in the North China Plain over the past 30 years. *Atmos. Chem. Phys. Discuss.* **2014**, *14*, 11949–11958. [CrossRef]

- Chambers, S.D.; Wang, F.; Williams, A.G.; Deng, X.; Zhang, H.; Lonati, G.; Crawford, J.; Griffiths, A.D.; Ianniello, A.; Allegrini, I. Quantifying the influences of atmospheric stability on air pollution in Lanzhou, China, using a radon-based stability monitor. *Atmos. Environ.* 2015, *107*, 233–243. [CrossRef]
- 32. Yang, S.; Ma, Y.L.; Duan, F.K.; He, K.B. Characteristics and formation of typical winter haze in Handan, one of the most polluted cities in China. *Sci. Total Environ.* **2017**, *613–614*, 1367–1375. [CrossRef]
- Li, W.; Liu, X.; Zhang, Y.; Sun, K.; Wu, Y.S.; Xue, R.; Zeng, L.; Qu, Y.; An, J. Characteristics and formation mechanism of regional haze episodes in the Pearl River Delta of China. *J. Environ. Sci.* 2018, 63, 236–249. [CrossRef]
- 34. Guo, L.C.; Zhang, Y.; Lin, H.; Zeng, W.; Liu, T.; Xiao, J.; Rutherford, S.; You, J.; Ma, W. The washout effects of rainfall on atmospheric particulate pollution in two Chinese cities. *Environ. Pollut.* **2016**, *215*, 195–202. [CrossRef]
- 35. Zhang, Y.; Yang, A.; Xiong, C.; Wang, T.; Zhang, Z. Feature selection using data envelopment analysis. *Knowl.-Based Syst.* **2014**, *64*, 70–80. [CrossRef]
- 36. Jin, J.; Zhou, D.; Zhou, P. Measuring environmental performance with stochastic environmental DEA: The case of APEC economies. *Econ. Modell.* **2014**, *38*, 80–86. [CrossRef]
- Charles, V.; Cornillier, F. Value of the stochastic efficiency in data envelopment analysis. *Expert Syst. Appl.* 2017, *81*, 349–357. [CrossRef]
- 38. Olesen, O.B.; Petersen, N.C. Stochastic data envelopment analysis—A review. *Eur. J. Oper. Res.* 2016, 251, 2–21. [CrossRef]
- 39. Wu, C.; Li, Y.; Liu, Q.; Wang, K. A stochastic DEA model considering undesirable outputs with weak disposability. *Math. Comput. Modell.* **2013**, *58*, 980–989. [CrossRef]
- 40. Chen, X.; Gong, Z. DEA efficiency of energy consumption in China's manufacturing sectors with environmental regulation policy constraints. *Sustainability* **2017**, *9*, 210. [CrossRef]
- 41. Gong, Z.; Chen, X. Analysis of Interval Data Envelopment Efficiency Model Considering Different Distribution Characteristics—Based on Environmental Performance Evaluation of the Manufacturing Industry. *Sustainability* **2017**, *9*, 2080. [CrossRef]
- 42. Wang, L.; Gong, Z.; Gao, G.; Wang, C. Can energy policies affect the cycle of carbon emissions? Case study on the energy consumption of industrial terminals in Shanghai, Jiangsu and Zhejiang. *Ecol. Ind.* **2017**, *83*, 1–12. [CrossRef]
- 43. Charnes, A.; Cooper, W.W.; Rhodes, E. Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **1978**, 2, 429–444. [CrossRef]
- 44. Banker, R.D.; Charnes, A.; Cooper, W.W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag. Sci.* **1984**, *30*, 1078–1092. [CrossRef]
- 45. Zhang, X.Y.; Wang, Y.Q.; Niu, T.; Zhang, X.C.; Gong, S.; Zhang, Y.M.; Sun, J.Y. Atmospheric aerosol compositions in China: Spatial/temporal variability, chemical signature, regional haze distribution and comparisons with global aerosols. *Atmos. Chem. Phys.* **2012**, *11*, 26571–26615. [CrossRef]
- 46. Lan, Y.X.; Wang, Y.M. Study of the relation between the chance constrained stochastic DEA efficiency and the risk level. *J. Syst. Eng.* **2014**, *29*, 423–432.
- 47. China Meteorological Science Data Sharing Service Network. China Daily Climate Data Set of Ground. Available online: http://data.cma.cn/data/cdcindex/cid/6d1b5efbdcbf9a58.html (accessed on 10 July 2018).
- 48. Statistics Bureau of Jiangsu Province. Jiangsu Statistical Yearbook. Available online: http://www.jssb.gov. cn/tjxxgk/tjsj/tjnq/nj2017/index_1508.html (accessed on 10 July 2018).
- 49. Mitropoulos, P.; Talias, M.A.; Mitropoulos, I. Combining stochastic DEA with Bayesian analysis to obtain statistical properties of the efficiency scores: An application to Greek public hospitals. *Eur. J. Oper. Res.* **2015**, 243, 302–311. [CrossRef]
- 50. Tiao, G.C.; Box, G.E.P.; Hamming, W.J. Analysis of Los Angeles photochemical smog data: A statistical overview. *J. Air Pollut. Control Assoc.* **1975**, *25*, 260–268. [CrossRef]
- 51. Helfand, W.H.; Lazarus, J.; Theerman, P. Donora, Pennsylvania: An environmental disaster of the 20th century. *Am. J. Public Health* **2001**, *91*, 553. [PubMed]
- Hunt, A.; Abraham, J.L.; Judson, B.; Berry, C.L. Toxicologic and epidemiologic clues from the characterization of the 1952 London smog fine particulate matter in archival autopsy lung tissues. *Environ. Health Perspect.* 2003, *111*, 1209–1214. [CrossRef] [PubMed]

- 53. Nemery, B.; Hoet, P.H.M.; Nemmar, A. The Meuse Valley fog of 1930: An air pollution disaster. *Lancet* 2001, 357, 704–708. [CrossRef]
- 54. Brunekreef, B.; Holgate, S.T. Air pollution and health. Lancet 2002, 360, 1233–1242. [CrossRef]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).