

國立臺灣大學公共衛生學院環境衛生研究所

碩士論文

Graduate Institute of Environmental Health

College of Public Health

National Taiwan University

Master thesis

運用傅立葉轉換紅外線光譜儀及輻射形光徑煙流分布

重建法定位污染源：重建演算法之評估

Using 2-D Radial Plume Mapping Technique with

OP-FTIR for Source Localization: Evaluation of

Reconstruction Algorithms

研究生：張世穎

Shih-Ying Chang

指導教授：吳章甫博士

Advisor: Chang-Fu Wu, PhD.

中華民國 97 年 10 月

October, 2008

## 中文摘要

輻射形光徑煙流重建法為使用光學遙測儀器進行污染物分布重建及污染源定位之技術。在此技術中，兩種主要之重建演算法被用來進行污染物分布重建-Smooth basis function minimization (SBFM)和 Non-negative least square (NNLS)。本研究之目的在比較此兩種主要重建法對於污染源定位及污染物分布重建之表現，此兩種演算法之相同處為重建出與測量得到之 PIC 相同的 PIC，差異處為 SBFM 使用預選之基本方程式描述污染物之分布，而 NNLS 則是直接重建逸散區域污染物之濃度。除此之外，在 SBFM 重建中，我們使用兩種不同之基本方程式(對稱和非對稱)來描述污染源。

本實驗由兩部份構成，第一部份為電腦模擬實驗，在電腦模擬中，首先產生 450 個基本分布並使用輻射形光徑煙流重建法重建污染物之分布及污染源。結果顯示 SBFM 使用非對稱基本方程式(bivariate lognormal distribution)做為基本方程式可以重建出完整之污染物；而當污染源位置接近 OP-FTIR 時，SBFM 使用對稱之基本方程式(bivariate Gaussian distribution)做為基本方程式時可以相當精確地重建出污染源之位置，但此時 NNLS 無法重建出正確之污染源。而當污染物遠離 OP-FTIR 時，使用 NNLS 進行重建會得到較好的結果而 SBFM 使用對稱之基本方程式無法定位污染源。

本研究之第二部份為實地實驗，使用開徑式傅立葉轉換紅外線光譜儀及輻射形光徑煙流重建法實地進行人為釋放之污染源定位。實地實驗之結果與電腦模擬之結果相似，當污染源遠離 OP-FTIR 時，使用 NNLS 重建演算法可以得到相當精確之污染源評估，而當污染源靠近 OP-FTIR 時，使用 SBFM 則可以得到較好的結果。此外，由三種重建方法所重建之污染源可以指出真實污染源之方向，佐以重建之污染源附近之較短之測線，可以加以判斷由何者重建出之污染源最接近真實污染源；當附近之短光徑沒有偵測到污染物時，可以選擇 NNLS 做為重建演算法，而當附近之短光徑測量到污染源時，則可使用 SBFM 做為重建演算法。

關鍵字：定位、空氣污染、開徑式傅立葉轉換紅外線光譜儀、輻射形光徑煙流重建法、光學遙測、逸散源

## Abstract

The OP-FTIR measurement combining the RPM technique is able to reconstruct the plume and thus localize the emission source. In this thesis, both the computational simulation and the field experiment are implemented. Two major kinds of the reconstruction algorithm used in RPM technique are evaluated. The first one is the smooth basis function minimization (SBFM) algorithm and the second one is the non-negative least square (NNLS) algorithm. The two algorithms are both implemented by fitting the reconstructed path integrated concentration (PIC) to the measured PIC. The differences are that the SBFM superimposes a basis function to describe the plume while the NNLS directly estimate the concentration value in the emission domain. In addition, two different kind of basis functions (symmetric and skewed) are used to describe the plume in SBFM reconstruction.

In the simulation analysis, 450 test distributions are generated to be localized by the RPM technique with different reconstruction algorithms. The result shows that the SBFM algorithm using the bivariate lognormal distribution as basis function gives the best result in both the aspects of plume reconstruction and source localization. Furthermore, when the plume is near the OP-FTIR, the SBFM reconstruction using bivariate Gaussian distribution as basis function may yield better result in the aspect of the source reconstruction comparing to the NNLS reconstruction. However, when the plume is far from the OP-FTIR, the NNLS reconstruction is able to localize the emission source more accurately than the SBFM using bivariate Gaussian distribution as basis function.

In the field experiment, four experiments with four pairs of different source locations are conducted to be localized by the RPM technique. The result shows that the reconstructed source locations by the three methods are able to point out the correct direction towards the real source. Furthermore, judging by the peripheral short monitoring lines, the reconstructed source location that is closest to the real source location can be chosen and gives the best estimation of the emission source location.

Keywords: source localization, plume reconstruction, optical remote sensing, OP-FTIR, air pollutant, RPM, CT

<b>Contents</b>	
中文摘要 .....	<b>I</b>
<b>Abstract</b> .....	<b>II</b>
<b>Chapter1. Introduction</b> .....	<b>1</b>
<b>1.1 Fourier transformed infrared spectrometer (FTIR)</b> .....	1
<b>1.2 Traditional methods for source localization</b> .....	4
1.2.1 The area sampling array method .....	4
1.2.2 The computed tomographic (CT) method .....	6
1.2.3 The application of the CT technique .....	9
<b>1.3 Radial plume mapping (RPM) technique</b> .....	10
1.3.1 RPM with SBFM reconstruction algorithm .....	10
1.3.2 The RPM with “grid based” reconstruction algorithm .....	12
<b>1.4 Study design and objectives</b> .....	15
<b>Chapter2. Materials and Methods</b> .....	<b>23</b>
<b>2.1 Data collection</b> .....	23
<b>2.2 Data analysis</b> .....	27
2.2.1 The computational simulation .....	27
2.2.2 The field experiment.....	45
<b>3.1 Computational simulation results</b> .....	53
3.1.1 The plume reconstruction .....	53
3.1.2 The reliability of source localization .....	60
3.1.3 The prior screening process .....	63
3.1.4 The uncertainty analysis of SBFM reconstruction.....	66
<b>3.2 Field experiment results</b> .....	71
3.2.1 The spectrum quantification .....	71
3.2.2 The source localization .....	72
3.2.3 The reliability of the reconstruction result .....	80
<b>Chapter4. Conclusions and suggestions</b> .....	<b>116</b>
<b>4.1 The simulation experiment</b> .....	116
<b>4.2 The field study</b> .....	117
<b>4.3 Suggestions</b> .....	117
<b>4.4 Limitations</b> .....	119
<b>References</b> .....	<b>120</b>

## Figure contents

Figure 1.1 The example beam geometry described by Todd et al. ....	17
Figure 1.2 The beam geometry proposed by Drescher et al. ....	18
Figure 1.3 The experimental set up of the experiment on the waste lagoon. ....	18
Figure 1.4 The beam geometry of 1-D RPM. ....	19
Figure 1.5 The beam geometry of 2-D RPM. ....	20
Figure 1.6 The reconstruction result in the studies using NNLS reconstruction .....	21
Figure 2.1 The beam geometries used in the computational simulation.....	48
Figure 2.2 The tracer gas releasing device.....	49
Figure 2.3 The experimental setup of the field experiment .....	50
Figure 2.4 The procedure of generating error map of single test distribution .....	51
Figure 2.5 The procedure of generating overall error map.....	52
Figure 3.1 The reconstruction results for different reconstruction methods.....	83
Figure 3.2 The reconstruction results for different reconstruction methods.....	84
Figure 3.3 The reconstruction results for different reconstruction methods.....	85
Figure 3.4 The error map of Geometry <sub>extend</sub> using different size of test distributions ....	86
Figure 3.5 The error map of Geometry <sub>center</sub> using different size of test distributions ....	87
Figure 3.6 The error map of Geometry <sub>extend</sub> using different size of test distributions ....	88
Figure 3.7 The error map of Geometry <sub>center</sub> using different size of test distributions ....	89
Figure 3.8 The error map of Geometry <sub>extend</sub> using different size of test distributions ....	90
Figure 3.9 The error map of Geometry <sub>center</sub> using different size of test distributions ....	91
Figure 3.10 The error map .....	92
Figure 3.11 The time series plot of PIC in SF <sub>6</sub> in four experiments.....	93
Figure 3.12 The time series plot of PIC in N <sub>2</sub> O in four experiments .....	94
Figure 3.13 The reconstruction result of SF <sub>6</sub> .....	95
Figure 3.14 The reconstruction result of SF <sub>6</sub> .....	96
Figure 3.15 The reconstruction result of N <sub>2</sub> O.....	97
Figure 3.16 The reconstruction result of N <sub>2</sub> O.....	98
Figure 3.17 The average result of SF <sub>6</sub> .....	99
Figure 3.18 The average result of N <sub>2</sub> O .....	100
Figure 3.19 The wind rose of the four field experiments .....	101

## Table contents

Table 3.1 The overall statistics of the reconstruction quality .....	102
Table 3.2 The t-test result for peak error between different reconstruction methods ...	103
Table 3.3 The summary statistics of the reconstruction.....	104
Table 3.4 The summary statistics of the reconstruction quality.....	105
Table 3.5 Sensitivity analysis of the center beam geometry .....	106
Table 3.6 Sensitivity analysis of the extend beam geometry .....	107
Table 3.7 The overall statistics of reconstruction quality .....	108
Table 3.8 The regression analysis result .....	109
Table 3.9 ANOVA analysis between different dPIC .....	110
Table 3.10 The detection limit of SF <sub>6</sub> and N <sub>2</sub> O in each monitoring line .....	110
Table 3.11 The summary statistics of PIC data in each experiment .....	111
Table 3.12 The reconstruction result of field experiment of two tracer gases .....	112
Table 3.13 The mean reconstruction result of field experiment of two tracer gases ....	113
Table 3.14 The chosen reconstructed result by the nearby beam paths.....	114
Table 3.15 The wind data of the field experiment.....	115



## Chapter1. Introduction

Air pollution is an important issue nowadays. Since the 18<sup>th</sup> century, the industrial revolution has promoted the life quality of human however the pollution caused by the manufacture and factory has also affected human's health. Thus, monitoring air pollutants plays an important role for preventing and controlling the air pollution. Several devices have been developed to reach the goal described above. Instruments such as time integrated samplers (i.e. charcoal tube, canister, impactor, sampler bag), direct reading instruments (i.e. Photo ionization detector) and optical remote sensing instruments (i.e. Fourier transform infrared spectroscopy, UV-differential optical absorption spectroscopy) are able to monitor the toxic gaseous in the ambient.



In order to provide sufficient information for pollution control, the source localization along with pollution mapping is also needed. Using the devices previously mentioned, scientists have developed several techniques for both source localization and pollution mapping. Besides, localizing the emission source in a timely fashion is also important to provide the information for pollution control.

### *1.1 Fourier transformed infrared spectrometer (FTIR)*

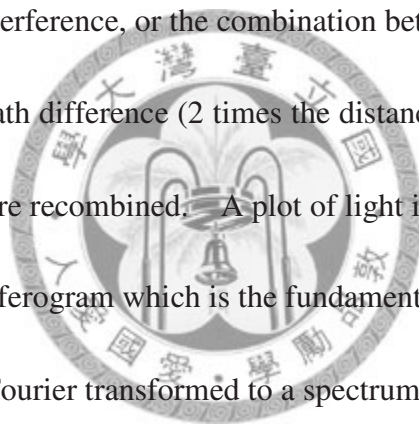
Fourier transformed infrared spectrometer (FTIR) is one of the optical remote sensing instrument which is able to detect the ambient chemical substance both qualitatively and quantitatively<sup>1</sup>. It works based on the theory of infrared spectroscopy which means the study of the interaction between the infrared and the matter. That is, when the infrared radiation interacts with the chemical substance it can be absorbed, causing the molecule bonds to vibrate furthermore, the functional groups in the chemical tend to absorb the infrared radiation in the same wavelength. Thus, there is an association between the structure of the chemical and the wavelength at which the chemical absorbs infrared radiation. This property of the substance allows the structure of the unknown chemical to be identified from the infrared spectrum. Along with the chemical identification, the infrared spectrum can also provide the information for quantification. The basic theory for the FTIR to quantify substance is Beer's law which describes the association between the absorbance of the infrared radiation and the chemical's concentration.

The open path FTIR (OP-FTIR) can only provide the concentration data in the form of path integrated concentration (PIC) which means the integrated concentration along the beam path. Thus it has a unit of "concentration" times "length" (e.g. ppm×meter). Once the beam path length is measured, the mean concentration along the monitoring line can also be calculated by dividing the PIC by the beam path length. As a result, the OP-FTIR can only provide the concentration information in the form of



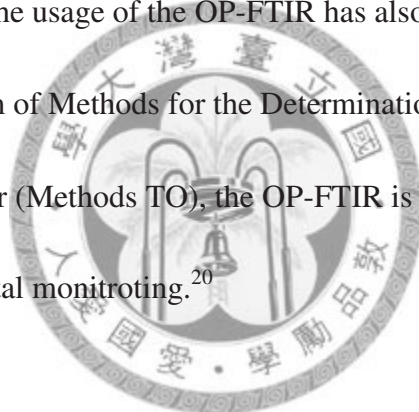
mean concentration which represents the mean concentration where the ray passes through.

The open path FTIR (OP-FTIR) is mainly composed of four parts which are infrared source, interferometer, sample compartment and detector. The most important part in the FTIR is the interferometer. In the interferometer, the infrared emitted from the source is first split into two beams by the beamsplitter. The two beams are then reflected by a fixed mirror and a moving mirror respectively. The constructive interference, destructive interference, or the combination between the both takes place depending on the optical path difference (2 times the distance the moving mirror travels) when the reflected beams are recombined. A plot of light intensity to the optical path difference is called an interferogram which is the fundamental of the FTIR measurement. The interferogram is then Fourier transformed to a spectrum. The sample compartment is used to contain samples however the OP-FTIR does not have one. Instead of passing through the sample compartment, the infrared emitted from the OP-FTIR passes through the ambient directly to measure the chemical substance. The last part of the OP-FTIR is the detector. The detector is able to detect the beam intensity of the entered infrared. To use the OP-FTIR to monitoring the ambient pollutant, a background spectrum must first be collected. By transforming the sample spectrum to the absorbance spectrum with the background spectrum, the chemical substance can be identified and quantified. The time for the OP-FTIR to obtain a



single spectrum depends on the resolution of the spectrum. For the instrument used in this study, it takes about 10 seconds for a single scan. Thus comparing to the time integrated samplers for monitoring ambient pollutant, the optical remote sensing instrument is able to provide real-time concentration data because the spectrum analysis process is faster and can be done *in situ*.

The OP-FTIR has now been used in many fields such as pollution monitoring<sup>2-5</sup>, exposure assessment<sup>6-8</sup>, pollution flux measurement<sup>9, 10</sup>, source localization and pollution mapping<sup>11-19</sup>. The usage of the OP-FTIR has also been verified by the US EPA. In the Compendium of Methods for the Determination of Toxic Organic Compounds in Ambient Air (Methods TO), the OP-FTIR is one of the accepted instrument for environmental monitoring.<sup>20</sup>



## **1.2 Traditional methods for source localization**

### **1.2.1 The area sampling array method**

Traditionally, the “area sampling array”<sup>7</sup> has been applied to localize the pollution source. When applying the “area sampling array”, a set of samplers are located in the emission site. After hours of sampling, the samples are then sent back to the lab for analysis. In 2005, Chen et al. placed 25 stainless steel canisters at a petrochemical

plant in order to localize the pollution source<sup>21</sup>. The sampling procedure lasted for 1 year. Each time after the sampling, the samples were sent back to be analyzed by the LC-Mass spectrometer. The concentration data was then input to the Surfer software to create the contour plot of the distribution of the chemicals. Although the result showed that this sampling method is quite accurate when comparing to real source location, there are still some limitations when applying this technology to localize the emission source. First of all, most of these kinds of point samplers are time integrated samplers, which can only provide concentration data with limited temporal resolution (canister in this case). In that study, each canister collected the samples for approximately 3 hours a day (11:00 am to 2:00 pm) thus having the temporal resolution of 3 hours. This poor temporal resolution data might ignore the short term emission and thus limit the application on risk assessment or exposure assessment. Second, the spatial resolution of concentration might also be limited because the limited amount of samplers. It is impossible for researchers to place a large amount of samplers among all of the area in the emission domain. Therefore, the concentration data at the location without samplers must be estimated by further interpolation or extrapolation. In that study, 25 canisters are located at certain sites in a petrochemical plant and a statistical method has been applied to estimate the concentration data at the un-sampled sites. These estimated concentration data might not be accurate comparing to the real concentration. Third, the sample transportation and sample analysis take time thus it is

not suitable for using this kind of technique in an emergency situation.

### *1.2.2 The computed tomographic (CT) method*

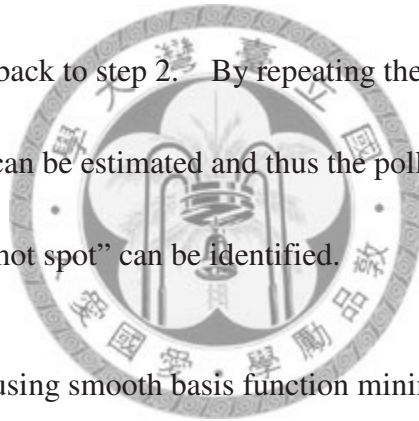
Considering the problem met in applying the area sampling array method, the computerized tomographic (CT) technology using optical remote sensing instrument has been proposed to map the pollutants. The CT technology is widely used in the medical field, which is a tool for imagine diagnosis. In the aspect of mapping pollutants, the CT technology means combining the optical remote sensing measurement and reconstruction algorithm to further map the pollutant. The multiple monitoring lines are arranged to form a network (called beam geometry) and the collected concentration data is input to the reconstructing algorithm to further map the pollutant. In 1979<sup>22</sup>, Byer et al. first proposed using CT technology to map the pollutant. The tunable-laser source is located in the middle of a circle in the measuring site. Several sets of cylindrical mirrors and detectors are equally located on the circumference of this circle to form a beam network. The laser beam emitted from the source is then reflected by the cylindrical mirror to the detector. By comparing the emitted and received beam intensity, the pollution in the measuring site can be mapped and the pollution source can be identified.<sup>22</sup> However, in order to map the pollution with a radius of 10 meters, more than 300 mirrors and detectors will be needed. Such huge amount of detectors

will be costly and the complicated beam geometry might limit the application of this technology. Since then, additional studies about using CT to map the pollutant are proposed and evaluated.<sup>10, 23-26</sup>

In all of the CT technique described in the last paragraph, several monitoring lines must be arranged to form a network that provides adequate concentration data for plume reconstruction. Many kinds of beam geometries have been proposed and all of them are composed of intersecting beam paths. To further investigate the impact of different beam geometry, Todd has evaluated the performance of different beam geometry by computational simulation<sup>23</sup>. A total of 13 different beam geometries with different number of intersecting beam path (120 to 288 paths) and different number of detectors (1 to 4 detectors) are applied to reconstruct a series of test maps. An example beam geometry is shown in Figure 1.1. The result shows that as the number of detectors increases, the performance of the reconstruction is improved despite the same number of beam path. However, when extra beam paths are added, the number of artifacts (reconstructed peaks that do not exist in the test map) and the peak location error (distance between real and reconstructed source locations) is decreased.

Another important part of the CT technique is the reconstructing algorithm. The most commonly used iterative algorithms in CT are Algebraic Reconstruction Technique (ART)<sup>27</sup>, Maximum Likelihood with Expectation Maximization (MLEM)<sup>10</sup>,

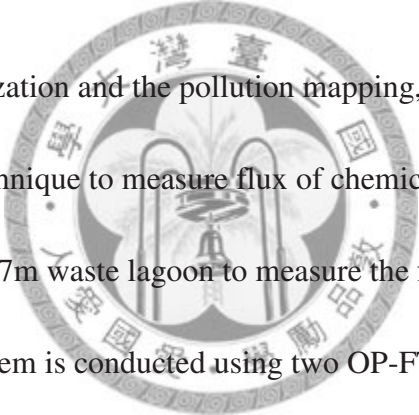
<sup>28</sup> and Multiplicative Algebraic Reconstruction Technique (MART)<sup>10, 29</sup>. In all these three reconstructing methods, the area of interested is first divided into several small grids and the concentration in each grid is assumed to be homogeneous and non negative. The first step of the reconstruction is to make an initial guess for the concentration in each grid. The second step is to calculate the reconstructed path integrated concentration (PIC) data for each beam path by summing up the concentration value of all the grids through which the ray passes. The third step is to adjust the concentration value in each grid by comparing the reconstructed PIC and the collected PIC data then go back to step 2. By repeating the second and third step, the concentration in each grid can be estimated and thus the pollutant can be mapped and the source location or the “hot spot” can be identified.



In 1996, an approach using smooth basis function minimization algorithm combining Fourier transformed infrared spectrometer (FTIR) measurement has been proposed to map the pollutant in the air.<sup>30</sup> Different from the “grid based” algorithm discussed in the previous paragraph, the SBFM algorithm applies a smooth basis function to describe the distribution of the plume. In that study, a series of experiments are conducted to evaluate the performance of the ART algorithm and the SBFM algorithm. In order to decrease the scanning time with limited amount of remote sensing instrument, only one OP-FTIR is located in the middle of the experimental domain to scan the flat mirrors and retroreflectors on the edge (Figure 1.2).

A total of 56 monitoring lines are conducted in that study. The result shows that the SBFM algorithm is able to reconstruct better in the aspect of plume's distribution and the number of artifacts under this beam geometry. The parameter used to describe the distributional difference between the real and reconstructed plume (the smaller the better, 0 represents perfect fit) is 0.11 for SBFM and 0.41 for ART.

### *1.2.3 The application of the CT technique*



Despite the source localization and the pollution mapping, researchers have also developed using the CT technique to measure flux of chemical emissions.<sup>10</sup> The study is conducted on a 255m×107m waste lagoon to measure the flux of nitrogen. A computed tomographic system is conducted using two OP-FTIR and total of 16 intersecting beam paths on the surface of the waste lagoon (Figure 1.3). Two tracer gases, SF<sub>6</sub> and CH<sub>4</sub> are released simultaneously during the FTIR measurement to provide the flux calculation information. Similar to the CT technique used to map the pollutant, the MLEM and MART algorithms are used to reconstruct the concentration of both the tracer gas and the target gas in each virtual grid. Although the flux calculation process has encountered many difficulties, this study is the first field implementation of this kind of CT system in the site of this scale.

### *1.3 Radial plume mapping (RPM) technique*

In most of the previous CT techniques, the applied beam geometry is quite complicated<sup>10, 12, 13, 22-24, 26, 30</sup> thus limits the field implementation. It can be seen that most of these applied beam geometries are composed of intersecting monitoring lines and multiple detectors or remote sensing instrument. Not only difficult to apply, setting up this kind of system would be expensive due to the multiple detectors.

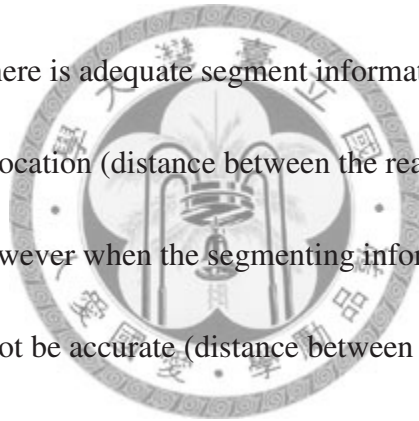
#### *1.3.1 RPM with SBFM reconstruction algorithm*

In 1999, Hashmonay et al. proposed using FTIR measurement combining wind data to localize the pollution source location.<sup>15</sup> In that study, a one-dimensional beam geometry is set up downwind to the releasing tracer gas. Three to four monitoring lines are arranged along a line (Figure 1.4). Each time when the wind direction changes the peak location on monitoring lines can be identified by the SBFM reconstruction<sup>31, 32</sup>. The line equation from the peak location reconstructed on the monitoring line to the orientation of wind direction is then calculated. The intersection of the line equations with different wind directions is the source location. Due to the relatively simple beam geometry at the down wind site, it is possible to localize the source location without setting the monitoring instruments or sampling device in the measuring site and the relatively simple beam geometry make it possible to apply.

A two dimensional beam geometry has also been proposed in 1999<sup>16</sup>. Similar to



the study in 1996<sup>30</sup>, the SBFM reconstruction uses bi-variate Gaussian distribution as smooth basis function to describe the pollution's plume. However the radial beam geometry, instead of the complicate beam geometry, has been applied (Figure 1.5). In that computational simulation study, the OP-FTIR is set in the corner of the experimental domain to limit the rotation of the FTIR in 45°. The radial beam geometry means to arrange the monitoring lines in a radial form which prevents the intersection of the monitoring beam path. A validation experiment conducted in a ventilation chamber has also been proposed to confirm the simulation result.<sup>18</sup> The result suggests that when there is adequate segment information, the SBFM algorithm is able to localize the source location (distance between the real and reconstructed source location = 0.3 meter). However when the segmenting information is limited, the reconstruction result may not be accurate (distance between the real and reconstructed source location = 1.2 meter). Another study using the non-overlapping geometry has also been conducted.<sup>33</sup> Different from the study of the one previously mentioned, the OP-FTIR is located in the middle of the experimental domain. Thirty random points are generated as the endpoint of monitoring lines and the OP-FTIR instrument is directed to scan the retrorereflectors in 360° rotation. Several experiments are conducted with tracer gas released at different locations. The collected PIC data is then reconstructed by the SBFM algorithm. The results show that the RPM technique under current beam geometry is able to reconstruct the distribution of the plume (The



correlation  $R = 0.8$ ). The relatively poor result may come from the inadequate PIC information. In those results with poor performance, the tracer gas is located on the edge of the experimental domain thus the monitoring line is not able to detect the plume.

Other than pollution mapping, the RPM technique has also been applied to measure the flux of the emission.<sup>34</sup> In this method, the monitoring lines are arranged in order to form vertical radial beam geometry. In that simulation study, the retroreflectors are placed on a tower with different elevations. Since the objective of this technique is to measure the flux of emission, the focus would be on the total concentration that emits from the area rather than the distribution of the plume. The SBFM reconstruction algorithm is also applied in this technique to estimate the total emission of the chemical of interest. This technology, known as Vertical Radial Plume Mapping (VRPM) is also described in the US EPA other test method 10: Optical Remote Sensing for Emission Characterization from Non-Point Sources<sup>35</sup> and is applied and demonstrated at various sites for measuring emission flux.<sup>36</sup>

### *1.3.2 The RPM with “grid based” reconstruction algorithm*

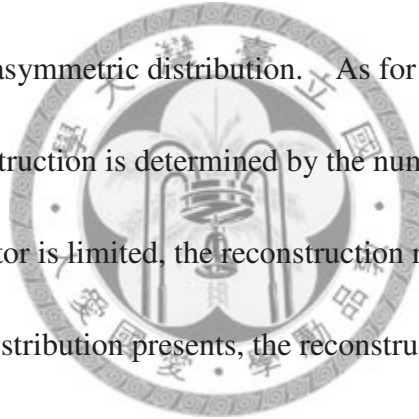
The relatively simple beam geometry of radial beam geometry makes the computed tomography for mapping pollutants possible in application. However the

performance between the conventional intersected beam geometry and the radial beam geometry still remains unknown. Another study further investigate this issue<sup>19</sup>. A computational simulation and a series of field experiments are conducted. In the computational simulation, two types of beam geometry are simulated with a conventional reconstructing algorithm. The first approach is the radial beam geometry (RCT) with 16 retroreflectors and the second is the conventional intersected beam geometry (CCT) with 16 beam paths. Instead of the SBFM reconstruction algorithm, the reconstructing algorithm used in both approaches is MART in which each measurement is compared to a predicted value computed from the current image estimate. Although the simulation result shows that the RCT approach performs better than the CCT approach (the modified correlation =1 and 0.94 respectively), the author suggests that due to the distribution of the plume in the real world is complicated (weather, wind effect etc.), it can not be concluded that the RCT is better than the CCT however the result from both reconstructing geometry is at least comparable. The conclusion of the computational simulation suggests that the relatively simple beam geometry of RCT is enough to map the plume. This is confirmed in the field experiment. In the field study, several experiments are conducted with the RCT technique at a 25m×35m domain. In each experiment, one tracer gas is released and the FTIR is used to collect the PIC data with nine retroreflectors. The result shows that the real and reconstructed source locations are in the same pixel of the experimental

domain (Figure 1.6(b)).

It can be seen that the “grid based” algorithm can also be applied in the radial plume mapping technique in last paragraph. This has also been described in the U.S EPA other test method (OTM10)<sup>35</sup>. In the OTM10, the radial plume mapping technique used to map the pollutant or searching for the “hot zone” is called horizontal radial plume mapping (HRPM). The non negative least square (NNLS) algorithm is used as the reconstruction algorithm in this method. Same as the “grid based” method described in the CT technique, the experimental domain must first be divided into several grids. The concentration in each grid is then reconstructed. However, the number of the grids divided must be as the same as the number of monitoring lines to prevent the under-determining situation. In the OTM10, an example of beam geometry composed of 9 retroreflectors is presented. The experimental domain is divided into 3×3 grids. Thus the resolution of the reconstruction would be quite coarse. The RPM using NNLS has also been implemented in a landfill by Hashmonay<sup>37</sup>. In that study, the RPM technology with NNLS reconstruction is practiced to locate the “hot spot” of a landfill in which the emission source is not a point source. The result has shown that the NNLS reconstruction is able to locate the “hot spot” of the emission and is comparable to the reconstruction by the RPM using MART. However, the reconstruction time is 100 times faster than the MART.

Although previous studies have proved that the radial plume mapping technique is able to localize the emission source location, there is still uncertainty in this technique. For example, when the segmenting information is limited, the RPM technique using SBFM reconstruction algorithm may not be able to map the pollutant accurately<sup>18</sup>. Furthermore, when applying SBFM, the bi-variate Gaussian distribution is used as basis function in most of previous studies<sup>16, 18</sup>. However the plume may not be a symmetric distribution when the wind presents. When wind presents the plume might be an asymmetric distribution. Thus it might cause error when using a symmetric distribution to describe an asymmetric distribution. As for the “grid based” method, the resolution of the reconstruction is determined by the number of the monitoring lines. If the available restroreflector is limited, the reconstruction result may be a coarse estimate. When a steep distribution presents, the reconstruction might underestimate the peak concentration.<sup>24, 25</sup>



#### ***1.4 Study design and objectives***

Figure 1.7 has shown the flow chart of this thesis. The main objective of this thesis is to further investigate the performance of the RPM technique. Not only the SBFM algorithm (non grid based) but also the NNLS reconstruction algorithm (grid based) is discussed in this thesis. A series of computational simulation study is conducted to compare the performance for the “non grid based” algorithm (i.e. SBFM)

and the “grid based“ algorithm (i.e. NNLS). Furthermore, the effect of using different smooth basis function (Skewed distribution V.S. symmetric distribution) in the SBFM algorithm is also investigated. A total of 450 artificially generated test distributions are reconstructed by the two algorithms respectively. We have also investigated the uncertainty when using the SBFM algorithm to reconstruct the plume in the computation simulation. After the simulation experiment, a series of field experiments are then implemented to verify the results in the computational simulation. A total of 4 experiments are conducted with four pairs of artificially released source locations. The RPM technique is applied to localize these emission sources. There are several innovative aspects in this study that is different from the others. First, two emission sources are released and localized simultaneously. Second, the effect of two different beam geometries is discussed in the computational simulation. Third, the SBFM and NNLS are both used to localize the emission source. Fourth, the uncertainty of the SBFM reconstruction to localize the source location is discussed by the concept of error map. The last, a skewed distribution (i.e. bivariate lognormal distribution) is used as the basis function in the SBFM reconstruction.

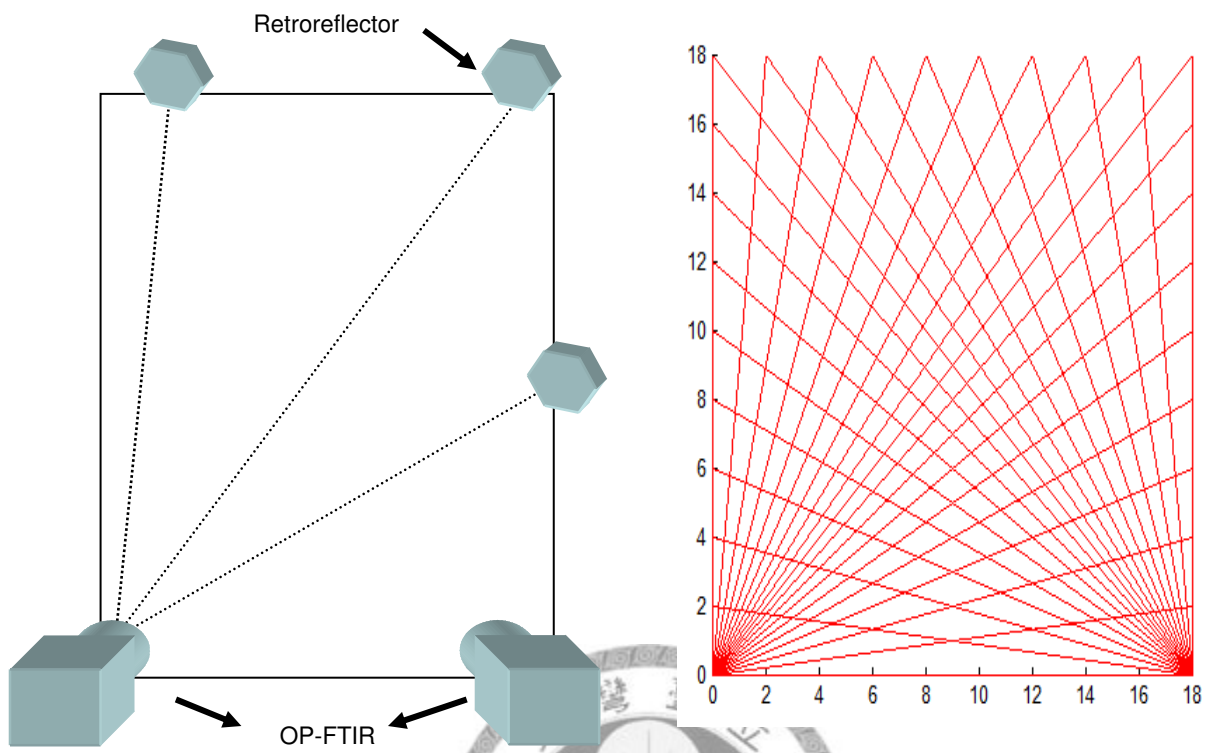


Figure 1.1 The example beam geometry described by Todd et al.<sup>23</sup> Two optical remote sensing instruments are located in the corner of the domain

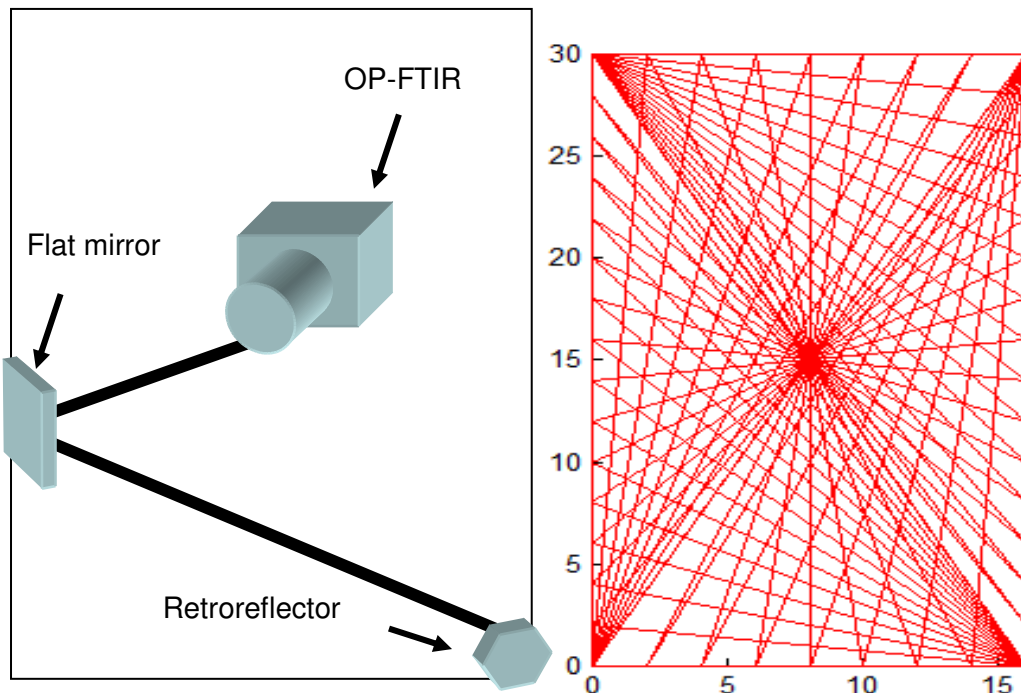


Figure 1.2 The beam geometry proposed by Drescher et al.<sup>30</sup> The optical remote sensing instrument is located in the middle of the domain. Four retroreflectors are in the corner and several flat mirrors are on the edge.

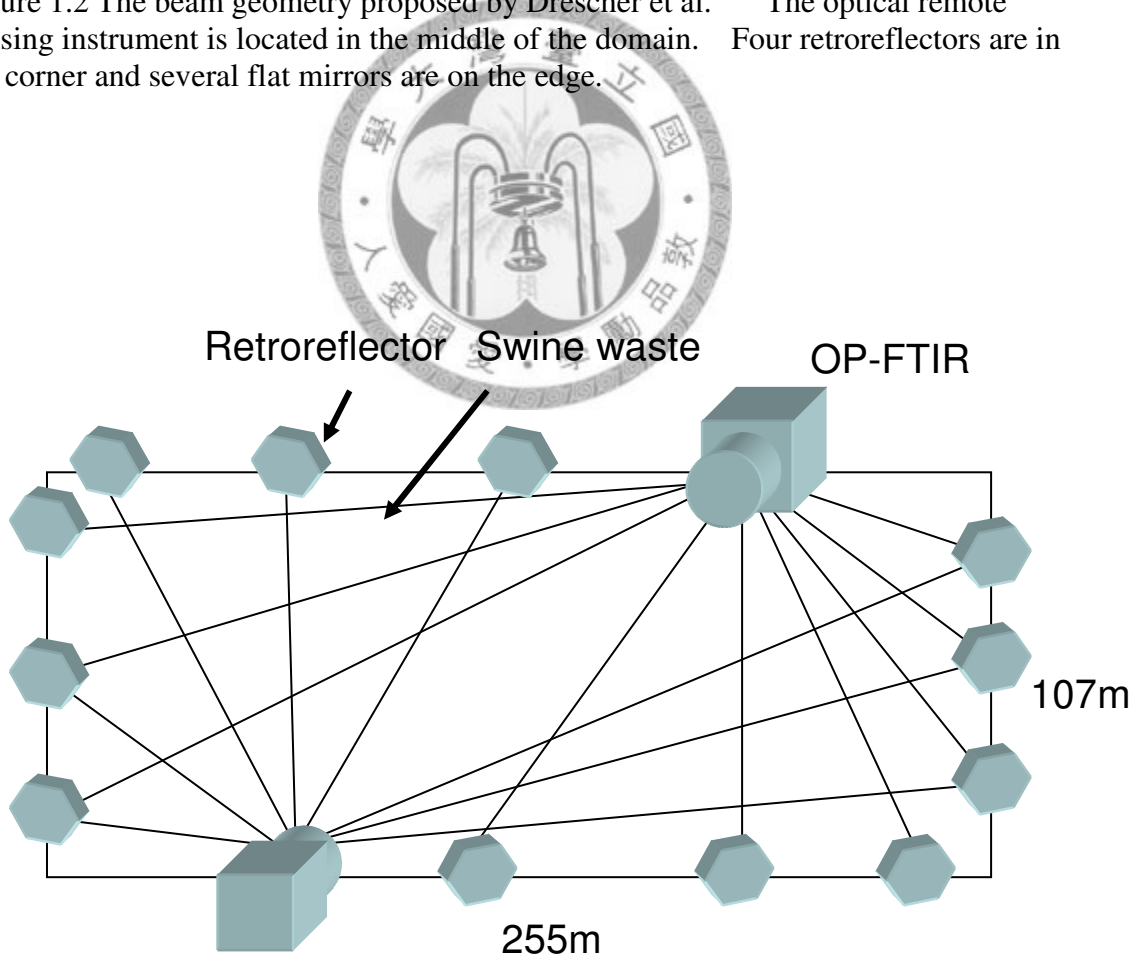


Figure 1.3 The experimental set up of the experiment on the waste lagoon. Modified from Todd et al.<sup>10</sup>



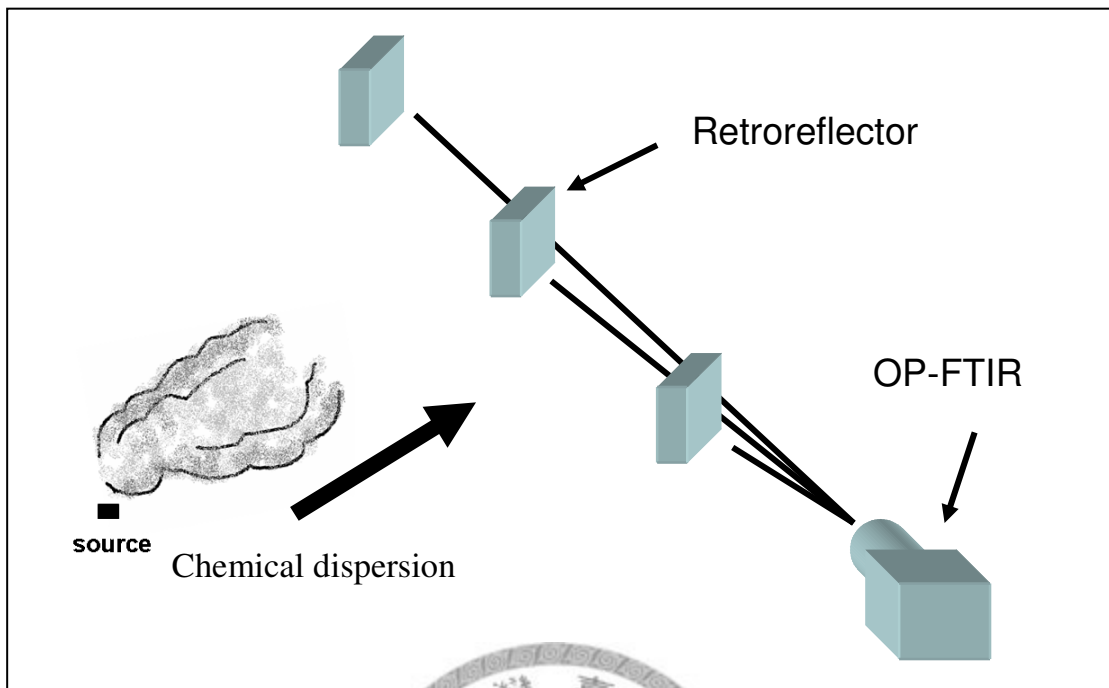
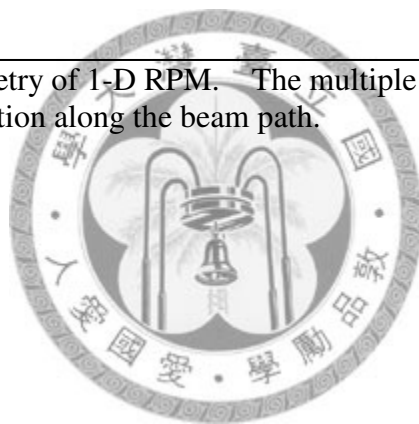


Figure 1.4 The beam geometry of 1-D RPM. The multiple ray paths are able to localize the peak concentration along the beam path.



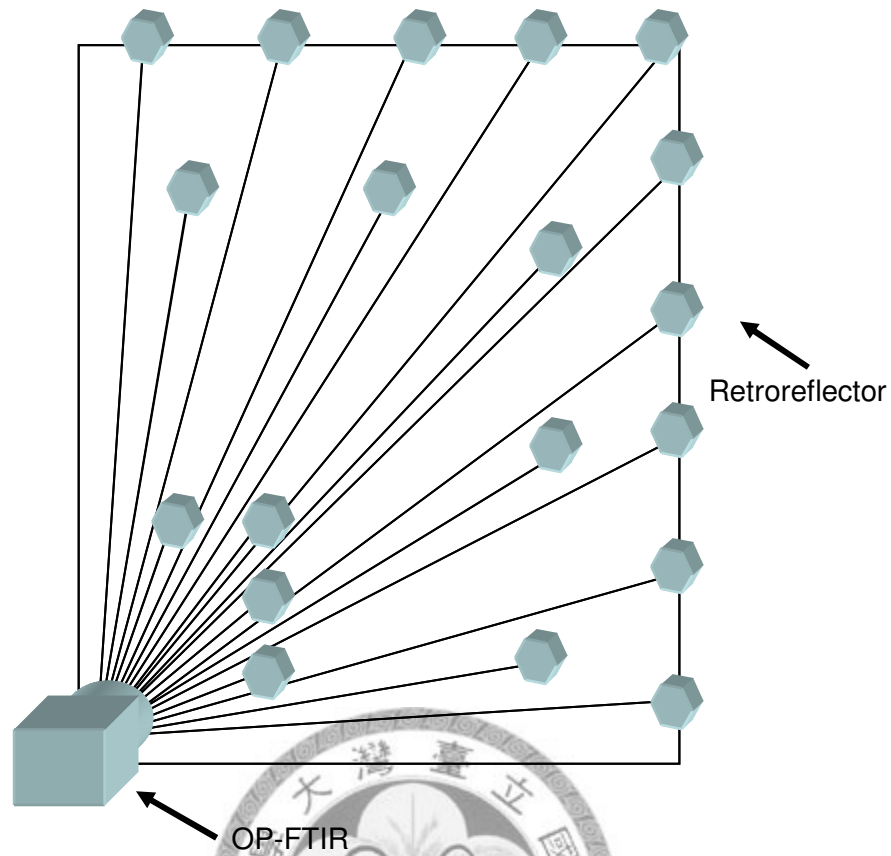
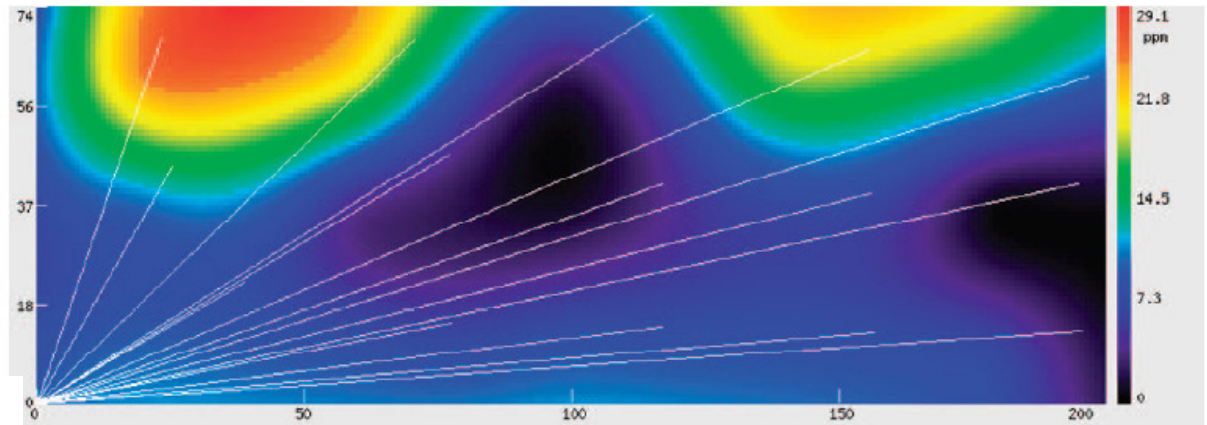


Figure 1.5 The beam geometry of 2-D RPM. The OP-FTIR is located in a corner of the domain thus limits the scan range within  $45^\circ$ . Each ray path ends at a retroreflector.

(a)



(b)

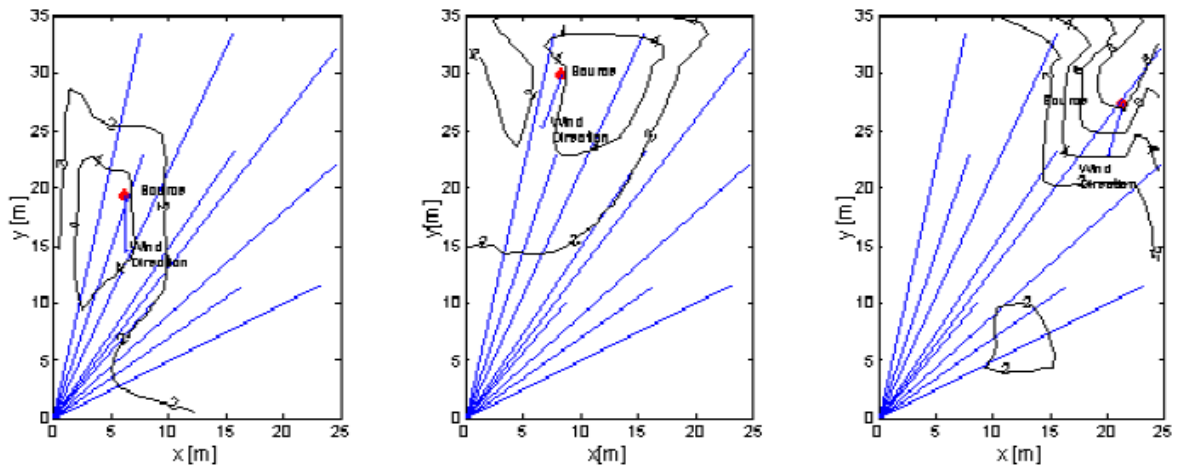


Figure 1.6 The reconstruction result in the studies using NNLS reconstruction. (a) The study “Theoretical Evaluation of a Method for Locating Gaseous Emission Hot Spots”<sup>37</sup> and (b) the successful reconstructed examples in the study “Radial Computed Tomography of Air Contaminants Using Optical Remote Sensing”<sup>19</sup>.

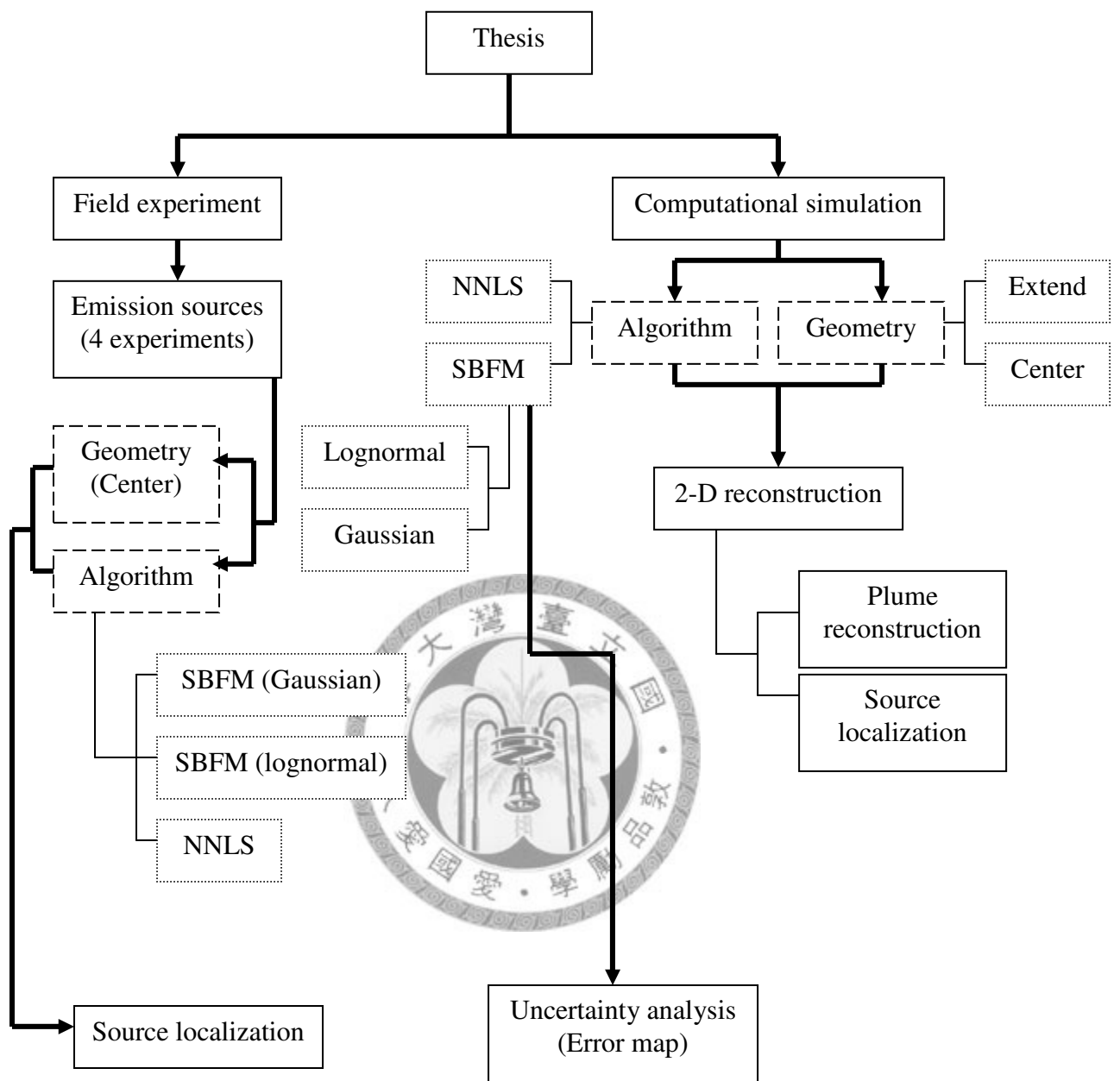



Figure 1.7 The flow chart of the thesis.

## Chapter2. Materials and Methods

This thesis is composed of two parts. The first part is the computational simulation. We proposed an “error map” concept to represent the uncertainty of current SBFM reconstruction. Also, the performance of different reconstruction algorithms and different beam geometries is evaluated. The second part is the field study. We release SF<sub>6</sub> and N<sub>2</sub>O as tracer gases and set up one of the beam geometry discuss in the computational simulation to validate the simulation results.

### 2.1 Data collection



The OP-FTIR used in the field experiment is manufactured by MASTEK Technology, Inc (MASTEK Technology, Inc, Wugu, Taiwan). It is a monostatic design with a mercury cadmium telluride (MCT) detector. The resolution is tunable and the highest resolution is 1 cm<sup>-1</sup> in wavenumber. The detection limit of this instrument is down to ppb level depending on the length of the monitoring line and the resolution of the spectrum. In order to direct the OP-FTIR telescope to scan multiple retroreflectors, a scanner composed of two step motors is used to rotate and elevate the OP-FTIR. The scanner is a custom built design which can direct the FTIR in 360° rotation and +60° to -30° elevation. The wind data (wind speed, wind direction) is collected by the meteorological station manufactured by R. M. Young (R.M. Young,

Traverse, Michigan, USA).

We set up a radial beam geometry with 9 retroreflectors (Retro) in a 30m×30m domain. Previous study<sup>18</sup> suggests that the segmenting information is needed when applying the SBFM reconstruction, thus the beam geometry must be composed of several sets of long and short beam paths. To do so, the experimental domain is divided into 3 ×3 grids (10m×10m for each grid). The 9 retroreflectors are located in the middle of these 9 squares (Figure 2.1 (a)) and are 170cm high from the ground.

The 9 retroreflectors are at the coordinate of (4.9,24.7) for Retro1, (4.8,14.9) for Retro2, (13.9,24.3) for Retro3, (11.5,16) for Retro4, (24.4,25) for Retro5, (5.1,4.5) for Retro6, (24.3,14.7) for Retro7, (14.5,7.5) for Retro8 and (25,4.9) for Retro9. In order to avoid the retro been arranged into a line, the retro in Pixel5 is shifted. The length of each beam path is 25.2m for path1, 15.6m for path2, 28.0m for Path3, 19.7m for Path4 34.9m for Path5, 6.8m for Paht6, 28.3m for Path7, 15.2m for Path8 and 25.5m for Path9. The FTIR is located at the origin of the experimental domain and mounted on the scanner previous mentioned which can direct the FTIR in a 360° rotation and +60° to -30° elevation. The FTIR scans the retroreflectors from 1<sup>st</sup> to 9<sup>th</sup> sequentially and repeatedly.

Pure nitrous oxide (N<sub>2</sub>O) and sulfur hexafluoride (SF<sub>6</sub>) are released from 3 plastic tubes which are vertical to the ground and at the same height of the retroreflectors

(Figure 2.2). The releasing flow rates for  $N_2O$  and  $SF_6$  are 10L/m and 0.5L/m respectively and are controlled by the mass flow controllers (Model 5850E series, Brooks, PA, USA). Four experiments are conducted with four different pairs of releasing locations (Figure 2.3 (a) to (d)). In each experiment, the two tracer gases are released simultaneously at different locations. This is to evaluate the ability of the RPM technique to localize two different emission sources under the same meteorological conditions. The Cartesian coordinate of  $SF_6$  and  $N_2O$  are (22.7,11.8) and (12.8,15.5) in Exp1, (7.4,9.9) and (11.6,22.7) in Exp 2, (13.6,6.1) and (4.7,14.9) in Exp 3 and (9.6,7.6) and (21.0,16.7) in Exp4. Before each experiment, we first begin the spectrum collection with the OP-FTIR to make sure that there is no background concentration for the two tracer gases in the experimental domain then we begin to release the tracer gases. Each experiment lasts for approximately two hours for the FTIR to scan the retroreflectors for approximately 25 complete sweeps. The concentration data which is collected after the tracer gas releasing rate is stable is used to reconstruct the source location. Thus for each experiment, approximately 20 sets of PIC are used for reconstruction. The spectra are collected with resolution of  $1\text{ cm}^{-1}$  in wavenumber. The time for each spectrum acquisition takes 20 seconds and the time for the FTIR move from one retroreflector to another is about 8 seconds. However it takes about 25 seconds for the FTIR to move from retro 9 to retro 1 due to the longer distance between these two retroreflectors. Thus it takes approximately 5 minutes for

a complete sweep.

Before the experiment, we follow the TO-16 method suggested by the EPA to implement the quality assurance procedure. The purpose of the quality assurance procedure suggested by the TO-16 is to determine how well the FTIR sensor is operating. For measuring the method noise, two spectra are collected with the same collection parameter (i.e. number of co-added scans, resolution). The absorbance spectrum is then created using either of these two spectra as background spectrum. The determination of method noise uses the statistical quantity called the root mean squared (RMS) deviation. The regions in the spectrum for the RMS calculation are 968–1008, 2480–2520, and 4380–4420  $\text{cm}^{-1}$ . Besides the method noise, the return beam intensity must remain proper during the spectrum collection. To make sure that the return beam intensity is not too strong, the region near the wavelength of 650 $\text{cm}^{-1}$  in the collected spectrum is examined to make sure that there is no indication of saturation. The detection limit is also considered in the TO-16. Following the Beer's law, the detection limit of the optical remote sensing instrument depends on the path length of the monitoring line. Thus to calculate the detection limit of the tracer gases, 16 spectra are collected continuously for each monitoring line. Using the previous collected spectrum as the background spectrum, 15 absorbance spectra are then created from the latter one(i.e. use the first spectrum as the background and create an absorbance spectrum from the second then use the second one as the background and create an



absorbance spectrum from the third one and so on). For these 15 absorbance spectra, the concentration of the SF<sub>6</sub> and N<sub>2</sub>O is calculated. The detection limit is 3 times the standard deviation of these 15 concentration values.

## 2.2 Data analysis

### 2.2.1 The computational simulation

In the computational simulation, a series of test distributions are generated using bivariate lognormal distribution<sup>38</sup> to mimic the skewed plume distribution when the wind presets. The SBFM and NNLS reconstruction are both used to reconstruct the test map under the beam geometry described in the data collection section and an example geometry described in the OTM-10 which is described in section 2.2.1.2. Furthermore, we use two different basis functions (bivariate lognormal distribution and bivariate Gaussian distribution) as the basis function when applying the SBFM reconstruction. The purpose of this computational simulation is (1) to compare the performance of the “grid based” reconstruction (i.e. NNLS) and the “non grid based” reconstruction (i.e. SBFM). (2) To evaluate whether using skewed distribution as basis function when applying SBFM reconstruction is able to better describe the plume. (3) To evaluate the performance of different beam geometries. (4) Evaluating the uncertainty of using the SBFM algorithm to localize the emission source by the concept of “error map” which is reported in the section 2.2.1.4: uncertainty analysis.

### 2.2.1.1 Generating test maps and derived PIC

The bivariate lognormal distribution is generated as the underlying distribution in a 30×30 domain. The probability density function of the bi-variate lognormal distribution is described by Aitchison and Brown in 1957 and has the following form<sup>38</sup>:

$$f(x_1, x_2) = \frac{1}{2\pi x_1 x_2 \sigma_{Y_1} \sigma_{Y_2} \sqrt{1 - \rho^2}} \exp\left(-\frac{q}{2}\right)$$

$$\text{Where } q = \frac{1}{1 - \rho^2} \left[ \left( \frac{\ln x_1 - \mu_{Y_1}}{\sigma_{Y_1}} \right)^2 - 2\rho \left( \frac{\ln x_1 - \mu_{Y_1}}{\sigma_{Y_1}} \right) \left( \frac{\ln x_2 - \mu_{Y_2}}{\sigma_{Y_2}} \right) + \left( \frac{\ln x_2 - \mu_{Y_2}}{\sigma_{Y_2}} \right)^2 \right] \quad (2-1)$$

$$x_1 > 0, x_2 > 0 \quad -1 < \rho < 1$$

where  $\mu_{Y_i}$  and  $\sigma_{Y_i}$  are the population mean and standard deviation of  $Y_i = \ln X_i$ ,  $i=1$  represents the X direction and  $i=2$  represents the Y direction.  $\rho$  is the population product-moment correlation coefficient of  $Y_1$  and  $Y_2$ . In this study, the lognormal distribution is multiplied by a scaling factor  $H$  which represents the peak height of the distribution. Thus the bivariate lognormal distribution used in this study has the following form:

$$f(x_1, x_2) = \frac{H}{2\pi x_1 x_2 \sigma_{Y_1} \sigma_{Y_2} \sqrt{1 - \rho^2}} \exp\left(-\frac{q}{2}\right)$$

$$\text{Where } q = \frac{1}{1-\rho^2} \left[ \left( \frac{\ln x_1 - \mu_{Y_1}}{\sigma_{Y_1}} \right)^2 - 2\rho \left( \frac{\ln x_1 - \mu_{Y_1}}{\sigma_{Y_1}} \right) \left( \frac{\ln x_2 - \mu_{Y_2}}{\sigma_{Y_2}} \right) + \left( \frac{\ln x_2 - \mu_{Y_2}}{\sigma_{Y_2}} \right)^2 \right] \quad (2-2)$$

$$x_1 > 0, x_2 > 0 \quad -1 < \rho < 1$$

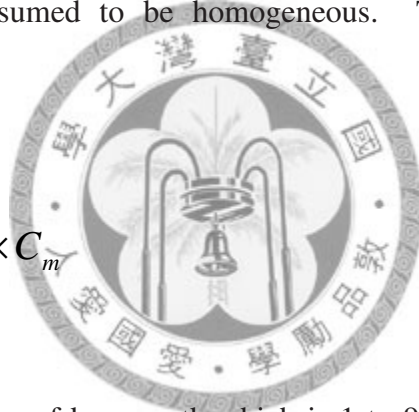
where  $x_1$  and  $x_2$  represent the location in Cartesian coordinate  $\mu_{Y_1}$  and  $\mu_{Y_2}$  represents the mean,  $\sigma_{Y_1}$  and  $\sigma_{Y_2}$  represents the standard deviation,  $\rho$  represents the correlation coefficient and  $H$  represents peak height. Note that the bivariate lognormal distribution can only be skewed to the direction of positive X axis and Y axis.

Two thousands of lognormal distributions are first generated by Monte Carlo method which is a process that relying on the repeated random sampling to generate results. A range of each parameters of the lognormal distributions are given, which are  $\ln(5)$  to  $\ln(45)$  with resolution of  $\ln(0.1)$  for  $\mu_{Y_1}$ ,  $\mu_{Y_2}$ , 0.1 to 1 with resolution of 0.05 for  $\sigma_{Y_1}$ ,  $\sigma_{Y_2}$ , -0.9 to 0.9 with resolution of 0.05 for  $\rho$  and 1 to 500 with resolution of 0.5 for  $H$ . The test maps are generated by randomly sample from the given range of these six parameters. That is, in each time of sampling, the six parameters are sampled from the given range of the parameters and are used to generate the first test map. Each time after the sampling, the sampled values are put back to the “parameter pool” for the next sampling process until a total of 2000 test maps are generated. As previously mentioned, the experimental domain is divided into 9 grids. Thus in each

grid, there are several test distributions with their peak locations in it. In order to make sure that the number of the test distributions is equal in each pixel, we randomly selected 50 test distributions in each pixel. Thus, a total of 450 (9 grids×50 test distributions) test distributions are randomly selected as the test map.

After the test distributions are generated, the observed PIC derived from these test distributions must be calculated for further reconstruction. To calculate the observed PIC, the experimental domain is first divided in to  $m$  pixels. In each pixel, the concentration value is assumed to be homogeneous. The observed PIC is then calculated as follow<sup>19</sup>:

$$PIC_{observed,i} = \sum_m K_{im} \times C_m \quad (2-3)$$



where  $i$  is the number index of beam path which is 1 to 9 (9 monitoring lines) in our study.  $m$  is the number index of pixel.  $K_{im}$  is the Kernel matrix representing the  $i^{th}$  beam path length in the  $m^{th}$  pixel.  $C_m$  is the concentration in the  $m^{th}$  pixel which is derived from the bivariate lognormal distribution (Equation 2-2). In this study, the concentration value in the middle of the  $m^{th}$  pixel is regarded to be the concentration in that pixel. Although the finer the pixel may yield more accurate PIC, however for the time consuming concern, we divided the domain into 60×60 pixels thus the size of each pixel is 0.5×0.5. The error of using this “grid integration” method was verified to

yield the error less than 2.5% (Appendix 2).

#### 2.2.1.2 The beam geometry

As mentioned previously, there are two kinds of beam geometries used in this study. The first one is the same described in the data collection section (Figure 2.2) and the second one is an example geometry described in the OTM-10. In the OTM-10, the “grid based” reconstruction algorithm, non-negative least square (NNLS), is used to reconstruct the “hot spot” of the emission. Thus, as previously mentioned, the experimental domain must first be divided into several smaller pixels. Once the number of the pixels is determined, the OTM-10 suggests that there should be at least one monitoring line that terminates within the boundary of each pixel to prevent the underdetermined situation. In the example geometry, the experimental domain is first divided into 3×3 pixels. Thus a total of 9 retroreflectors are located within the boundary of each pixel. To maximize the spread of the optical beams inside the area of emissions, one optical beam is set to pass through the middle of each pixel and terminates at the retroreflector on the boundary of each pixel. (Figure 2.1 (b)) The Cartesian coordinate of each retroreflector is (5.7,28.5) for Retro1, (5.6,17.1) for Retro2, (16.0,27.9) for Retro3, (8.1,8.4) for Retro4, (19.6,19.6) for Retro5, (28.8,27.6) for Retro6, (27.9,16.9) for Retro7, (16.7,5.1) for Retro8 and (28.8,5.6) for Retro9. The difference between the geometry used in the field experiment (named Geometry<sub>center</sub>)

(Figure 2.1 (a)) and that described in the OTM-10 (named Geometry<sub>extend</sub>) (Figure 2.1 (b)) is the location of each retroreflector. In the Geometry<sub>center</sub>, the retroreflector is located in the middle of each pixel while in the Geometry<sub>extend</sub>, the retroreflector is located near the boundary of each pixel.

### 2.2.1.3 The reconstruction algorithms

#### (1) Smooth basis function minimization (SBFM) algorithm

The SBFM reconstruction is first proposed by Drescher et al.<sup>30</sup> Different from the traditional “grid based” reconstruction algorithm, a known smooth basis function with unknown parameters must be first chosen to describe the plume of the pollutant when applying the SBFM algorithm. Since the concentration data obtained from the OP-FTIR is in the form of path integrated concentration (PIC), the PIC derived from the chosen basis function (called the predicted PIC) can be regarded as the integration of the basis function. Thus by fitting the predicted PIC to the measured PIC (observed PIC), the unknown parameters of the basis function can be estimated. To do so, an error function, the sum of squared errors (SSE), is conducted to be minimized.

$$SSE = \sum_i (PIC_{observed,i} - PIC_{predicted,i})^2 \quad (2-4)$$

where  $i$  is the ray number index,  $PIC_{observed,i}$  is the  $i^{th}$  observed PIC derived from the test distribution and  $PIC_{predicted,i}$  is the  $i^{th}$  predicted PIC derived from the chosen basis

functions (bivariate lognormal distribution and bivariate Gaussian distribution). By iteratively changing the parameters of the chosen basis function, the parameters that make the predicted PIC best fit to the observed PIC (minimum SSE) can be evaluated.

The solution searching process described above is called optimization which means finding the best solution for the given function. In one previous study, different optimizations algorithm including simplex algorithm and simulated annealing have been applied and evaluated in the RPM-SBFM technique.<sup>16</sup> In our study, we applied the built-in optimization function, `lsqnonlin`, in the optimization toolbox of MATLAB software to minimize the SSE. When using the `lsqnonlin` optimization, the upper and lower bound along with the first guess of the solutions must first be given to the algorithm. Once the observed PIC is input to the algorithm, the algorithm will begin to search the solution iteratively from the given first guess within the given upper and lower boundary. The searching procedure will stop when certain criteria are reached. The first criterion is that the solution perfectly yields a SSE of zero. The second one is that the change of the residual is less than the given tolerance. In this study, the tolerance is set to  $10^{-10}$  (default= $10^{-6}$ ). The third one is the number of iteration has reached to the given value which represents a possibility of reconstruction failure. The maximum number of iteration is set to 240000 (default=400) in this study.

To evaluate whether the reconstruction algorithm has reconstructed a reasonable

solution, the concordance correlation factor (CCF)<sup>39</sup> between the observed and predicted PIC is calculated (CCF<sub>PIC</sub>). The CCF is similar to the Pearson correlation coefficient however it is adjusted to account for the shifts in location and scale. The CCF values are limited between -1 and 1 yet it does not exceed Pearson correlation factor. The CCF equaling 1 represents a perfect match between the two sets of data. Since the fundamental procedure of the SBFM reconstruction is to minimize the SSE between the observed and predicted PIC, the CCF<sub>PIC</sub> is expected to near 1. Thus a low CCF<sub>PIC</sub> may indicate a failure of the reconstruction. To evaluate the performance of the reconstruction, the CCF between the test map and the reconstructed map is also calculated (CCF<sub>RPM</sub>). The CCF<sub>RPM</sub> equaling 1 represents a perfect match between the test map and the reconstructed result. Despite CCF<sub>RPM</sub>, another indicator, Nearness,<sup>23, 40</sup> is also calculated in this study to represent the performance of the reconstruction. The Nearness (refer to Nearness<sub>RPM</sub>) describes the discrepancy between the test map and the reconstructed map<sup>23, 40</sup>.

$$Nearness_{RPM} = \sqrt{\frac{\sum_m (c_m^* - c_m)^2}{\sum_m (c_m^* - c_{avg}^*)^2}} \quad (2-5)$$

where  $m$  is the number index of the pixel,  $c_m^*$  is the concentration value of the test map in the  $m^{th}$  pixel,  $c_m$  is the concentration value of the reconstructed map in the  $m^{th}$  pixel and  $c_{avg}^*$  is the average concentration of the test map. The smaller the Nearness<sub>RPM</sub>



represents the better the performance of the reconstruction thus the  $\text{Nearness}_{\text{SRPM}}$  equals 0 represents the perfect match between the test map and the reconstructed map.

The performance of the source localization is evaluated by two indicators. The first one is the distance between the real and reconstructed source location (peak error). The second one is to measure whether the reconstructed peak location “hit” the same pixel of the real source location. The number of the distribution out of 450 reconstructed distributions is calculated as the indicator “HIT”. Note that the pixel here refers to the pixel described in the generating test map section which is 3 by 3.

In the computational simulation, we apply both the bivariate Gaussian and bivariate lognormal distributions as the SBFM fitting functions.

(1.a) The bivariate lognormal distribution

The bivariate lognormal distribution is the same distribution used to generate the test distribution (Equation 2-2). Although the integration method described in section 2.2.1.3(1.b) may provide the actual PIC information, we are not able to apply it because there is no analytical solution when using bivariate lognormal distribution as basis function. Thus to calculate the predicted PIC from the bivariate normal distribution, the method, same as the one used to calculate the observed PIC in Equation 2-3, is applied. The experimental domain is first divided in to  $m$  pixels and the  $\text{PIC}_{\text{predicted},i}$  derived from the bivariate lognormal distribution is:

$$PIC_{predicted,i}(p_j) = \sum_m K_{im} \times L(X_m, Y_m, p_j) \quad (2-6)$$

where  $j$  (1 to 6) is the parameter number index,  $i$  (from 1 to 9) is the number index of beam path and  $m$  ( $3600=60 \times 60$ ) is the number index of pixel;  $p_j$  is the  $j^{th}$  parameter of the lognormal distribution;  $K_{im}$  is the Kernel matrix representing the path length for  $i^{th}$  beam path in the  $m^{th}$  pixel.  $L(X_m, Y_m, p_j)$  is the bivariate lognormal distribution in which  $X_m$  and  $Y_m$  represent the Cartesian coordinate in the middle of the  $m^{th}$  pixel.

The peak location, mode, of the bivariate lognormal distribution is associated with all of the distribution's parameters. To calculate the peak location, the concentration in each pixel derived from the reconstructed parameters is first calculated. The Cartesian coordinate in the middle of the pixel with the maximum concentration is regarded as the peak location. Since the size of each pixel is  $0.5 \times 0.5$ , the error between the actual and the calculated peak location will be less than 0.35 (i.e. the distance from the middle of the pixel to the corner of the pixel). In the scale of  $30m \times 30m$ , an error of 35cm is acceptable.

#### (1.b) The bivariate Gaussian distributions

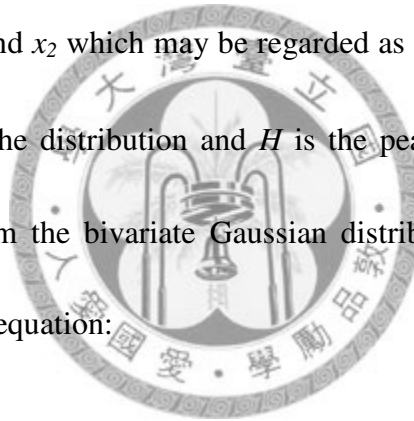
The bivariate Gaussian distribution, similar to the lognormal distribution, is also described by six parameters: the peak location, two standard deviations, the peak height, and the angle between the x-axis and the principal axis of the distribution and has the

following form:

$$G(x_1, x_2) = \frac{H}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{z}{2(1-\rho^2)}\right]$$

$$z = \frac{(x_1 - \mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \quad (2-7)$$

where  $x_1$  and  $x_2$  are the Cartesian coordinate of the location in the experimental domain,  $\mu_1, \mu_2$  and  $\sigma_1, \sigma_2$  are the peak location and the standard deviation for both X and Y axis,  $\rho$  is the correlation of  $x_1$  and  $x_2$  which may be regarded as the angle between the x-axis and the principal axis of the distribution and  $H$  is the peak height. To calculate the predicted PIC derived from the bivariate Gaussian distribution, Equation 2-7 is first reformed by the following equation:



$$\begin{aligned} x_1 &= r \cos \theta \\ x_2 &= r \sin \theta \end{aligned} \quad (2-8)$$

Thus the bivariate Gaussian is reformed and presented in polar coordinate:

$$G(r, \theta) = \frac{H}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{z}{2(1-\rho^2)}\right]$$

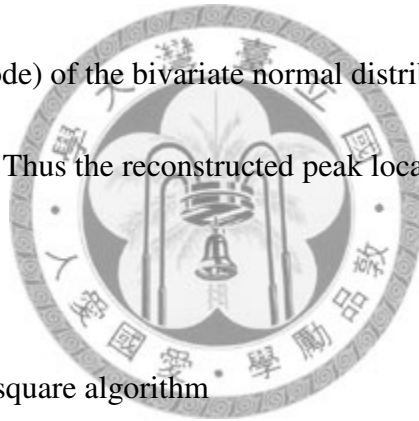
$$z = \frac{(r \cos \theta - \mu_1)^2}{\sigma_1^2} - \frac{2\rho(r \cos \theta - \mu_1)(r \sin \theta - \mu_2)}{\sigma_1\sigma_2} + \frac{(r \sin \theta - \mu_2)^2}{\sigma_2^2} \quad (2-9)$$

and the  $PIC_{predicted,i}$  for the bivariate Gaussian distribution is<sup>16</sup>:

$$PIC_{predicted,i}(p_j) = \sum \int_0^{L_i} G(r, \theta_i, p_j) dr \quad (2-10)$$

where  $j$  is the parameter number index (1 to 6) and  $i$  (1 to 9) is the number index of beam path;  $L_i$  is the  $i^{th}$  beam path length and  $p_j$  is the  $j^{th}$  parameter of the bivariate Gaussian distribution;  $G(r, \theta_i, p_j)$  is the bivariate Gaussian distribution in polar coordinates  $r$  and  $\theta$ .

The peak location (mode) of the bivariate normal distribution is the same as the mean of this distribution. Thus the reconstructed peak location is  $(\mu_1, \mu_2)$  in Cartesian coordinate.

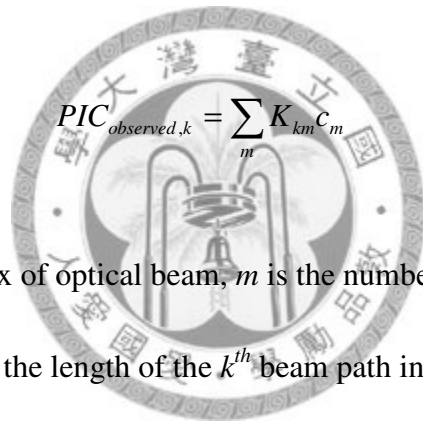


(2) The non-negative least square algorithm

The non-negative least square algorithm is described in the US EPA other test method 10 (OTM10)<sup>35</sup> which is an iterative algebraic deconvolution algorithm<sup>41</sup>. The NNLS algorithm is also a built-in algorithm for solving nonnegative least-squares constraints problem. For example, if  $K$  is a  $k$ -by- $m$  matrix and  $D$  is a vertical vector with  $k$  elements. The NNLS algorithm is used to solve the vertical vector  $C$  where  $K_{km} \times C = D_k$ .

Similar to the traditional “grid based” algorithm, when using the NNLS algorithm

the experimental domain must first be divided into several pixels. For each pixel, the concentration is assumed to be non-negative and homogeneous. By implementing the reconstruction algorithm using collected PIC data, the concentration value of each pixel can be estimated. In an example of OTM-10, the experimental domain is divided into 9 pixels. And to prevent the underdetermined situation, there must be more than 1 monitoring line terminates within the boundary of each pixel. Once the beam geometry is determined, the measured PIC (observed PIC), as a function of the field of concentration, is given by:

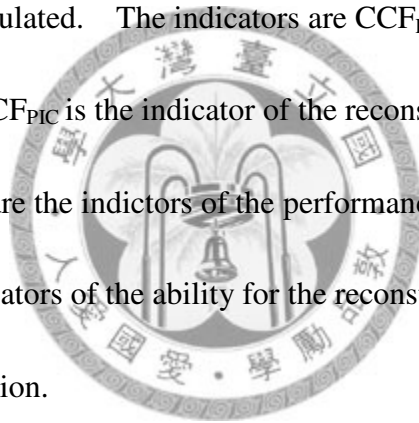
$$PIC_{observed,k} = \sum_m K_{km} c_m \quad (2-11)$$


where  $k$  is the number index of optical beam,  $m$  is the number index of pixel,  $K_{km}$  is the Kernel matrix representing the length of the  $k^{th}$  beam path in the  $m^{th}$  pixel and  $c_m$  is the average concentration in the  $m^{th}$  pixel. In our study, the  $m=1$  to 9 (9 pixels) and  $k=1$  to 9 (9 optical beams). Thus the  $PIC_{observed}$  for each beam path derived from the test distribution ( $PIC_{observed,k}$ ) and the Kernel matrix ( $K_{km}$ ) are input to the NNLS algorithm and the mean concentration of each pixel ( $c_m$ ) can be estimated as output.

The second stage of the NNLS reconstruction involves the interpolation among the reconstructed pixel's mean concentration, providing a peak location not limited in the middle of each pixel. The interpolation method used in this study is Kriging method which is a geostatistical technique to interpolate or extrapolate the value of a random

field. The Krigin process is implemented by the free software “DACE”<sup>42</sup> which is a MATLAB toolbox. For the comparison purpose, we interpolate and extrapolate the mean concentration value of each pixel to the same scale of the test map which is 60×60 in the emission area. Thus the peak location of the reconstructed plume is assumed to be the location in the middle of the interpolated pixels with highest concentration in Cartesian coordinate.

The indicators used in the SBFM reconstruction to evaluate the performance of the reconstruction are also calculated. The indicators are  $CCF_{PIC}$ ,  $CCF_{RPM}$ ,  $Nearness_{RPM}$ ,  $DIST$  and  $HIT$ . Where  $CCF_{PIC}$  is the indicator of the reconstruction’s reasonableness,  $CCF_{RPM}$  and  $Nearness_{RPM}$  are the indicators of the performance of the reconstruction and  $DIST$  and  $HIT$  are the indicators of the ability for the reconstruction algorithm to reconstruct the source location.



#### 2.2.1.4 The uncertainty analysis of SBFM algorithm

We conduct an uncertainty analysis to construct the error map of the beam geometry when applying the SBFM reconstruction. The two geometries  $Geometry_{center}$  and  $Geometry_{extend}$  are both investigated in this uncertainty analysis. The main concept of this uncertainty analysis is to mimic the SBFM reconstructing process yet without implementing the optimization algorithm. As mentioned in the SBFM algorithm, the error function, SSE between the observed and predicted PIC is the

fitting function for the algorithm to find the solution. By minimizing the SSE, the parameters of the chosen basis function can be estimated. To mimic this solution searching process, we deploy the bivariate Gaussian distribution as basis function. A range for each of the bivariate Gaussian distribution's parameter is then introduced as the possible solution. The procedure of generating possible solution and their derived PIC is as follow:

Step 1: Give a range of the parameters of bivariate distributions (six parameters) which are: 0 to 30 with resolution of 2 for peak location  $\mu_1$  and  $\mu_2$ , 1 to 6 with resolution of 1 for standard deviation  $\sigma_1$  and  $\sigma_2$ , 10 to 500 with resolution of 10 for peak height  $H$  and -0.9 to 0.9 with resolution of 0.3 for the angle  $\rho$ .

Step2: Calculate the derived PIC set (9 PICs for each combination) from the combination of the parameters in step 1 with Equation 2-10. Thus for each combination of the  $\mu_1$  and  $\mu_2$  ( $256=16\times 16$ ), there are a series of distributions ( $12600=6\times 6\times 50\times 7$ ) regarded as the predicted distributions with different shape ( $\sigma_1$  and  $\sigma_2$ ), peak height ( $H$ ) and different angle ( $\rho$ ).

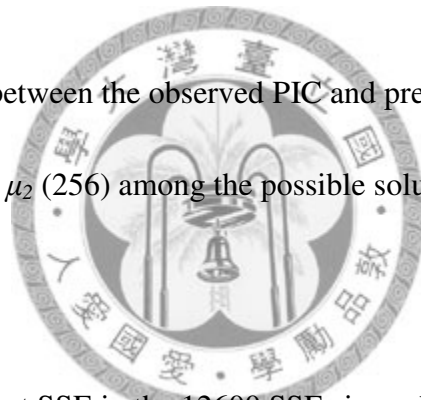
The next step of the uncertainty analysis is to calculate the SSE between the predicted and observed PIC. There are two kinds of test distributions used in this uncertainty analysis which are the single and multiple test distribution. The procedure for generating single test distribution is as follow:

Step 1: Divide the experimental domain into 3 by 3 pixels, generate a bivariate Gaussian distribution with the peak location in the middle of Pixel 1.

Step 2: The PIC set (9 PICs) derived from the test distribution is calculated with Equation 2-10 as observed PIC.

The error map of the single test distribution is able to visually demonstrate the possible reconstructed source location when a test distribution presents. The error map is generated as follow (Figure 2.4):

Step 3: Calculate the SSE between the observed PIC and predicted PIC. For each combination of  $\mu_1$  and  $\mu_2$  (256) among the possible solutions, 12600 SSEs are calculated.



Step 4: Calculate the smallest SSE in the 12600 SSEs in each combination of  $\mu_1$  and  $\mu_2$ . For example, a total of 12600 SSEs are calculated as the candidate SSE at (0,0) in Cartesian coordinate (the first combination of  $\mu_1$  and  $\mu_2$ ). The combination of the parameters that yields the minimum SSE among these 12600 SSEs are regarded as the best answer the algorithm can find at (0,0).

Step 5: For each of the 16×16 combinations of  $\mu_1$  and  $\mu_2$  (source location), there is a minimum SSE representing the smallest SSE the SBFM algorithm can find at this location.



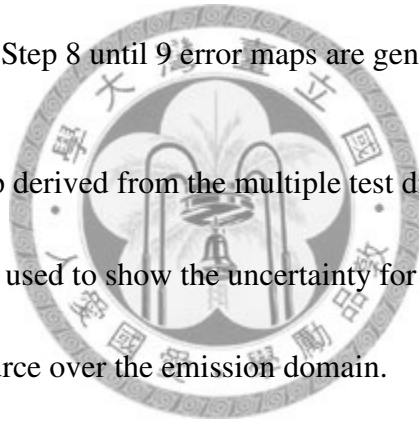
Step 6: Standardize these 16×16 SSEs by dividing them by the maximum value among them. The location with smaller standardized SSE represents the possible reconstructed source location.

Step 7: The contour plot of this SSE matrix is called error map. The area with SSE less than  $10^{-2}$  in the error map represents the possible reconstructed source location.

Step 8: Generate the same test distribution with peak location in the middle of Pixel 2.

Repeat from Step 1 to Step 8 until 9 error maps are generated.

The other kind of error map derived from the multiple test distribution is called the overall error map. This is used to show the uncertainty for the current beam geometry to localize the emission source over the emission domain. The procedure to generate the observed PIC derived from the multiple test distribution is as follow:



Step 1: Generate a series of the same test distributions in the aspect of peak height, width and the size with peak locations shifting from (0,0) to (30,30) with resolution equaling 2. Thus total of 256 test distributions are generated as observed PIC.

Step 2: Calculate the derived PIC sets (9 for each) of the observed distributions with Equation 2-10 as observed PIC.

The procedure of generating the overall error map is as follow (Figure 2.5):

Step 1: Calculate the SSE between the observed PIC with peak location at (0,0) and predicted PIC. The predicted PIC is the same as described in generating the error map of the single test distribution.

Step 2: For each combination of the  $\mu_1$  and  $\mu_2$  ( $256=16\times 16$ ) of the predicted distributions, 12600 SSEs are calculated. Use the smallest SSE among these 12600 SSEs as the best answer the SBFM algorithm can find when the predicted distribution is at this location.

Step 3: For each combination of the  $\mu_1$  and  $\mu_2$ , there is one SSE representing the best answer the SBFM algorithm can find. Standardize the SSE matrix by dividing them by the maximum value among them.

Step 4: Calculate the area with standardized SSE smaller than  $10^{-2}$  which represents the possible area where the reconstructed source location might appear when the real source location is at (0,0).

Step 5: Repeat Step 1 to Step 4 however the observed PIC with peak location at (0,2) is used as test distribution. Repeat the process until peak location of the observed PIC reaches (30,30).

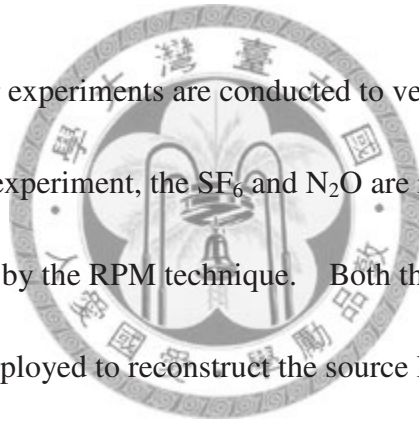
Step 6: For each test distribution, there is a value representing the area that the reconstructed source location might appear. The larger the value represents the

larger uncertainty for the SBFM to reconstruct the emission source at this location.

In order to further investigate the effect of the different plume (i.e. the size and the shape of the plume), we use the test distribution with ratio of  $\sigma_X$  and  $\sigma_Y$  being 1:1, 1:2, 1:3, 2:1 and 3:1 respectively. This is to generate the test distribution with different shape. And in each ratio of the  $\sigma_X$  and  $\sigma_Y$ , there are at least two distributions with different size.

### 2.2.2 The field experiment

In the field study, four experiments are conducted to verify the performance of RPM technique. In each experiment, the SF<sub>6</sub> and N<sub>2</sub>O are released simultaneously as tracer gases to be localized by the RPM technique. Both the reconstruction algorithms of NNLS and SBFM are deployed to reconstruct the source locations.



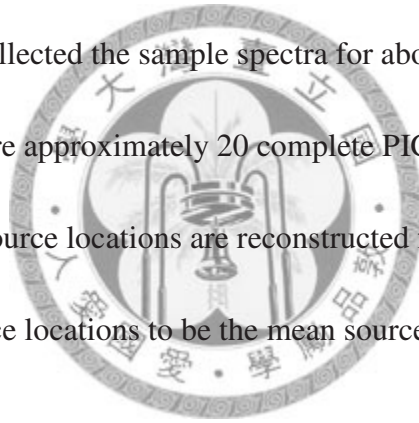
#### 2.2.2.1 The spectrum quantification

For the spectrum quantification, the classical least square (CLS) method is applied by the Ir-soft software (ITRI, Hsingchu, Taiwan). Each set of observed PIC is consisted with 9 PIC data for which the path length are 25.2m, 15.6m, 28.0m, 19.7m, 34.9m, 6.8m, 28.3m, 15.2m and 25.5m respectively. For the N<sub>2</sub>O, the spectral range of 2120-2228 cm<sup>-1</sup> is used for the quantification with the reference spectrum at the PIC of 142 ppm-m. Since the IR absorbance of this region is overlapped with CO, the

reference spectrum of CO (462.4 ppm-m) is also included for the N<sub>2</sub>O quantification to eliminate the influence of the substance. As for SF<sub>6</sub>, the spectral region from 935 to 955 cm<sup>-1</sup> is used for quantification. The reference spectra are with PIC of 10.4 ppm-m.

#### 2.2.2.2 The source location reconstruction

After the quantification of the PIC data, the observed PIC (i.e. measured PIC) is then input to both the NNLS and SBFM reconstruction. A complete set of PIC (i.e. PIC collected from Retro1 to Retro9) may yield one reconstructed source location. In the field experiment, we collected the sample spectra for about 2 hours in each experiment. Thus there are approximately 20 complete PIC sets in each experiment. And as a result, about 20 source locations are reconstructed for each tracer gas. We then average these 20 source locations to be the mean source location.



The SBFM reconstruction is implemented as described in the computational simulation. The difference is that the observed PIC is no longer calculated from the test distribution but is from the direct measurement of the tracer gases. The bivariate lognormal and bivariate Gaussian distributions are both deployed as the basis function. The maximum iteration time is set to 240000 and the tolerance is set to 10<sup>-10</sup> as described in the computational simulation. The time for a single reconstruction takes about 80 seconds when using the bivariate lognormal distribution and 2.5 seconds when using the bivariate Gaussian distribution. The difference in reconstruction time is

because that we use the analytical method to generate the PIC when using the bivariate Gaussian distribution while the “grid integrated” method is applied when using the bivariate lognormal distribution as basis function. The distance between the real and reconstructed source location along with “HIT” (whether the reconstructed source location is at the correct pixel) are calculated as indicators of the ability for the source localization.

The NNLS reconstruction is also implemented for source localization. The experimental domain is first divided into 9 pixels. The reconstructed concentration is then projected in these 9 grids. The Kriging interpolation is then applied to interpolate these 9 reconstructed concentrations into 60×60 pixels. As described in the OTM-10, the peak location after the kriging interpolation will not be limited in the middle of the 9 pixels. Same as the SBFM reconstruction, the distance between the reconstructed and real source location along with “HIT” are calculated as the indicators of the reconstruction performance.

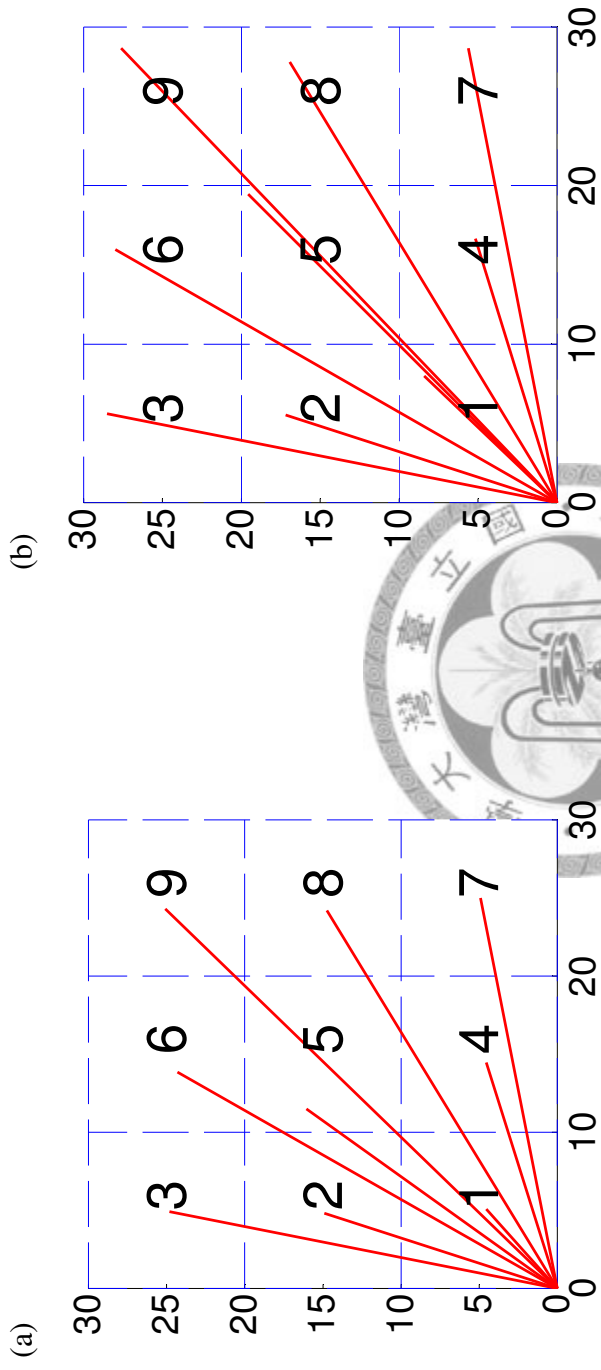


Figure 2.1 The beam geometries used in the computational simulation (a) Geometry<sub>center</sub>. (b) Geometry<sub>extend</sub>. The number in each pixel represents the number index of each pixel



Figure 2.2 The tracer gas releasing device. The three vertical plastic tubes are at the same height of the retroreflectors.



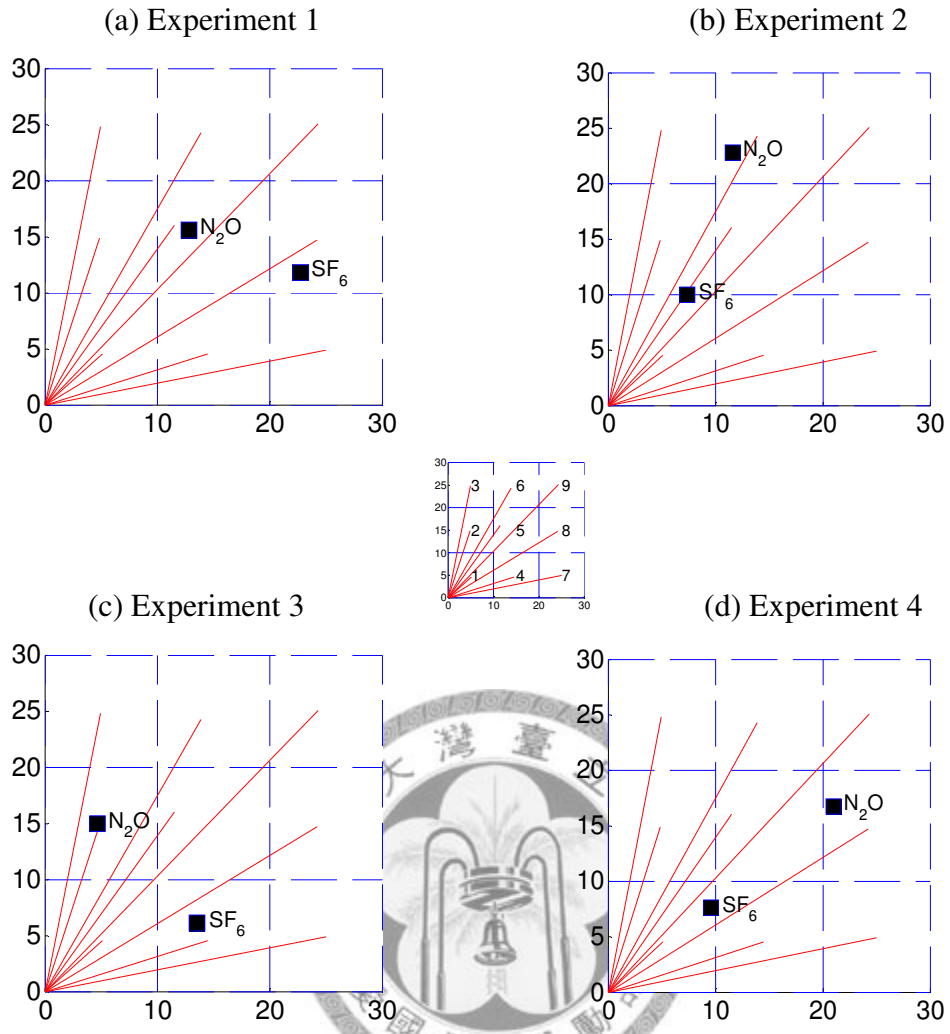


Figure 2.3 The experimental setup of the field experiment (a) Experiment 1 (b) Experiment 2 (c) Experiment 3 (d) Experiment 4. The black square represents the source location



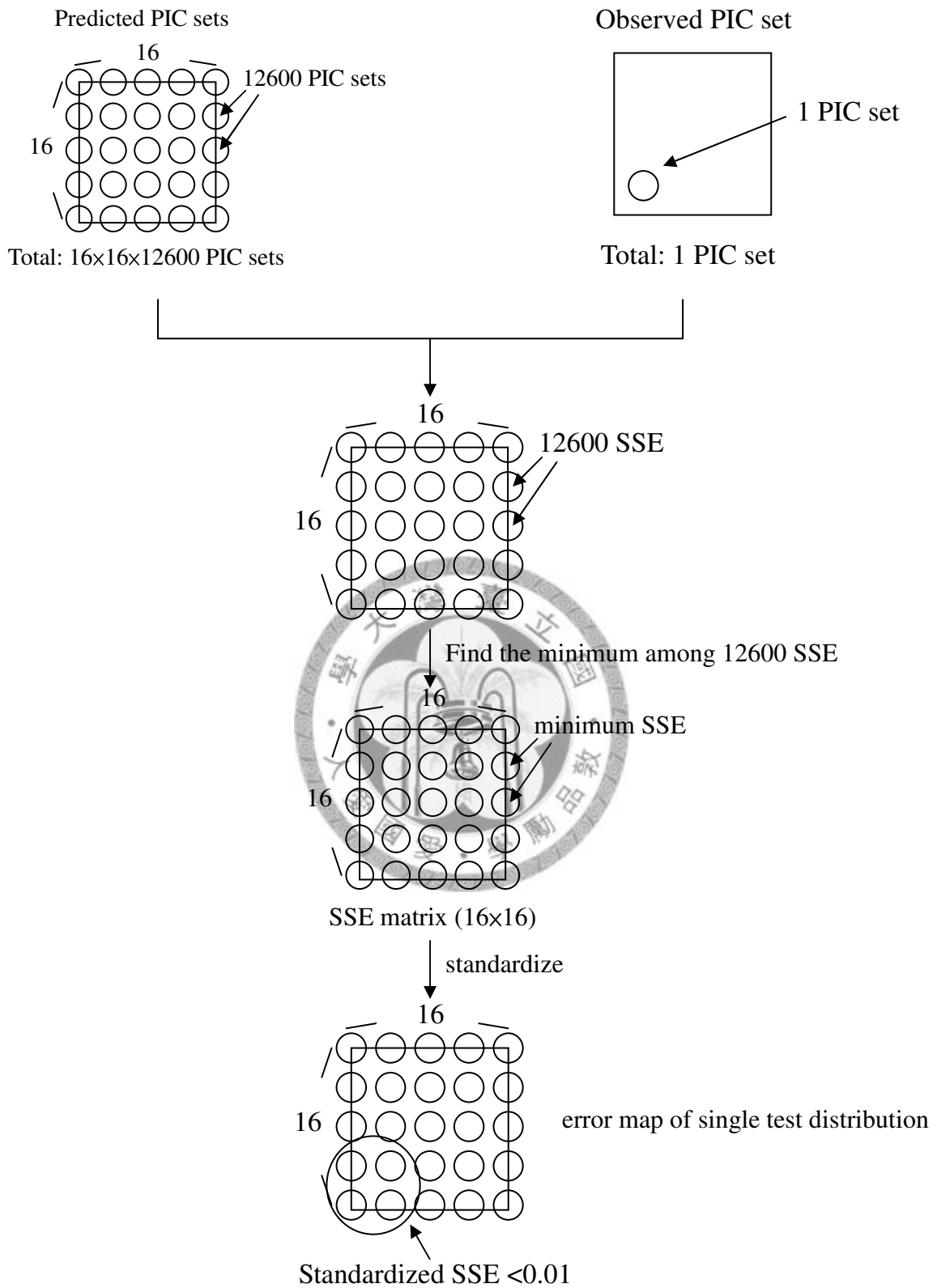
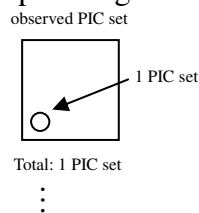


Figure 2.4 The procedure of generating error map of single test distribution



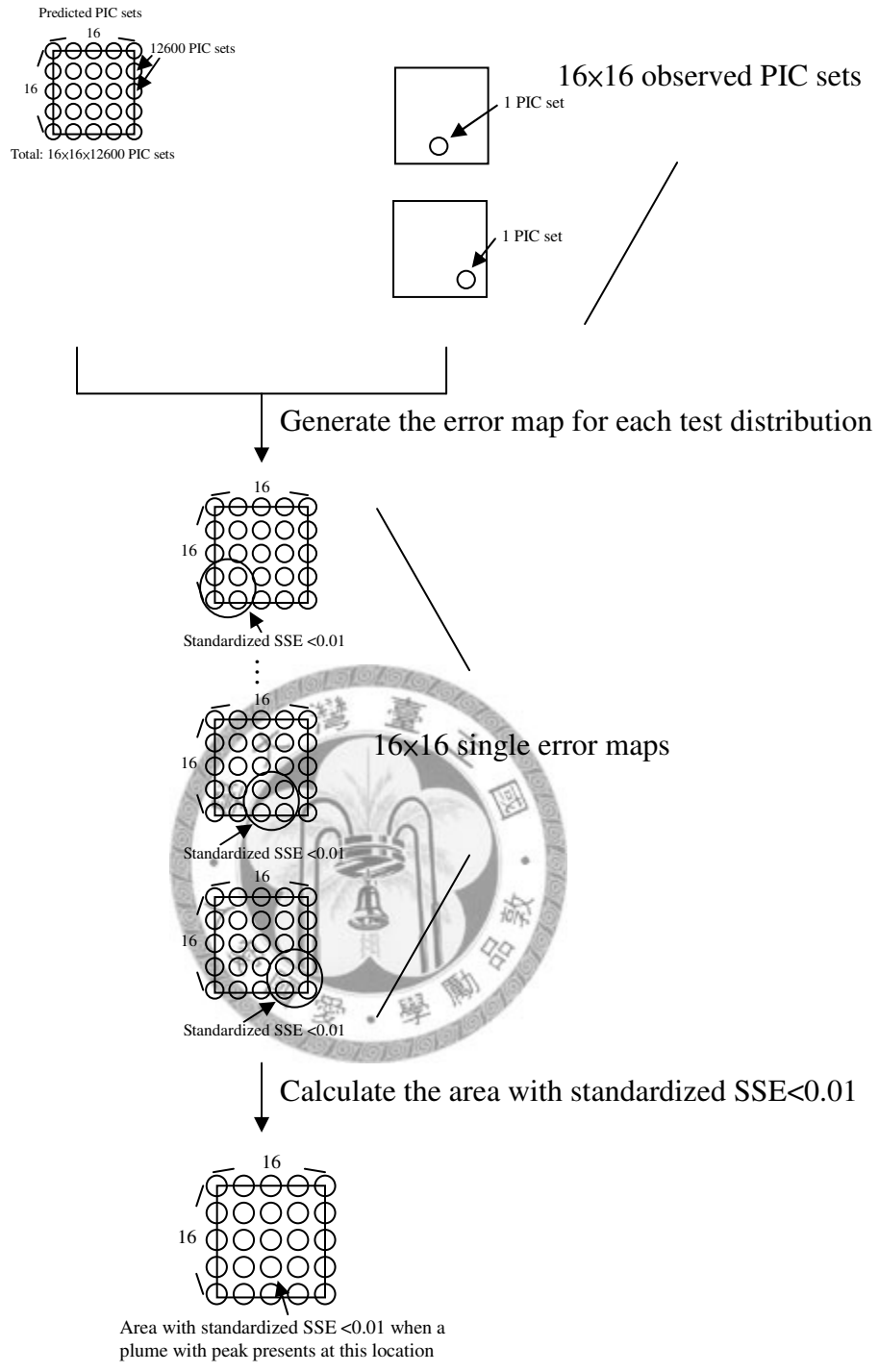


Figure 2.5 The procedure of generating overall error map

## Chapter 3. Results and discussion

### 3.1 Computational simulation results

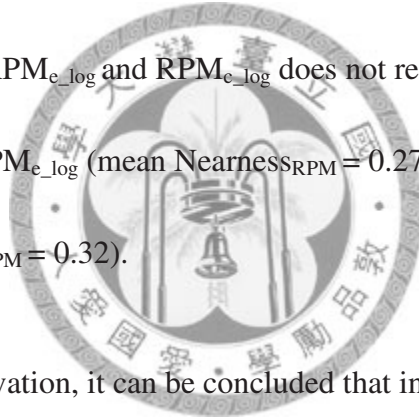
#### 3.1.1 The plume reconstruction

Figure 3.1 shows one example of the reconstruction results using the reconstruction methods described above under the same test map. The test map in this example is from pixel 5, which is in the middle of the experimental domain and is a wide plume. The best result is given by both the  $RPM_{c\_log}$  and  $RPM_{e\_log}$  (Figure 3.1 (d) (e)). The four indicators show that the reconstructed distributions by these two methods have perfectly matched the test map in both the aspects of reconstruction quality and source localization ability. The worst result is given by the  $RPM_{c\_Gauss}$  (Figure 3.1 (b)), although it yields a high  $CCF_{PIC}$  ( $CCF_{PIC}=0.92$ ), it fails to reconstruct the plume ( $CCF_{RPM}=0.12$   $Nearness_{RPM}=2.30$ ) but give a comparable source location estimation to that by the  $RPM_{c\_NNLS}$  and  $RPM_{e\_NNLS}$  (peak error=4.55, hit the right pixel). As for the two reconstructions made by the NNLS algorithm under the two different beam geometries, although the peak errors (5.38 for  $Geometry_{center}$  and 5.40 for  $Geometry_{extend}$ ) are slightly worse than that by the  $RPM_{c\_Gauss}$ , they both fail to reconstruct the source location that is at the right pixel. In general, for the reconstruction of the plume, the SBFM using bivariate lognormal as basis function is

able to give the best result while the worst result is given by the SBFM using bivariate Gaussian as basis function. As for the reconstruction of the source location, although the SBFM using bivariate Gaussian gives the worst result, it can still reconstruct a comparable or even better result to that of NNLS algorithm.

The summary statistics of the reconstruction result is shown in Table 3.1. The first two rows indicate the reconstruction using SBFM algorithm with bivariate Gaussian distribution as basis function under the extend beam geometry ( $RPM_{e\_Gauss}$ ) and the center beam geometry ( $RPM_{c\_Gauss}$ ) respectively. The 3<sup>rd</sup> and 4<sup>th</sup> rows indicate the reconstruction using SBFM algorithm with bivariate lognormal distribution as the basis function under extend ( $RPM_{e\_log}$ ) and center beam geometry respectively ( $RPM_{c\_log}$ ). The last two rows indicate the reconstruction using NNLS algorithm also under extend ( $RPM_{e\_NNLS}$ ) and center beam geometry ( $RPM_{c\_NNLS}$ ) respectively. From the column of ray fit, the  $CCF_{PIC}$  in all of the reconstruction method is higher than 0.90 which implies that the result has fulfilled the basic criteria of reconstruction that is fitting the predicted PIC to the observed PIC. For the two indicators of the reconstruction performance,  $CCF_{RPM}$  (the larger represents the better quality) ranges from 0.27 to 0.92 and  $Nearness_{RPM}$  (the smaller represents the better quality) ranges from 0.27 to 17.23. The best result is from the  $RPM_{e\_log}$  ( $CCF_{RPM} = 0.92$  and  $Nearness_{RPM}=0.27$ ) while the worst reconstruction result is from that by  $RPM_{c\_Gauss}$  ( $CCF_{RPM} = 0.27$  and  $Nearness_{RPM}=17.23$ ). Despite the different definition of the two

indicators, the Nearness and  $CCF_{RPM}$  shows conformity between each other. In general, under the same beam geometry, the SBFM reconstruction using lognormal distribution as basis function gives the best result followed by NNLS and SBFM using bivariate Gaussian distribution as basis function. As for the beam geometry, under the same reconstruction algorithm, the  $Geometry_{extend}$  gives the better result than that given by  $Geometry_{center}$ . The difference of the  $CCF_{RPM}$  between either pair of the reconstruction method has all reached statistical significant level at  $\alpha=0.05$  under paired t-test. Similar result is observed in  $Nearness_{RPM}$ , although the difference of the  $Nearness_{RPM}$  between the  $RPM_{e\_log}$  and  $RPM_{c\_log}$  does not reach the significant level, it can still be seen that the  $RPM_{e\_log}$  (mean  $Nearness_{RPM}=0.27$ ) performs better than the  $RPM_{c\_log}$  (mean  $Nearness_{RPM}=0.32$ ).



From the above observation, it can be concluded that in the aspect of reconstruction quality, the geometry from EPA ( $Geometry_{extend}$ ) performs better than the geometry from our original design ( $Geometry_{center}$ ) using the same reconstruction algorithm. Since the only difference between these two geometries is the ray length, the result shows that the longer the ray length yields better performance. We will discuss about the possible reasons in the latter paragraph. The other finding is that the SBFM algorithm may perform better than the NNLS reconstruction by choosing the suitable basis function. In the computational simulation, under the same beam geometry, the ranking of the reconstruction performance is (1) SBFM using bivariate

lognormal distribution as basis function (2) NNLS (3) SBFM using bivariate Gaussian as basis function (Table 3.1). The result is reasonable because the test distribution used in the simulation is the bivariate lognormal distribution which is a skewed distribution. Thus it can be predicted that a symmetric distribution may not fit well and as a consequence, the SBFM using bivariate Gaussian distribution gives the worst result. The NNLS algorithm is able to reconstruct better because it is a “grid based” reconstruction technique. Instead of superimposing a basis function, the NNLS directly estimates the concentration in the experimental domain thus it is not affected by the plume’s distribution. As for the best result reconstructed by the SBFM using bivariate lognormal distribution as basis function, this may be because that the same basis function is used to generate the test map thus the reconstructed plume is able to perfectly fit the test distribution (mean  $CCF_{PIC}=1$ ) and gives a nearly perfect result (mean  $CCFRPM>0.89$ ). Since the best result is given by the SBFM reconstruction using the same basis distribution as the test map, it can only be concluded that using SBFM with bivariate lognormal distribution as basis function is better when the plume is skewed to the same direction as the bivariate lognormal distribution.

For the source localization accuracy, despite the high  $CCF_{PIC}$  for each reconstruction methods, the mean peak error varies from 1.53 to 8.02 (Table 3.1). The best result is given by  $RPM_{e\_log}$  (mean peak error=1.53) while the worst result is given by  $RPM_{c\_Gauss}$  (mean peak error=8.02). Table 3.2 shows the paired t-test result of the

peak location between each reconstruction method. The difference between  $RPM_{e\_Gauss}$  and  $RPM_{c\_NNLS}$  does not reach the statistical significant level at  $\alpha=0.05$ . As for the other indicator used to describe the accuracy of the source localization ability, HIT, the best result is also given by  $RPM_{e\_log}$  (HIT=406 (90.2%)) while the worst is also given by  $RPM_{c\_Gauss}$  (140(31.1%)) (Table 3.1).

From the above observation, the reconstruction method using SBFM with bivariate lognormal distribution gives superior result in the aspect of the source localization accuracy under the same beam geometry. However, there is no significant difference between the NNLS reconstruction and the SBFM using the bivariate Gaussian as basis function which implies that using SBFM with bivariate Gaussian as basis function may still reconstruct a comparable result comparing to the NNLS reconstruction. The opposite result given by  $RPM_{e\_NNLS}$  and  $RPM_{e\_Gauss}$  in the two indicators of source localization accuracy seems to be illogical. It is still reasonable because although the reconstructed source location does not “hit” the correct pixel, it may still close to the real source location at nearby pixel. Comparing to the reconstruction quality, although the NNLS reconstruction performs better than the SBFM using bivariate Gaussian as the basis function, the SBFM using bivariate Gaussian as basis function can still reconstruct a comparable source location to the NNLS reconstruction. This indicates that although failing to reconstruct the whole plume, the SBFM with bivariate Gaussian as basis function can still reconstruct the source location. In the aspect of finding the

emission source, SBFM using a symmetric basis function is able to give a comparable result to the grid based method. In the aspect of plume reconstruction, SBFM using a symmetric basis function might lead to an estimation bias. Again, although the SBFM using bivariate lognormal as basis function gives the best result in both indicators, it can not be concluded that this is the best reconstruction algorithm for source localization because the test distribution is exactly the same as the basis function. It can only be concluded that the SBFM using bivariate lognormal as basis function is better when the plume is skewed to the same direction as the bivariate lognormal distribution in the real world.

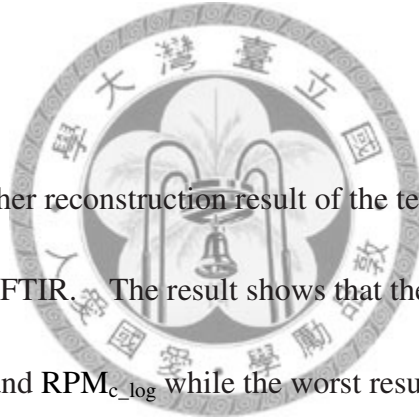


Figure 3.2 shows another reconstruction result of the test map in pixel 1, which is close to the location of OP-FTIR. The result shows that the best performance is implemented by  $RPM_{e\_log}$  and  $RPM_{c\_log}$  while the worst result is given by  $RPM_{e\_NNLS}$  and  $RPM_{c\_NNLS}$  in both the aspects of reconstruction quality and source localization accuracy. This result is somehow different from the overall result suggested in Table 3.1. Another example is shown in Figure 3.3, where the test map is in pixel 9 and is far away from the OP-FTIR. The result shows that the best performance is implemented by  $RPM_{e\_log}$  and  $RPM_{c\_log}$  while the worst is given by  $RPM_{e\_Gauss}$ . The  $RPM_{e\_NNLS}$  and  $RPM_{c\_NNLS}$  both give comparable results to that given by the  $RPM_{e\_log}$  and  $RPM_{c\_log}$  ( $CCF_{RPM}=0.97$  and  $0.91$ ,  $Nearness_{RPM}=0.25$  and  $0.46$ , peak error=4 and 4.6 and hit the correct pixel) while the SBFM using Gaussian distribution not only fail



to reconstruct the plume but also failed to localize the source location ( $CCF_{RPM}=0.002$  and  $0.003$ ,  $Nearness_{RPM}=21.0$  and  $30.9$ , peak error= $18.58$  and  $17.36$ )

The above result seems to suggest that there is an association between the reconstruction quality and the real source location. Table 3.3 and Table 3.4 respectively, have shown the summary statistics of the reconstruction quality and source localization accuracy of the reconstruction of  $RPM_{c\_Gauss}$  and  $RPM_{c\_NNLS}$  classified by the pixel index. From Table 3.3, the mean  $CCF_{PIC}$  ranges from  $0.89$  to  $0.99$  which suggests that the reconstruction has reached the criteria of fitting the predicted PIC to observed PIC. However, the mean  $CCF_{RPM}$  and mean  $Nearness_{RPM}$  ranges from  $0.05$  to  $0.57$  and  $44.95$  to  $1.31$  respectively. In terms of the SBFM reconstruction using bivariate Gaussian distribution as basis function, the best result is in pixel 1 where the mean  $CCF_{RPM}$  and mean  $Nearness_{RPM}=0.57$  and  $1.31$  however the  $CCF_{RPM}$  drops dramatically to  $0.05$  in pixel 9 and the  $Nearness_{RPM}$  rises to  $44.95$  in pixel 9 which indicates that the SBFM using bivariate Gaussian distribution is not able to reconstruct the plume which is far away from the OP-FTIR. Despite the reconstruction quality, the source localization accuracy has the almost same trend. The mean peak error rises from  $3.11$  to  $15.54$  from pixel 1 to pixel 9 and no reconstructed source location has hit the correct pixel in pixel 9.

A completely different result is reconstructed by the NNLS algorithm. From

table 3.4, the test maps in pixel 9 can be best reconstructed (mean  $CCF_{RPM}=0.80$ ,  $Nearness_{RPM}=0.55$ , mean peak error=4.58) by the NNLS algorithm while the test maps in pixel 1 is not able to be reconstructed well (mean  $CCF_{RPM}=0.22$ ,  $Nearness_{RPM}=1.39$ , mean peak error=15.35) by the NNLS algorithm. The result suggests that the NNLS reconstruction is not able to reconstruct the plume which is near the OP-FTIR however it is able to reconstruct well when the plume is far away from the OP-FTIR. This may be because that although the source location is near the FTIR, there still might be a few monitoring lines (i.e. the monitoring lines on the edge) of the pixel that has detected the low PIC. All of the monitoring lines have passed Pixel 1, thus the high concentration value in Pixel 1 would rise the reconstructed PIC when deploying the NNLS reconstruction. As a result, the NNLS algorithm can only subject a low concentration data to Pixel 1 to satisfy the monitoring line that has detected less emission to yield low reconstructed PIC. However when the source is far away from the OP-FTIR, the concentration in each pixel is correlated to certain monitoring line only. Thus the NNLS is able to subject a high concentration value in that pixel.

### *3.1.2 The reliability of source localization*

From section 3.1.1, it can be seen that the SBFM reconstruction using the lognormal distribution has provided a perfect match between the real and reconstructed plume. This may be because that the test map is bivariate lognormal distribution. As for

the SBFM reconstruction using bivariate Gaussian as basis function and NNLS reconstruction, there are still certain unsolved problems. For example, the NNLS reconstruction is not able to localize the emission source which is near the OP-FTIR while the SBFM algorithm using bivariate Gaussian distribution as basis function is not able to localize the emission source which is far away from the OP-FTIR. While in application, it is important to know whether the reconstructed source location is reliable. In the aspect of source localization, to investigate the reliability of the different reconstruction methods, the sensitivity rate and specificity rate along with the false positive rate and false negative rate in each pixel are calculated to provide the screening information after the reconstruction process.

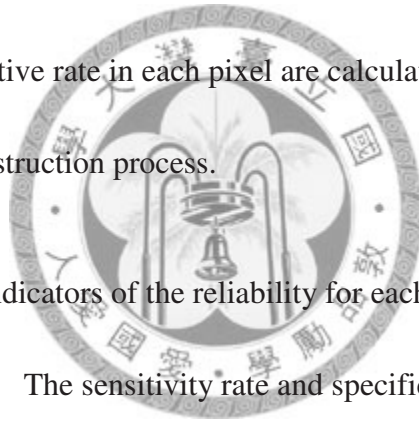


Table 3.5 shows the indicators of the reliability for each reconstruction method under the center geometry. The sensitivity rate and specificity rate are the indicators to represent the goodness of the method. The higher the sensitivity rate represents the better tool. The sensitivity rate of each pixel represents the ability of current reconstruction method to localize the emission source in current pixel. Note that the ability to localize the emission source is defined as the ability of this method to reconstruct the source location that hit the correct pixel. For example, Table 3.5 (a) shows the reliability of the  $RPM_{c\_Gauss}$ . The sensitivity rate of pixel 1 "52%" represents that 52% of the plumes with peak location in pixel 1 can be localized accurately (hit the correct pixel) by the  $RPM_{c\_Gauss}$ . The specificity rate represents the

ability of the reconstruction method to distinguish the source location that is not in current pixel. For example, Table 3.5 (a), the specificity rate of pixel 1 is 94.3% which represents that 94.3% of the plumes with peak location outside of pixel 1 can be distinguished. Associated with the specificity rate, the false positive rate is also calculated. The false positive rate means the possibility that the reconstructed source location is within the current pixel however the true source location is not. Along with the false negative rate, they are the indicators for the user to distinguish whether the reconstructed source location is reliable or not. For example, Table 3.5 (a), the false positive rate of pixel 1 is 5.8% which represents that 5.8% of the plumes outside pixel 1 are reconstructed as having source location in pixel 1. The last indicator is false negative rate which represents the possibility that the reconstructed source location is outside the current pixel however the true source location is within the current pixel. For example, Table 3.5 (a), the false negative rate of pixel 1 is 48% which represents that 48% of the plumes within pixel 1 are reconstructed as having source location outside pixel 1. A reliable method is considered to having both high sensitivity and specificity. As a result, the higher the sensitivity and specificity rate yield lower false negative and false positive rate.

Table 3.5 (a) to (c) are able to give the guideline of the reconstruction with different methods while in application. For example, in pixel 1, the false positive rate of each reconstruction method is 5.8%, 1.3% and 0 % for  $RPM_{c\_Gauss}$ ,  $RPM_{c\_log}$  and

$RPM_{c\_NNLS}$  respectively which suggests that once a reconstructed source location is in pixel 1 it is highly possible that the real source location is in pixel 1. However from the sensitivity rate (52%, 100%, 2% respectively), it can be seen that the NNLS algorithm might not be able to localize the plume with source location in pixel 1. The best method that is able to localize the emission source is SBFM reconstruction using bivariate lognormal distribution as basis function. From Table 3.5 (b), both the sensitivity rate and specificity rate are high ( $> 62\%$ ). The worst method is SBFM reconstruction using bivariate Gaussian distribution as basis function. From Table 3.5 (a), only pixel 1, 2, 4 and 5 yield both sensitivity rate and specificity rate larger than 50% which suggest that this method may work best when the source location of the plume is in pixel 1, 2, 4 and 5. The NNLS is capable of localizing the source location far from the OP-FTIR where pixel 6, 8 and 9 yield sensitivity rate and specificity rate larger than 70%. Table 3.6 (a) to (c) have also demonstrated the reliability analysis results for  $Geometry_{extend}$ . Similar to the results of  $Geometry_{center}$  where the NNLS is able to localize the source location far from the OP-FTIR (Table 3.6 (c)) while the SBFM using bivariate Gaussian distribution as basis function is able to localize the source location that is near the OP-FTIR (Table 3.6(a)). The SBFM using bivariate lognormal distribution as basis function has given the best result in every pixel (Table 3.6(b)).

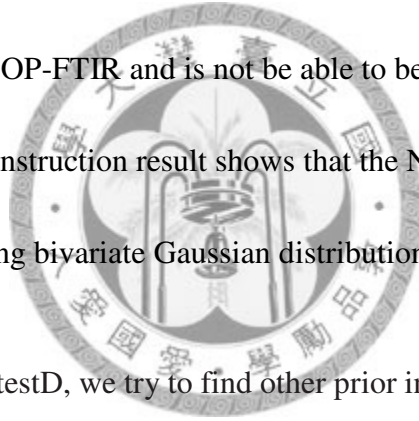
### 3.1.3 The prior screening process

From table 3.1, 3.3 and 3.4, although the high  $CCF_{PIC}$  has suggested a good match between the observed and predicted PIC however the reconstruction quality and the source localization ability still varies. Furthermore, we attempted to find possible indicators that provide information to predict the performance of the reconstruction prior to the reconstruction process.

Section 1.2 suggested that the plume with peak location near the OP-FTIR may not be localized accurately by the NNLS algorithm while the plume with peak location far away from the OP-FTIR may not be localized accurately by the SBFM algorithm using bivariate Gaussian distribution as basis function. To take advantage of this information when in field application, the prior information about whether the emission source is near or far away from the OP-FTIR must be obtained. To do so, we try to use whether the shortest monitoring is able to detect the plume (shortestD) as an indicator representing the closeness of the emission source to the OP-FTIR. Table 3.7 shows the summary statistics for the performance of reconstruction and source localization ability. It can be seen that there is significant difference in both the reconstruction quality and source localization accuracy between the test map with shortestD and that without shortestD in most of the cases (Table 3.7). Similar to the previous analysis, the SBFM with shortestD performs better than that without shortestD in both aspects of reconstruction quality and source localization accuracy.

Furthermore, the SBFM reconstruction using bivariate lognormal distribution has also

shown the same trend. The NNLS reconstruction however shows an opposite trend to the SBFM reconstruction where the test map without shortestD yields better performance than that with shortestD. This also matches the observation from section 3.1.1 and 3.1.2. For example, Figure 3.2 is one example in which the test distribution is in Pixel 1 and the shortest monitoring line for both the beam geometry is not able to detect the plume. And the reconstruction result shows that the SBFM using bivariate normal distribution as basis function is able to localize the source while the NNLS may not. Another example is shown in Figure 3.3 where the source location is in Pixel 9 which is far away from the OP-FTIR and is not be able to be detected by the shortest monitoring line. The reconstruction result shows that the NNLS is able to localize the plume while the SBFM using bivariate Gaussian distribution as basis function may not.



In addition to the shortestD, we try to find other prior information that is helpful to screen out the unsuitable PIC data or to choose a proper reconstruction method. The number of detected PIC (dPIC) and the number index of monitoring line which detects the highest concentration of PIC (maxCON) are considered. These two factors are strongly associated with the plume's location. For example, most of the plumes in pixel 1 are able to be detected by all of the monitoring line and most of the highest concentration is detected by the 4<sup>th</sup> monitoring line. We use the multiple linear regression models to analyze the fitness of possible factors which are dPIC, maxCON and ShortestD. Table 3.8 shows the regression result using dPIC, maxCON and

ShortestD as predictor variables. The  $R^2$  ranges from 0.14 to 0.44 for  $CCF_{RPM}$ , 0.05 to 0.42 for  $Nearness_{RPM}$  and 0.12 to 0.47 for peak error. The low  $R^2$  value suggests that the chosen factors may not explain the variation of the indicators used to describe the performance of reconstruction. Thus we are not able to predict the performance of the reconstruction with the three chosen factors.

### 3.1.4 The uncertainty analysis of SBFM reconstruction

Figure 3.4 shows the error map of the  $Geometry_{extend}$  using bivariate Gaussian distribution with the ratio of  $\sigma_x$  and  $\sigma_y$  being 1:1 (a circle distributed distribution from the top view) as test distribution. The figures on the left (Figure 3.4 (a) (c) (e)) show the error map with the peak location in the middle of each pixel with growing size ( $\sigma_x=1, 2$  and  $3$   $\sigma_y=1, 2$  and  $3$  respectively). The black color represents the area with standardized SSE smaller than 0.01 which represents the possible reconstructed source location when applying SBFM reconstruction. The larger the area with black color represents the larger uncertainty when applying SBFM reconstruction. The figures on the right (Figure 3.4 (b) (d) (f)) represent the overall error map with test distributions corresponding to the right figure. The white color represents the area with standardized SSE smaller than 0.01 which represents the smaller uncertainty when applying SBFM algorithm. That is, the test distribution with peak location in the white area can be localized by the SBFM algorithm more accurately comparing to those with



peak location in the black area.

From Figure 3.4, it can be seen that as the size of the test distribution grows the uncertainty decreases. Although some of the black area increases (Figure 3.4 (a) (c) (e)), the white area increases in Figure 3.4 (b) (d) (f). This indicates that although some of the test distribution in the error map of the single test distribution has increasing uncertainty, the overall uncertainty (Figure 3.4 (b) (d) (e)) still decreases. This may be because that the larger plume is able to be detected by multiple monitoring lines even it's far away from the OP-FTIR (Figure 3.4 (f)). However the smaller plume can only be detected by limited monitoring lines thus increases the uncertainty. Similar result is also observed in Figure 3.5, which is the error map of the same distribution to Figure 3.4 however the beam geometry is different.

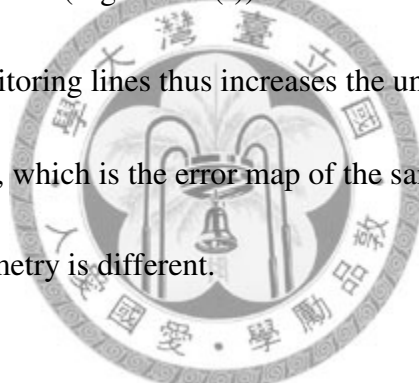
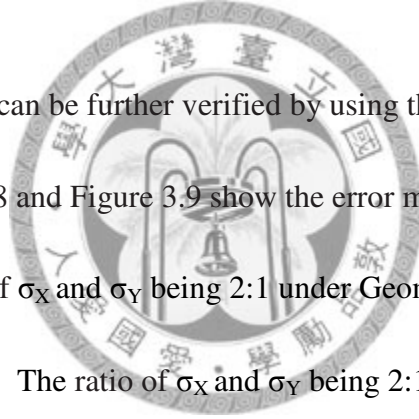


Figure 3.6 and Figure 3.7 shows the error map with the ratio of  $\sigma_x$  and  $\sigma_y$  of the test distribution being 1:2 (The principle axis of the distribution is perpendicular to the X axis). Similar results are observed where the larger the test distribution yields smaller overall uncertainty. However it is worth noticed in Figure 3.6 (d), the white area between monitoring line 7 and 9 extends towards the far end of the domain. The white area between monitoring line 7 and 9 even extends further when the plume grows larger (Figure 3.6 (f)). Since the test distribution in Figure 3.6 is parallel to the Y axis, it is likely that the monitoring line 7 and 9 can still detect the plume because they are

approximately perpendicular the Y axis. This can be confirmed by the white area between monitoring line 1 and 3 where the white area does not extend like it does between monitoring line 7 and 9. This may because monitoring line 1 and 3, same as the test distribution, are approximately parallel to the Y axis. The monitoring lines are not able to detect the plume when the plume is far away from the OP-FTIR. The same result can be observed under the different beam geometry (Figure 3.7). The white area between monitoring line 7 and 9 extends towards the far end of the domain in Figure 3.7 (d) and (f) while that between monitoring line 1 and 3 does not.

The above observation can be further verified by using the test distributions with different angle. Figure 3.8 and Figure 3.9 show the error map using the test distribution with the ratio of  $\sigma_x$  and  $\sigma_y$  being 2:1 under Geometry<sub>extend</sub> and Geometry<sub>center</sub> respectively. The ratio of  $\sigma_x$  and  $\sigma_y$  being 2:1 represents that the principle axis of the test distribution in Figure 3.8 and 3.7 are perpendicular to that in Figure 3.6 and 3.7. As expected, the white area between monitoring line 1 and 3 extends towards the farther end of the domain but not that between monitoring line 7 and 9 (Figure 3.6 (d) (f) and Figure 3.7 (d) (f)). We further use the test distribution with different ratio to generate the error map. Figure 3.10 (a) shows the error map using test distribution with the ratio between  $\sigma_x$  and  $\sigma_y$  being 1:2 ( $\sigma_x=2$   $\sigma_y=4$ ) and Figure 3.10 (b) shows the error map using test distribution with the ratio between  $\sigma_x$  and  $\sigma_y$  being 1:3 ( $\sigma_x=2$   $\sigma_y=6$ ) which represents the longer test distribution. As



expected, the white area between monitoring line 7 and 9 in Figure 3.10 (b) extends more than that in Figure 3.10 (a). This is because that the longer test distribution makes it possible to be detected by monitoring line 7 and 9 even it's far away from the OP-FTIR. Figure 3.10 (c) and (d) show the error map using test distribution that is perpendicular to that in Figure 3.10 (a) and (b) respectively. Again, the white area between monitoring line 1 and 3 in Figure 3.10 (d) extends more than that in Figure 3.10 (c). Same result is observed under the other beam geometry (Figure 3.10 (e) (f) and Figure 3.10 (g) (h)).

The above analysis suggests that the SBFM reconstruction is able to localize the plume that is close to the OP-FTIR. As the size of the plume grows, the SBFM reconstruction can also work even when the plume is far from the OP-FTIR. In addition, the plume with peak location that is far from the OP-FTIR might still be localized if it is large enough to be detected by at least two monitoring lines. This finding might be helpful for further predict the performance of the SBFM reconstruction when the plume is far from the OP-FTIR.

The same situation is not observed in the overall simulation reconstruction analysis in section 3.1.1 (Table 3.1). This may be because that the test distribution (bivariate lognormal distribution) used in the reconstruction simulation is different from the basis function (bivariate Gaussian distribution). Under this situation, a smaller plume of the

basis function close the OP-FTIR might be more flexible to fit the predicted PIC to the observed PIC. For example, in Figure 3.2 (b) and (c), the SBFM reconstruction fails to reconstruct the plume in Pixel 9. The  $CCF_{PIC}$  (0.99 for both) indicates that the SSE between the observed and reconstructed PIC is minimized.

As previously mention, we might be able to distinguish whether the plume is in the proximity of the OP-FTIR by examining whether the shortest monitoring line has detected the plume (ShortestD). Since the error map suggests that the plume far from the OP-FTIR can still be localized if it is detected by at least two monitoring line, the test distributions without ShortestD in the simulation experiment might still be localized. We further implement the ANOVA analysis to see whether there is a difference in peak error between different dPIC (how many monitoring line that detects the plume) for the test distributions that is not able to be detected by the shortest monitoring line in the simulation experiment. Table 3.9 shows the results of the ANOVA analysis on peak error with  $RPM_{e\_Gauss}$  and  $RPM_{c\_Gauss}$ . It can be seen that the  $RPM_{c\_Gauss}$  reconstruction method has shown significant difference between different dPIC and the reconstruction has shown that the dPIC of 8 yields the best result. However if we further analyze the test distributions of dPIC=6, it shows that the test distributions with dPIC=6 and 8 has included some test distributions that are in the proximity of the OP-FTIR. Thus dPIC=8 and 6 yields better result. The worst result is given by dPIC=2 and 1 respectively and most of the test distributions with dPIC=2 and 1 are

from the pixel of 9 which suggests that the test distribution far from the OP-FTIR may not be localized when the dPIC is limited even it's able to be detected by 2 monitoring lines. This has also shown that the generated test distributions may not be homogenous enough. The plume tends to grow widely when it is far from the origin even when the given  $\sigma_x$  and  $\sigma_y$  are small. Thus there is only little plume with dPIC=1 which causes the error.

### 3.2 *Field experiment results*

#### 3.2.1 *The spectrum quantification*

The detection limit for the two tracer gases in each monitoring line is shown in Table 3.10. The measured PIC that is below detection limit is subjects to 0 to prevent possible influence on the source localization procedure. All of the collected spectra are confirmed to show no sign of saturation by visually examine the spectral range in the proximity of  $650\text{ cm}^{-1}$ . The maximum of the measured PIC for  $\text{SF}_6$  is  $4.02\pm 5.49$  ppm×m for Retro 5 while the minimum is  $0.28\pm 0.84$  ppm×m for Retro 6, The maximum measured PIC for  $\text{N}_2\text{O}$  is  $27.30\pm 52.41$  ppm×m for retro 3 while the minimum is  $0.49\pm 2.01$  ppm×m for retro 6. Table 3.11 (a) and (b) have shown the summary statistics of PIC data for  $\text{SF}_6$  and  $\text{N}_2\text{O}$  respectively. Figure 3.11 and Figure 3.12 have further shown the time series plot of the PIC data. In all of the experiments, the PIC data of the first few sweeps (1 to 3) for both the tracer gases is 0 which indicates that

there is no background concentration increment for the two tracer gases in the experimental domain. The concentration rises after the releasing of the tracer gases which implies that the measured concentration is from the artificially released tracer gases. The PIC data that is measured after the release of tracer gases is used to localize the source. Thus for each experiment, there are at least 21 sets of PIC data that can be used for source localizing. However, the N<sub>2</sub>O has run out during experiment 4 thus there are only 17 sets of PIC information available.

### 3.2.2 The source localization

The emission source reconstruction result is shown in Table 3.12. The  $CCF_{PIC}$  indicates the fitness between the observed and predicted PIC. The highest  $CCF_{PIC}$  is given by the NNLS reconstruction of N<sub>2</sub>O in experiment 2 ( $CCF_{PIC}=0.99\pm 0.01$ ) and SBFM reconstruction using bivariate lognormal distribution as basis function ( $SBFM_{lognormal}$ ) of N<sub>2</sub>O in Exp 4 ( $CCF_{PIC}=0.99\pm 0.02$ ) while the lowest mean  $CCF_{PIC}$  is given by the NNLS reconstruction of N<sub>2</sub>O in Exp 3 ( $CCF_{PIC}=0.77\pm 0.26$ ). The high  $CCF_{PIC}$  value in each reconstruction method shows a good fit between observed and predicted PIC however the source localization accuracy varies. The best result is given by  $SBFM_{Gaussian}$  reconstruction for N<sub>2</sub>O in Exp 3 (mean peak error= $5.26\pm 3.78$ ) while the worst result is given by NNLS reconstruction for SF<sub>6</sub> in the Exp 4 (mean peak error= $19.93\pm 4.41$ ). This shows conformity to the simulation experiment where high

$CCF_{PIC}$  does not promise an accurate source localization result.

The poor performance of NNLS reconstruction for  $SF_6$  in Exp 4 may be because that the source is in pixel 1 which is close to the OP-FTIR. As shown in the simulation experiment, the NNLS reconstruction is not able to localize the emission source that is close to the OP-FTIR. This may be because that the pixel close to the OP-FTIR is passed by all of the monitoring lines and as a result, the concentration in pixel 1 is correlated to the PIC data for all of the beam paths. In this case, a high concentration in pixel 1 may increase the predicted PIC in all of the ray paths. However, there lies a possibility that the monitoring lines on the upwind site may not be able to detect the plume even when the plume is with peak location in the proximity of OP-FTIR. The algorithm thus subjects a low concentration value in pixel 1 to avoid high reconstructed PIC in the monitoring lines on the upwind site. And for the monitoring lines with high PIC, the NNLS algorithm tends to subject the high concentration in the pixel that is only passed by that monitoring line. Figure 3.14 (f) shows the result using NNLS algorithm to localize the source location of  $SF_6$  in Exp 4 and Table 3.10 (d) shows the time series plot of the PIC data of  $SF_6$  in Exp 4. From table 3.10 (d), the PIC rises mostly in Path 3, 4, 7 and 9 but not 1 and 2 and the highest concentration is detected by Path 7. Thus it can be predicted that although the emission source is in Pixel 1, the reconstructed source location by NNLS might still subject a low concentration value in Pixel 1 to fit the low observed PIC in Path 1 and 2. Furthermore, the reconstructed source location

can be predicted to be near Pixel 8 which is passed by Path 7 only (Figure 3.14 (f)).

Another example is shown in Exp 2 SF<sub>6</sub> in which the emission source is also in Pixel 1

(Figure 3.12 (f)). The NNLS still gives the worst result in all three kinds of

reconstruction methods (mean peak error= $17.44 \pm 7.02$ ). From the time series plot of

PIC data (Figure 3.11 (b)), the PIC data rises mostly in Path 4, 5 and 7 but not Path 1

and 2 and the highest concentration is detected by Path 5. As a result, most of the

reconstructed source locations are in Pixel 5 and 9 which correlate more to Path 5. As

stated in the simulation experiment, the NNLS reconstruction is able to localize the

source that is far from the OP-FTIR. For example, the NNLS reconstruction gives the

best result for SF<sub>6</sub> in Exp 1 (mean peak error= $8.88 \pm 5.41$ ) in which the source location is

in pixel 8 (Figure 3.13 (c)). From the series plot (Figure 3.11 (a)), the concentration

rises mostly in Path 7 and 9 and the highest concentration is detected by Path 7. Thus

most of the reconstructed source locations are in Pixel 8 which is only passed by Path 7

and some of the reconstructed source locations are in Pixel 7 which is passed only by

Path 9.

Regarding to the studies of Hashmonay et al. in 2002<sup>19</sup> and 2008<sup>37</sup>, the result found in this study for using NNLS algorithm to locate the pollutant may indicate the uncertainty of the result in the previous two. In the study of Hashmonay in 2008, the RPM technique using NNLS is implemented in a landfill to locate the unknown “hot spot” of the pollutant. The result shows that the reconstruction is almost identical to

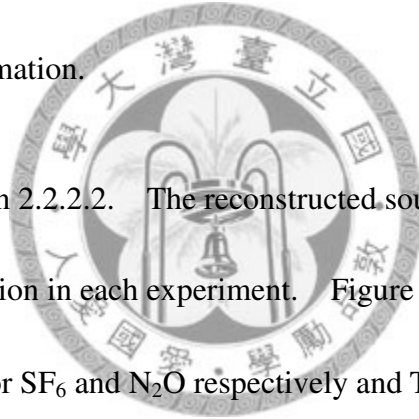


that reconstructed by MART algorithm. However, it can be seen that in this study, the reconstructed “hot spot” of the pollutant is in the farther end of the experimental domain (Figure 1.6 (a)). Comparing to the result in our study, there is a possibility that the real “hot spot” of the pollutant may be close to the optical remote sensing instrument however the NNLS reconstruction is not able to tell it apart. In the other study<sup>19</sup>, a total of 5 experiments are implemented using the RPM technique to localize the artificially released source. In each experiment, the SF<sub>6</sub> is released as tracer gas to be localized. Figure 1.6 (b) has shown the successful reconstruction results to the artificially released tracer gases. Also, in these success examples, the tracer gases are located in the farther end of the emission domain. And the results of the other two examples are not discussed. Thus the result of having the emission source in the proximity of the emission domain still remains unknown. Although in this case, the reconstruction algorithm is MART instead of NNLS, the study in 2008 has proved that the MART is able to reconstruct an identical result to that using NNLS algorithm.

The SBFM reconstruction using bivariate Gaussian distribution as basis function gives the significantly better result than the other methods in Exp 2, 3 and 4 for SF<sub>6</sub> (Table 3.12(a)). In all of these cases, the source locations are near the OP-FTIR (Pixel 1, Pixel 4 and Pixel 1 respectively). This has also been verified in the simulation experiment where the source location that is near the OP-FTIR yields the shortest peak error. In addition, the error map analysis has also supported this result where in Figure

3.9, the white area covers the region that in the proximity of the OP-FTIR. This may be because that the closer the emission source to the OP-FTIR, the higher opportunity that all of the monitoring line is able to detect the plume. Once the plume is detected by all of the monitoring lines, the SBFM may not be able to tell apart whether the plume is a narrow plume near the OP-FTIR or a wide plume far from the OP-FTIR. In this case, the SBFM tends to reconstruct the source in the region that is close to the OP-FTIR because the smaller plume may be more flexible to fit the predicted PIC to the observed PIC. In Figure 3.13 (d) SF<sub>6</sub> in Exp 2, although the SBFM<sub>Gaussian</sub> gives the significantly better result than the others, most of the reconstructed source locations are in Pixel 5 rather than Pixel 1. This can be explained by the measured PIC data. From Figure 3.11 (b), Path 6 does not detect any tracer gas in most of the times thus giving the segmenting information to avoid the source being reconstructed in Pixel 1. The star sign in Figure 3.13 (d) shows a result that is reconstructed in the correct pixel. From Figure 3.11 (b), it is able to be detected by all of the monitoring lines. For the cases that the SBFM<sub>Gaussian</sub> gives worse result, the source locations are all far from the OP-FTIR. In this case, the uncertainty analysis suggests that if the plume is able to be detected by at least two monitoring line, it is still possible to be localized by the SBFM reconstruction. An example is shown in Exp 1 SF<sub>6</sub> (Figure 3.13 (a)) in which, the relatively poor result is given by SBFM<sub>Gaussian</sub> (mean peak error=9.00±6.38). However, several reconstruction results are still close to the real source location. For example,

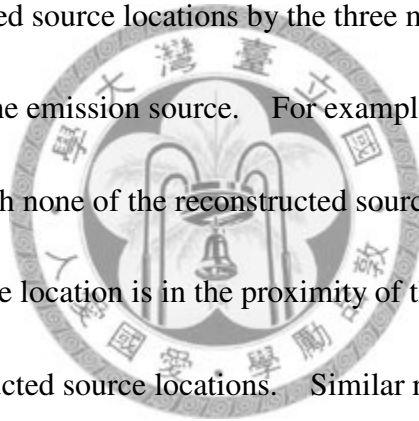
the 20<sup>th</sup> set of input PIC has yields the result with peak error equaling 8 (Figure 3.13 (a) star sign). In Figure 3.11 (a), the 20<sup>th</sup> set of input PIC is detected by path 5 and 7 only, which confirms the uncertainty analysis in the computational simulation. Furthermore, the reconstruction using the 21<sup>th</sup> set of input PIC is detected by path 5 and 7 (Figure 3.11 (a)) which are the longer monitoring lines in the beam geometry. The reconstruction with this set of PIC has yielded the reconstruction result with peak error equaling 8 (Figure 3.13 (a) triangle sign) which suggests that when the emission source is far from the FTIR, the SBFM may still reconstruct the source location if there is adequate segmenting information.



As described in section 2.2.2.2. The reconstructed source locations are averaged to give a mean source location in each experiment. Figure 3.17 and 3.18 show the averaged source location for SF<sub>6</sub> and N<sub>2</sub>O respectively and Table 3.13 (a) and (b) show the reconstruction performance of the averaged result of SF<sub>6</sub> and N<sub>2</sub>O. In table 3.13 (a), the best result of the SF<sub>6</sub> source reconstruction is given by NNLS reconstruction (peak error=2.90) in Exp 1 while the worst result is given by NNLS reconstruction (peak error=18.03) in the Exp 4. In table 3.13 (b), the best result of the N<sub>2</sub>O reconstruction is given by SBFM reconstruction using bivariate lognormal distribution as basis function (peak error=0.49) in Exp 1 while the worst result is given by NNLS reconstruction in Exp 3 (peak error=10.59). Similar trend of the performance of the different reconstruction methods to the overall reconstruction is observed. The NNLS

is able to localize the plume that is far from the OP-FTIR while the SBFM is able to localize the plume that is close to the OP-FTIR.

Carefully looking at the averaged reconstruction result, it can be seen that the NNLS has reconstructed the source location mostly in pixel 6, 8 and 9 while the SBFM using bivariate Gaussian distribution as basis function has reconstructed the source location mostly in pixel 2, 4 and 5. The SBFM using bivariate lognormal distribution as basis function has reconstructed the source location between the other two methods. In addition, the reconstructed source locations by the three methods are able to point out the proximate location of the emission source. For example, for the SF<sub>6</sub> in experiment 2 (Figure 3.17 (b)), although none of the reconstructed source location has hit the correct pixel, the real source location is in the proximity of the extension of the line between the three reconstructed source locations. Similar results can be seen in all other experiments. With this finding, combining the measured PIC, we might be able to choose which reconstruction result is the closest to the real source location. For example, the SF<sub>6</sub> in Exp4 (Figure 3.17 (b)), the reconstructed source location by the 3 methods are between monitoring line 5 and 7. From the measured PIC data (Figure 3.11(d)), it can be seen the nearby short beam paths have also detected the plume (Path4 and Path8). Thus we might be able to judge that the source location is in the proximity of the SBFM reconstruction using bivariate Gaussian distribution as basis function. Another example, Exp2 N<sub>2</sub>O (Figure 3.18 (b)), the reconstructed source locations are

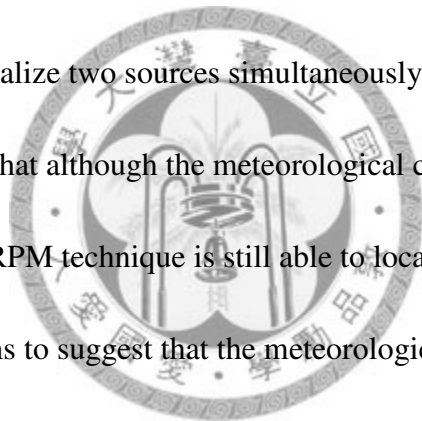


between monitoring line 3 and 4. Judging by the measured PIC (Figure 3.12 (b)), the plume is not detected by nearby short paths (Path4 and Path6), thus we might be able to choose the reconstructed location by NNLS as the real source location. From the above example, we might be able to judge which reconstructed source location is the closest to the real source location by looking at the peripheral short paths. Table 3.14 has further shown the result by based on this rule. In all four experiments, the chosen results based on this hypothesis are all the closest to the real source location.

Another possible solution is to apply the 1-D radial plume mapping described in previous studies<sup>15, 31, 32, 35</sup>. In Figure 3.17 and 3.18, a virtual line can be created by linking the reconstructed source locations of the three methods. The distance between the line and the real source location is less than 5 meters. If the 1-D RPM is applied by setting up the beam geometry on this line the result might be more accurate. In previous studies, the SBFM reconstruction in a 1-D scenario using different input PIC strategies and different basis functions have been proposed<sup>15, 31, 32</sup>. Since the real source location is close to the monitoring line, we might be able to simply localize the peak concentration location along the monitoring line as the source location. Also, the 1-D RPM combining wind data can also be implemented to further localize the source.

Table 3.15 has shown the wind data of the four experiments. It can be seen that the average wind speed ranges from 0.43 to 0.86 m/s which indicates that the

experiment was conducted in the field with weak wind. The wind direction varies rapidly in the four experiments. The standard deviation of the wind direction ranges from 110 to 133 degrees. Figure 3.19 has further shown the wind rose plot of the four experiments. It can be seen that the mode of the wind direction is 0 degree (due north) which is approximately perpendicular to the monitoring lines. Comparing the reconstruction result in each experiment, it can be seen that the RPM technique is able to localize the source with different reconstruction algorithm. This suggests that under the same meteorological condition, although the source location varies, the RPM technique is still able to localize two sources simultaneously. Comparing the different experiment, it can be seen that although the meteorological condition is significantly different (Table 3.15), the RPM technique is still able to localize the source. Although the above observation seems to suggest that the meteorological condition cause no effect for the RPM technique to localize the source, it still needs to be noticed that the segmenting information is needed when applying the RPM technique. If the wind direction is due to the location of OP-FTIR, the segmenting information is limited. The reason that the experiment is less affected by the wind may be because that in most of the time, the wind direction is perpendicular to the monitoring lines. Thus the segmenting information is adequate for the SBFM to localize the source.

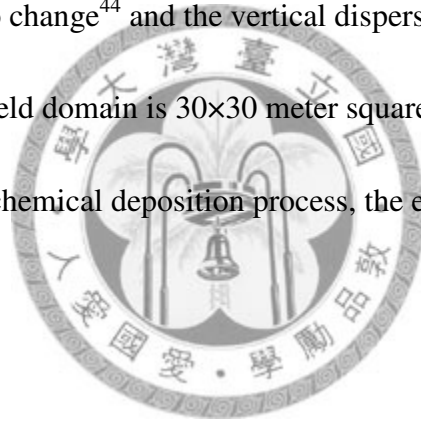


### *3.2.3 The reliability of the reconstruction result*

As mentioned in the simulation experiment, the section of reliability of the reconstruction result, table 3.5 is able to provide a guideline of the reconstruction. In order to validate whether the reconstruction result is reliable, the mean source location of each reconstruction method is first calculated as described in section 2.2.2.2. Figure 3.17 and 3.18 show the averaged result for SF<sub>6</sub> and N<sub>2</sub>O respectively. It is difficult to tell whether the reconstructed source location is reliable or not by table 3.5. For example, Figure 3.16 (a) shows the averaged result for SF<sub>6</sub> in experiment 1. The SBFM<sub>Gaussian</sub> reconstructed the source location in pixel 4 and Table 3.5 (a) has shown 13.0% of false positive rate which shows the uncertainty of the reconstructed source location. Indeed, the real source location is not in pixel 4 however the 13.0% of possibility makes it hard to judge whether to be convinced by the reconstruction result. Table 3.5 (b) has shown the reliability of the SBFM<sub>lognormal</sub> reconstruction. The high sensitivity rate and low false positive rate has suggested that the reconstruction result may be highly reliable. However, in the field experiment, the SBFM<sub>lognormal</sub> has only reconstructed two out of eight source locations that are able to hit the correct pixel. As mentioned previously, the high reliability of the reconstruction method from the simulation experiment may come from the similarity between the test distribution and the basis function which both are the bivariate lognormal distribution. As for the NNLS reconstruction, similar to the SBFM<sub>Gaussian</sub>, the high sensitivity and specificity rate of pixel 8 suggests that the NNLS is able to reconstruct the source location in Pixel

8. The 26.5% of false positive rate makes it difficult to rely on the reconstruction by the NNLS algorithm. Again, the NNLS reconstruction tends to reconstruct the source location far away from the OP-FTIR. Four out of eight of the emission sources are reconstructed in Pixel 8 however none of them hit the correct pixel.

Despite the error caused by the RPM technique, the property of the pollution might also influence the performance of the RPM technique to localize the source. For example, the acid deposition of the chemical substances<sup>43</sup>, the turbulence causing the distribution of the plume to change<sup>44</sup> and the vertical dispersion of the pollutant<sup>45</sup>. In our field experiment, the field domain is 30×30 meter square. Although the relatively small area might limit the chemical deposition process, the error might still exist.





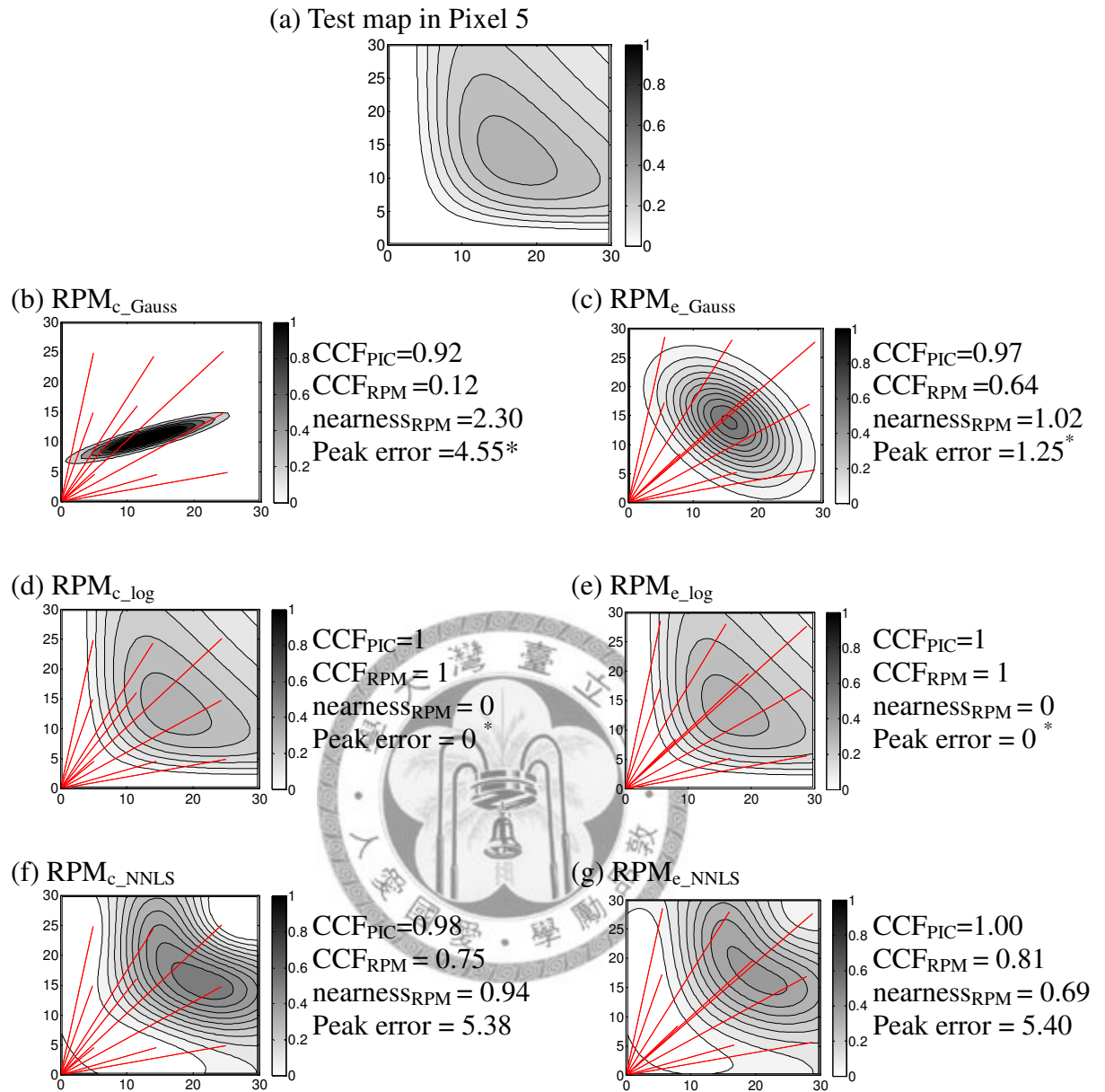


Figure 3.1 The reconstruction results for different reconstruction methods (a) The original test map. (b) (c) The reconstruction by the SBFM using bivariate Gaussian distribution as basis function under  $Geometry_{center}$  and  $Geometry_{extend}$  respectively. (d) (e) The reconstruction by the SBFM using bivariate lognormal distribution as basis function under  $Geometry_{center}$  and  $Geometry_{extend}$  respectively. (f) (g) NNLS reconstruction under  $Geometry_{center}$  and  $Geometry_{extend}$  respectively. The \* sign after the peak error indicates that the reconstructed peak location has “hit” the correct pixel.

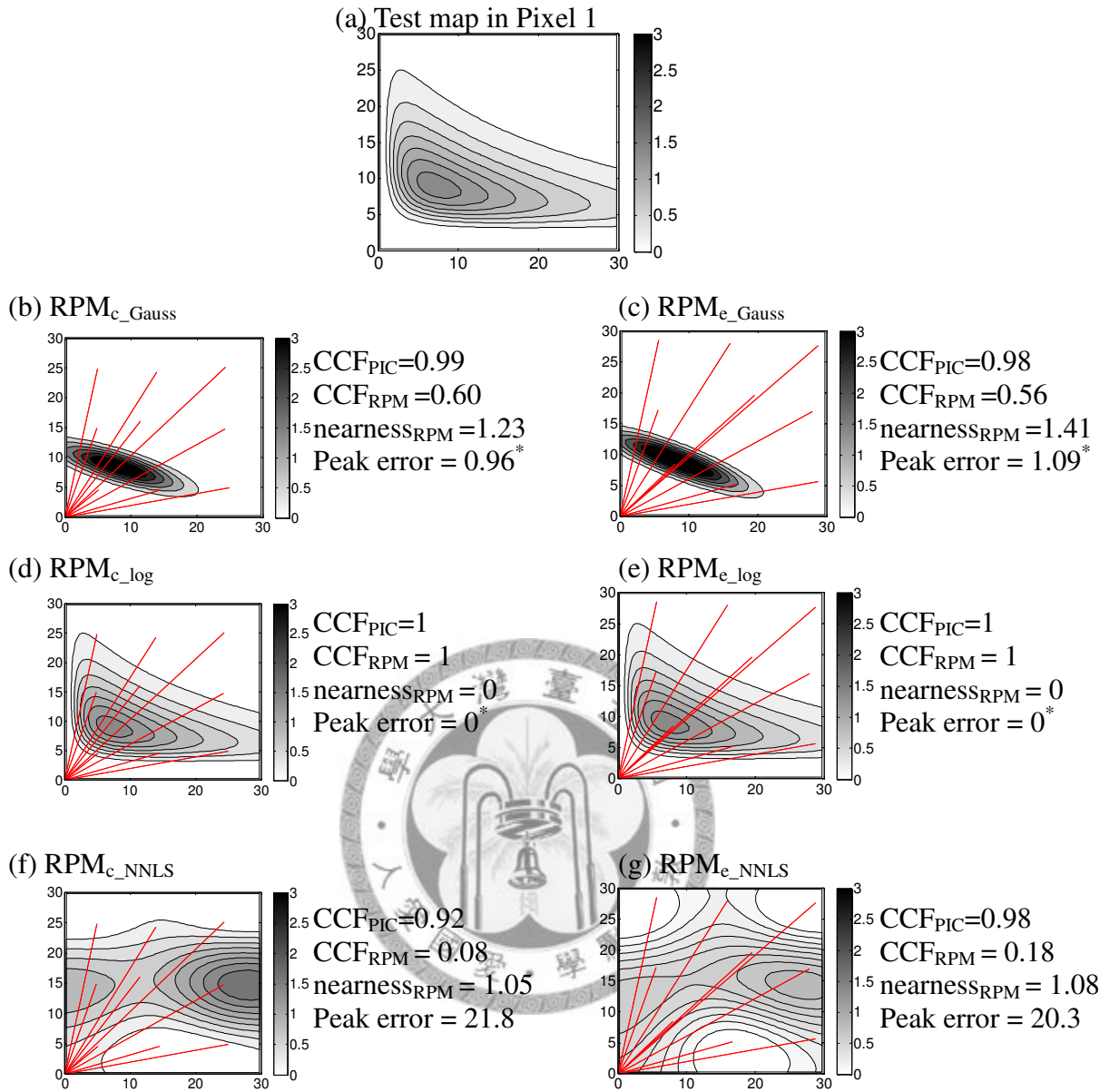


Figure 3.2 The reconstruction results for different reconstruction methods (a) The original test map. (b) (c) The reconstruction by the SBFM using bivariate Gaussian distribution as basis function under  $Geometry_{center}$  and  $Geometry_{extend}$  respectively. (d) (e) The reconstruction by the SBFM using bivariate lognormal distribution as basis function under  $Geometry_{center}$  and  $Geometry_{extend}$  respectively. (f) (g) NNLS reconstruction under  $Geometry_{center}$  and  $Geometry_{extend}$  respectively. The \* sign after the peak error indicates that the reconstructed peak location has “hit” the correct pixel.

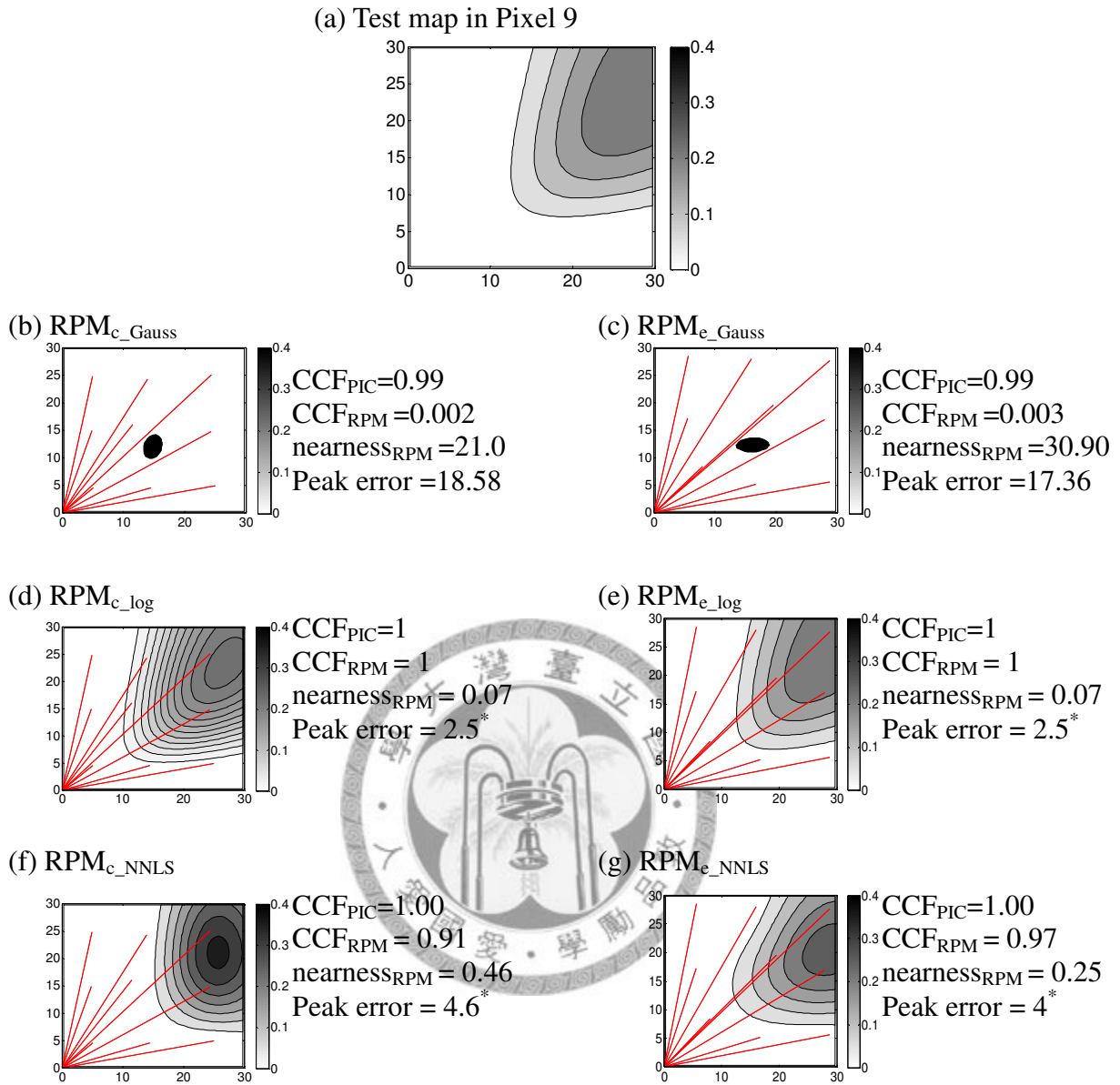


Figure 3.3 The reconstruction results for different reconstruction methods (a) The original test map. (b) (c) The reconstruction by the SBFM using bivariate Gaussian distribution as basis function under  $\text{Geometry}_{\text{center}}$  and  $\text{Geometry}_{\text{extend}}$  respectively. (d) (e) The reconstruction by the SBFM using bivariate lognormal distribution as basis function under  $\text{Geometry}_{\text{center}}$  and  $\text{Geometry}_{\text{extend}}$  respectively. (f) (g) NNLS reconstruction under  $\text{Geometry}_{\text{center}}$  and  $\text{Geometry}_{\text{extend}}$  respectively. The \* sign after the peak error indicates that the reconstructed peak location has “hit” the correct pixel.

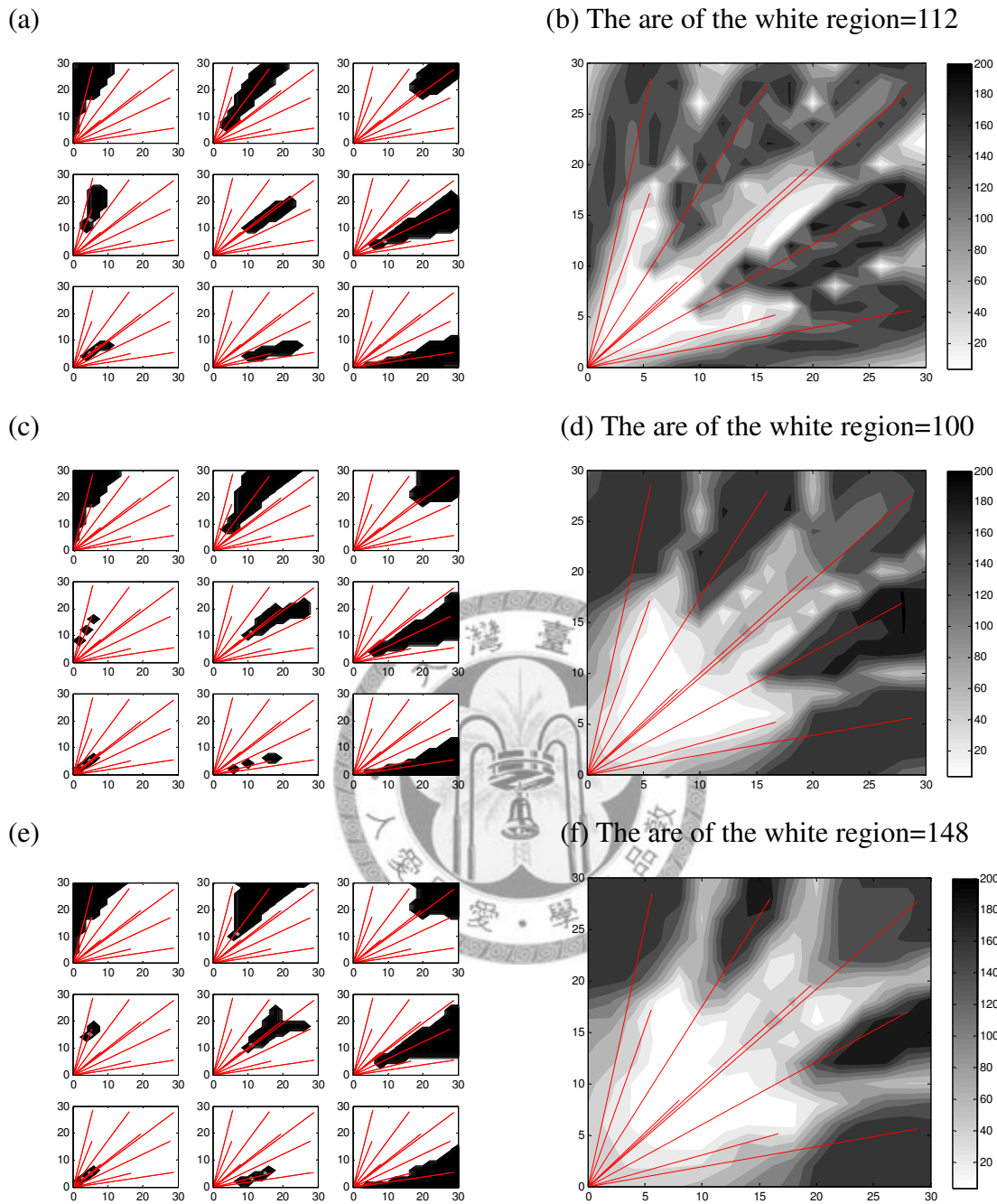
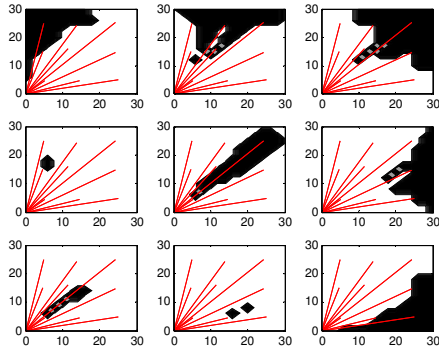
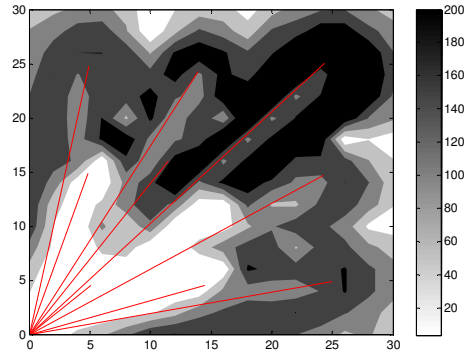


Figure 3.4 The error map of  $\text{Geometry}_{\text{extend}}$  using different size of test distributions with the ratio of  $\sigma_X$  and  $\sigma_Y$  being 1:1. The figures on the left are the error maps with only one test distribution in the middle of each pixel. The black color represents the standardized SSE  $< 0.01$ . The figures on the right are the overall error map. The white color representing the area with standardized SSE  $< 0.01$ . (a) (b)  $\sigma_X=1$   $\sigma_Y=1$  (c) (d)  $\sigma_X=2$   $\sigma_Y=2$  (e) (f)  $\sigma_X=3$   $\sigma_Y=3$

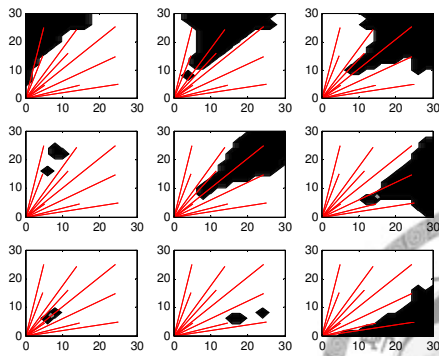
(a)



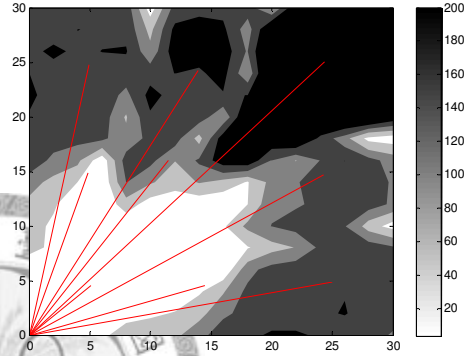
(b) The are of the white region =120



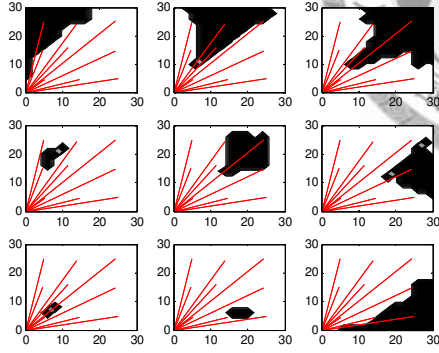
(c)



(d) The are of the white region =128



(e)



(f) The are of the white region =140

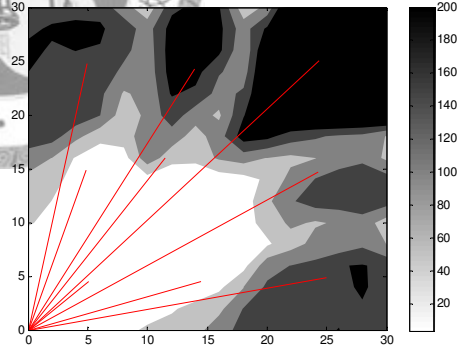
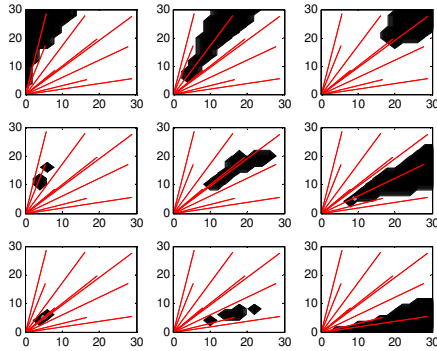
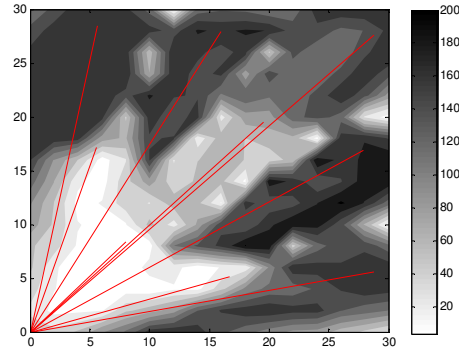


Figure 3.5 The error map of Geometry<sub>center</sub> using different size of test distributions with the ratio of  $\sigma_X$  and  $\sigma_Y$  being 1:1. The figures on the left are the error maps with only one test distribution in the middle of each pixel. The black color represents the standardized SSE < 0.01. The figures on the right are the overall error map. The white color representing the area with standardized SSE < 0.01. (a) (b)  $\sigma_X=1$   $\sigma_Y=1$  (c) (d)  $\sigma_X=2$   $\sigma_Y=2$  (e) (f)  $\sigma_X=3$   $\sigma_Y=3$

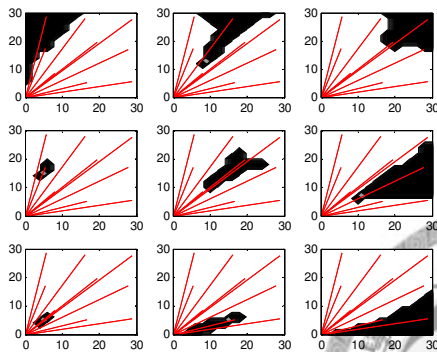
(a)



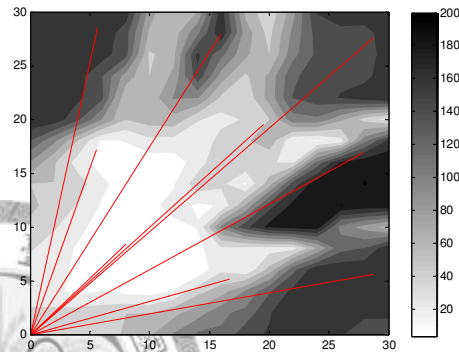
(b) The are of the white region =112



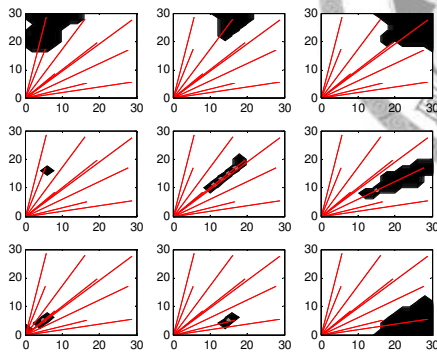
(c)



(d) The are of the white region =132



(e)



(f) The are of the white region =248

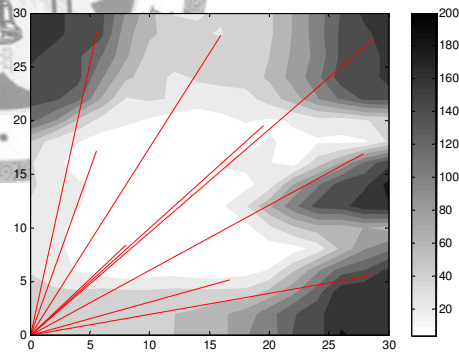


Figure 3.6 The error map of  $Geometry_{\text{extend}}$  using different size of test distributions with the ratio of  $\sigma_X$  and  $\sigma_Y$  being 1 to 2. The figures on the left are the error maps with only one test distribution in the middle of each pixel. The black color represents the standardized SSE  $< 0.01$ . The figures on the right are the overall error map. The white color representing the area with standardized SSE  $< 0.01$ . (a) (b)  $\sigma_X=1$   $\sigma_Y=2$  (c) (d)  $\sigma_X=2$   $\sigma_Y=4$  (e) (f)  $\sigma_X=3$   $\sigma_Y=6$

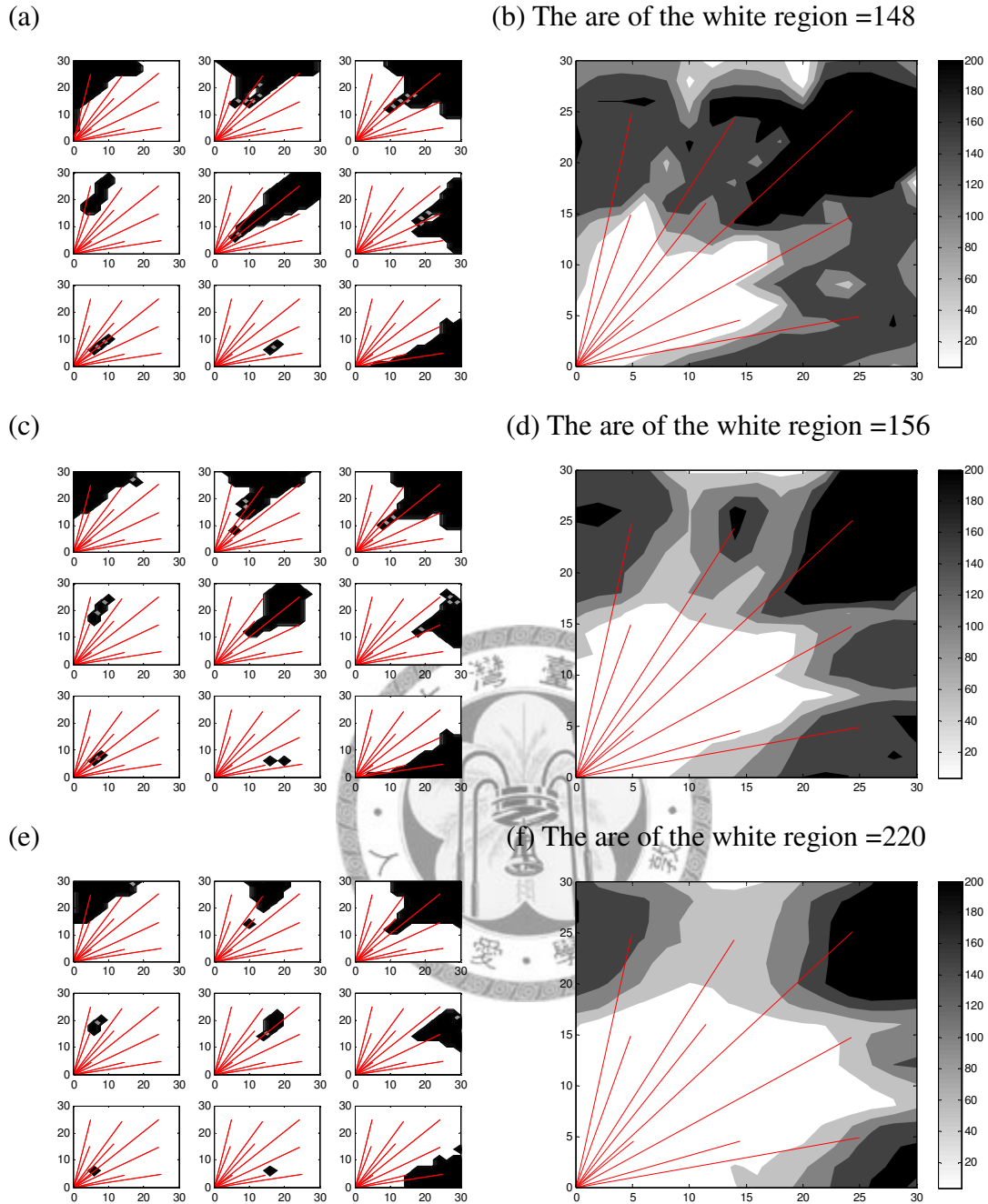


Figure 3.7 The error map of  $\text{Geometry}_{\text{center}}$  using different size of test distributions with the ratio of  $\sigma_X$  and  $\sigma_Y$  being 1 to 2. The figures on the left are the error maps with only one test distribution in the middle of each pixel. The black color represents the standardized SSE < 0.01. The figures on the right are the overall error map. The white color representing the area with standardized SSE < 0.01. (a) (b)  $\sigma_X=1$   $\sigma_Y=2$  (c) (d)  $\sigma_X=2$   $\sigma_Y=4$  (e) (f)  $\sigma_X=3$   $\sigma_Y=6$

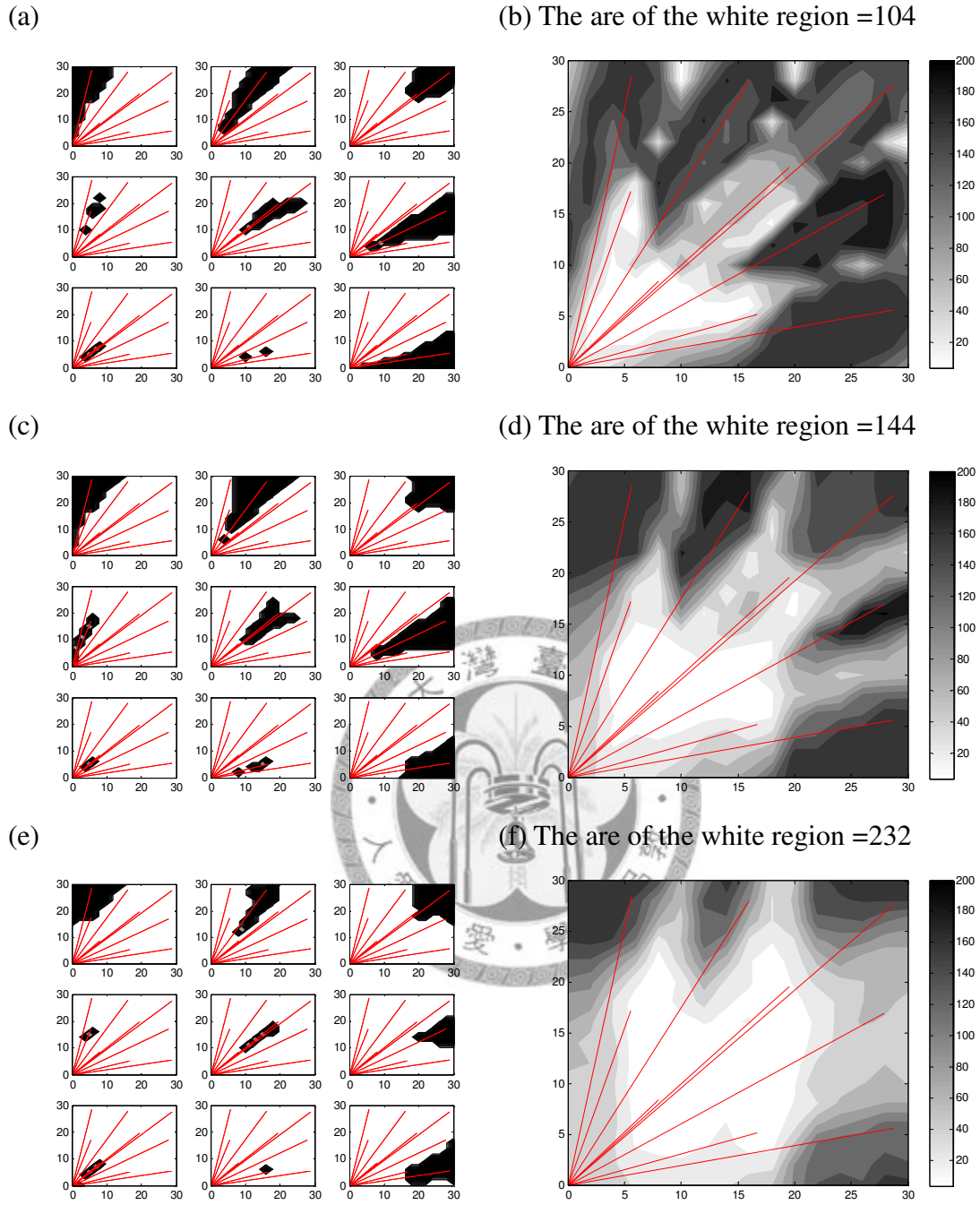


Figure 3.8 The error map of  $\text{Geometry}_{\text{extend}}$  using different size of test distributions with the ratio of  $\sigma_X$  and  $\sigma_Y$  being 2 to 1. The figures on the left are the error maps with only one test distribution in the middle of each pixel. The black color represents the standardized SSE < 0.01. The figures on the right are the overall error map. The white color representing the area with standardized SSE < 0.01. (a) (b)  $\sigma_X=2$   $\sigma_Y=1$  (c) (d)  $\sigma_X=4$   $\sigma_Y=2$  (e) (f)  $\sigma_X=6$   $\sigma_Y=3$



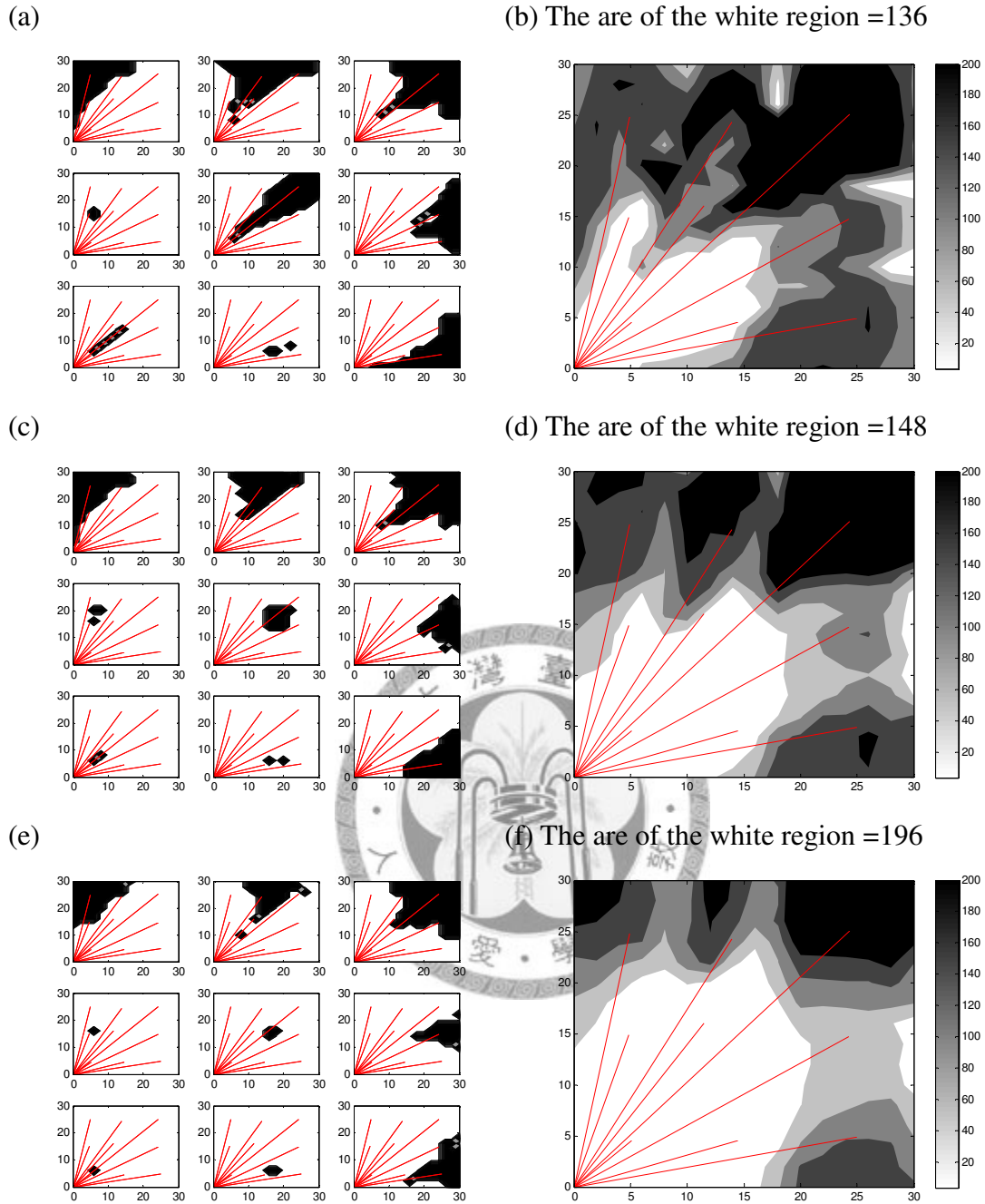
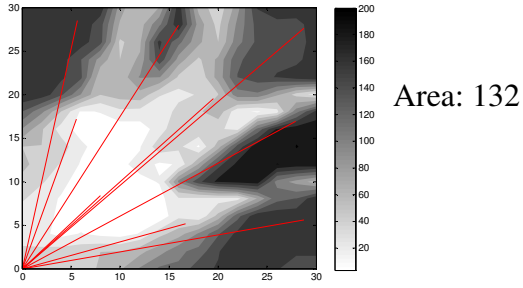
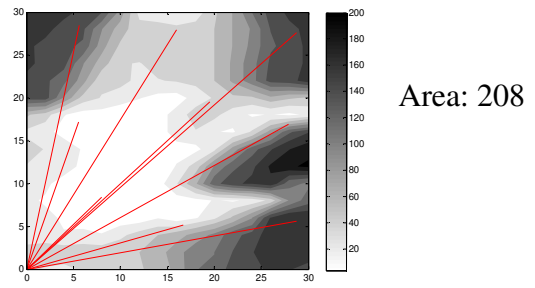


Figure 3.9 The error map of  $\text{Geometry}_{\text{center}}$  using different size of test distributions with the ratio of  $\sigma_X$  and  $\sigma_Y$  being 2 to 1. The figures on the left are the error maps with only one test distribution in the middle of each pixel. The black color represents the standardized SSE  $< 0.01$ . The figures on the right are the overall error map. The white color representing the area with standardized SSE  $< 0.01$ . (a) (b)  $\sigma_X=2$   $\sigma_Y=1$  (c) (d)  $\sigma_X=4$   $\sigma_Y=2$  (e) (f)  $\sigma_X=6$   $\sigma_Y=3$

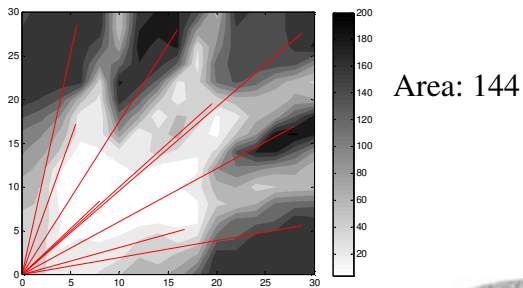
(a)  $\sigma_X=2 \sigma_Y=4$  Geometry<sub>extend</sub>



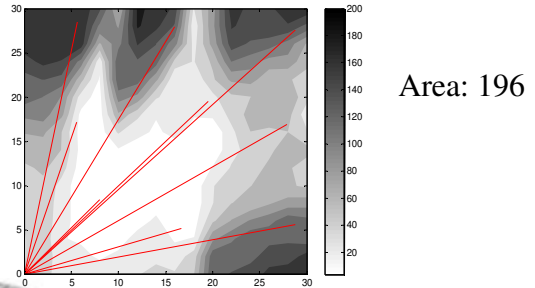
(b)  $\sigma_X=2 \sigma_Y=6$  Geometry<sub>extend</sub>



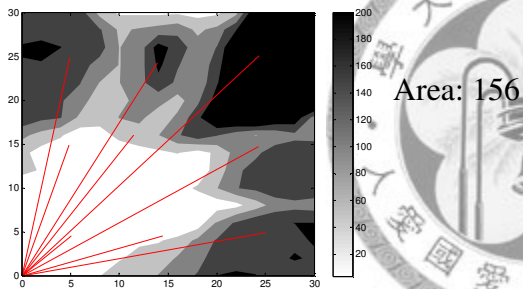
(c)  $\sigma_X=4 \sigma_Y=2$  Geometry<sub>extend</sub>



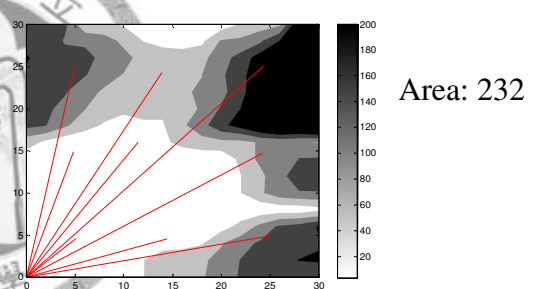
(d)  $\sigma_X=6 \sigma_Y=2$  Geometry<sub>extend</sub>



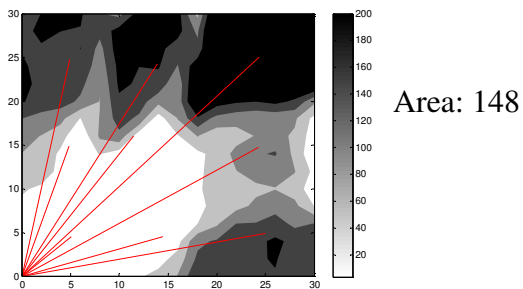
(e)  $\sigma_X=2 \sigma_Y=4$  Geometry<sub>center</sub>



(f)  $\sigma_X=2 \sigma_Y=6$  Geometry<sub>center</sub>



(g)  $\sigma_X=4 \sigma_Y=2$  Geometry<sub>center</sub>



(h)  $\sigma_X=6 \sigma_Y=2$  Geometry<sub>center</sub>

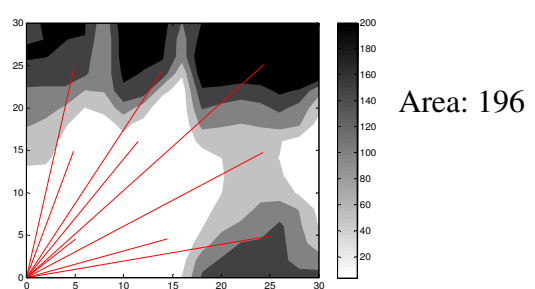
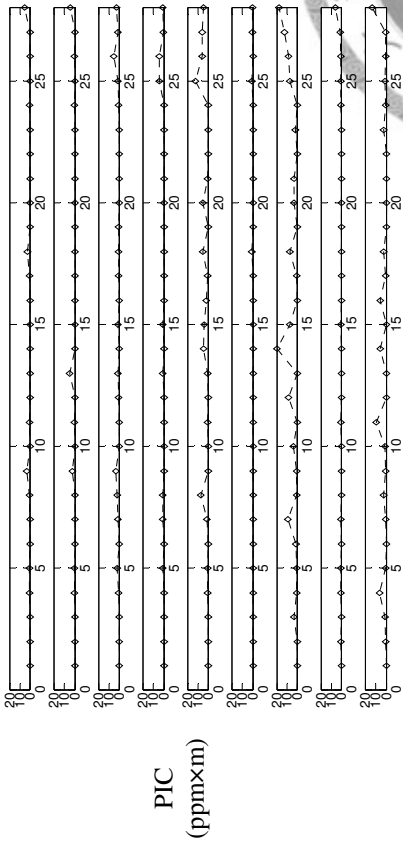
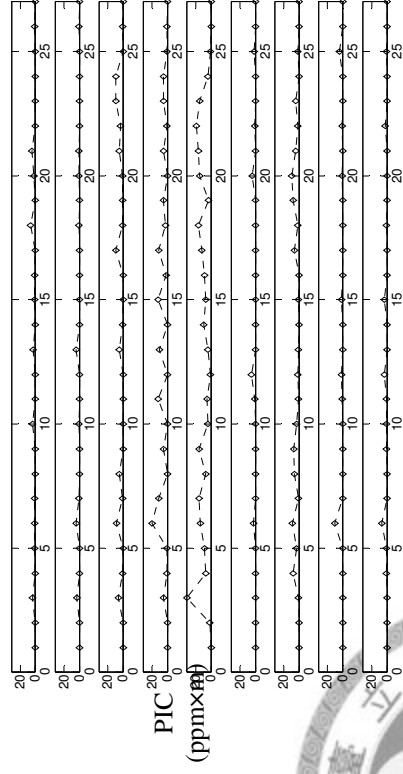


Figure 3.10 The error map using test distribution with the ratio of  $\sigma_X$  and  $\sigma_Y$  being (a) 1 to 2 (b) 1 to 3 (c) 2 to 1 and (d) 3 to 1 under Geometry<sub>extend</sub> (e) 1 to 2 (f) 1 to 3 (g) 2 to 1 (h) 3 to 1 under Geometry<sub>center</sub>

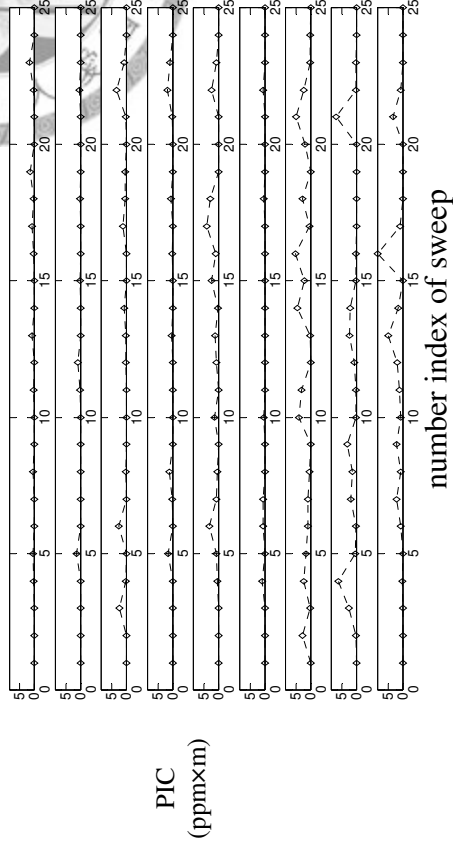
(a) Exp1 SF<sub>6</sub>



(b) Exp2 SF<sub>6</sub>



(c) Exp3 SF<sub>6</sub>



(d) Exp4 SF<sub>6</sub>

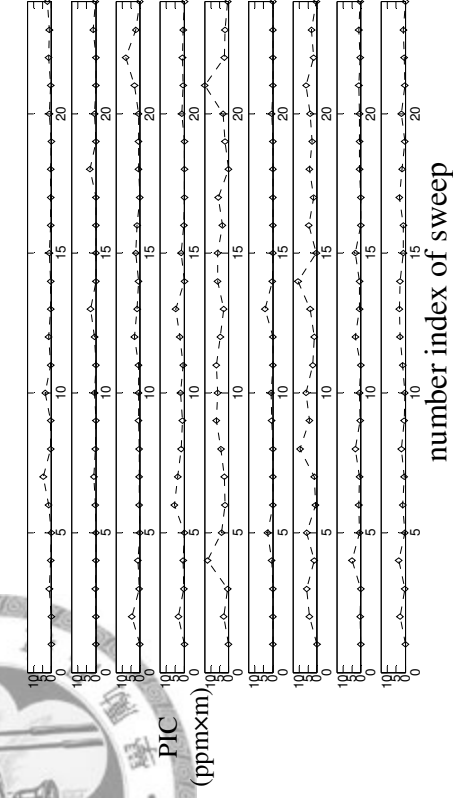


Figure 3.11 The time series plot of PIC in SF<sub>6</sub> in four experiments (a) Exp1 (2) Exp2 (3) Exp3 (4) Exp4

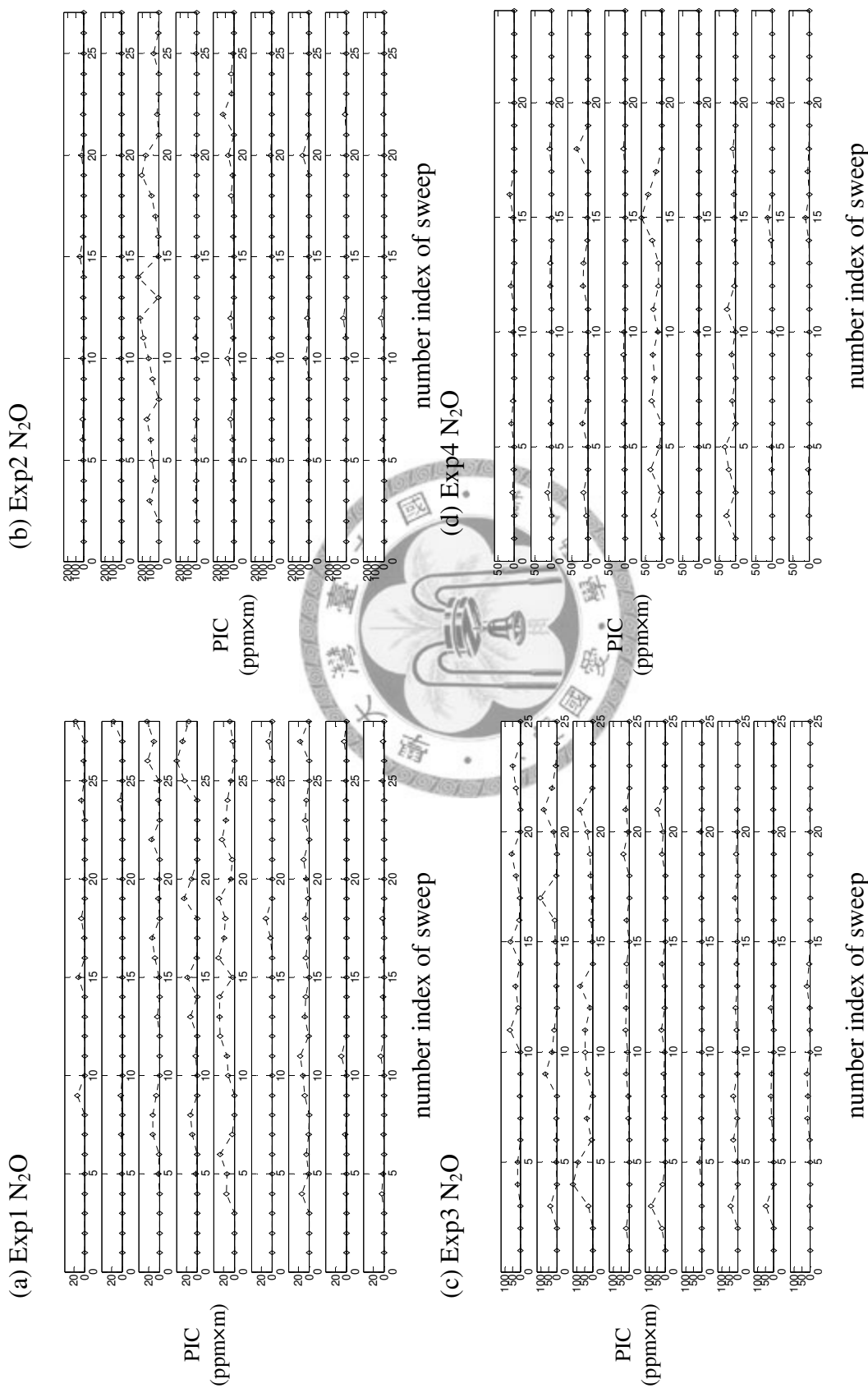


Figure 3.12 The time series plot of PIC in N<sub>2</sub>O in four experiments (a) Exp1 (2) Exp2 (3) Exp3 (4) Exp4

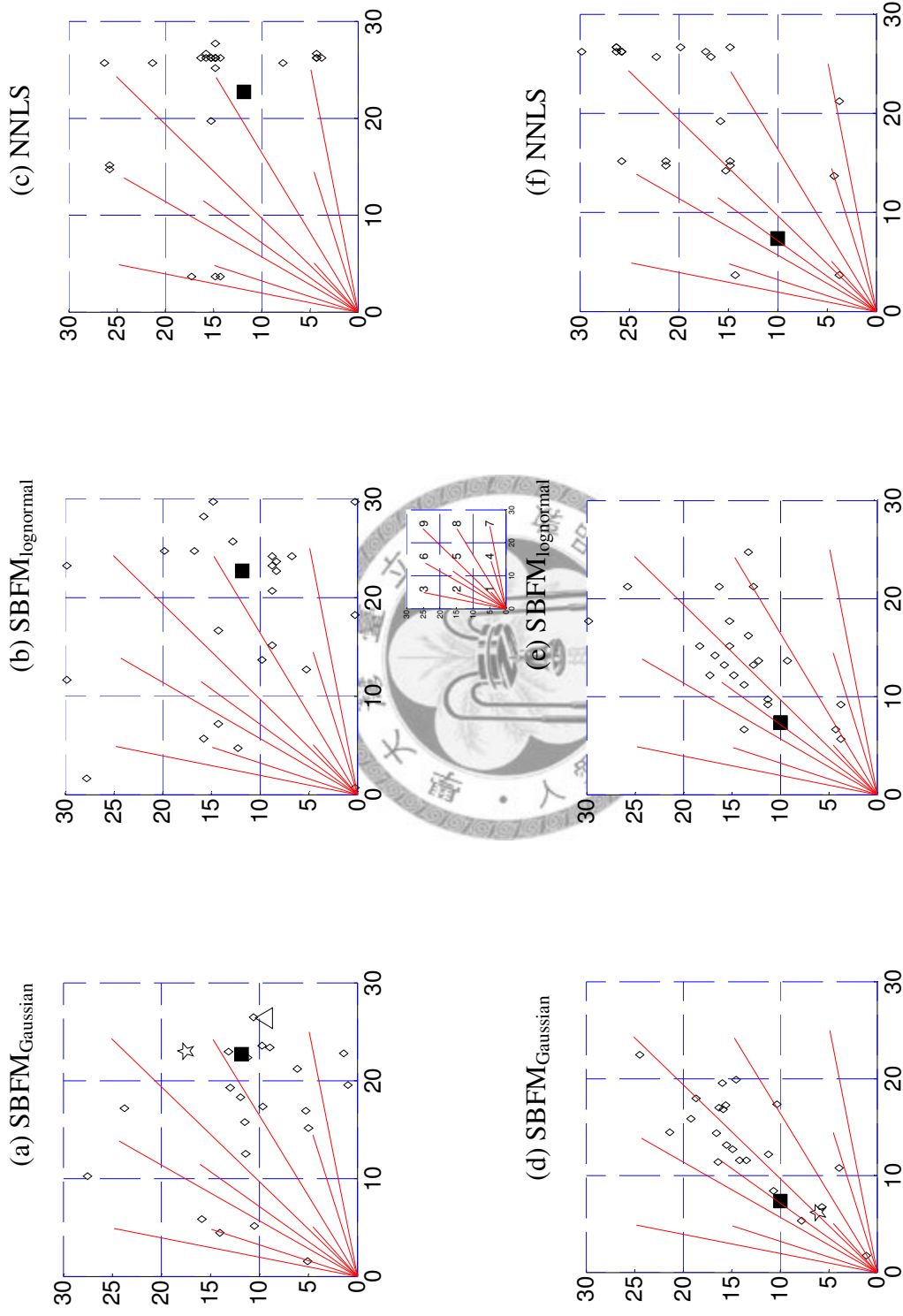


Figure 3.13 The reconstruction result of  $\text{SF}_6$  in experiment 1 (a) (b) (c) and 2 (d) (e) (f). The solid square represents the real source location. The diamonds are the reconstructed source locations by the three reconstruction algorithms. The star and triangle signs represent the cases that successfully reconstruct the source location.

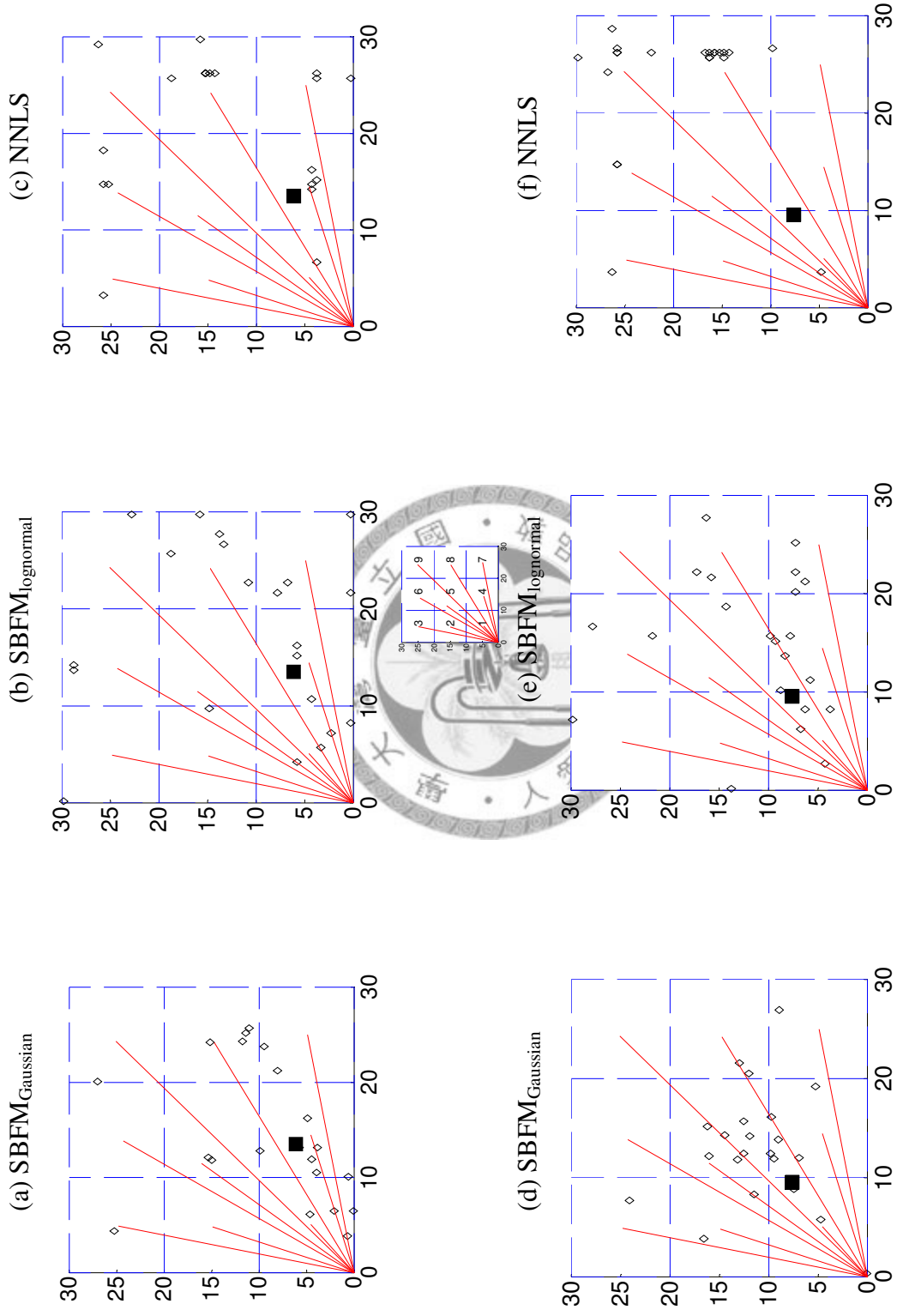


Figure 3.14 The reconstruction result of  $\text{SF}_6$  in experiment 3 (a) (b) (c) and 4 (d) (e) (f). The solid square represents the real source location. The diamonds are the reconstructed source locations by the three reconstruction algorithms.

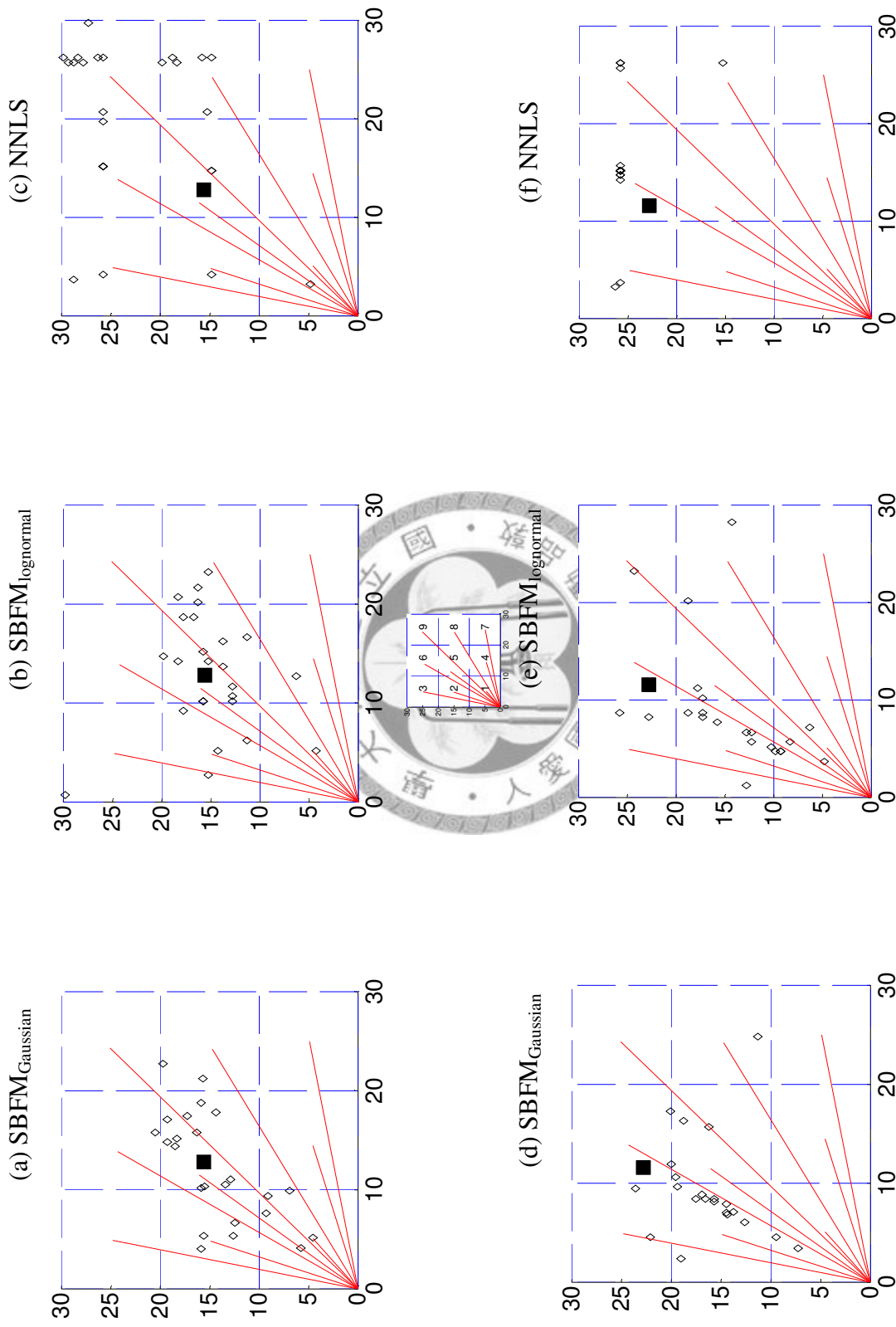


Figure 3.15 The reconstruction result of  $N_2O$  in experiment 1 (a) (b) (c) and 2 (d) (e) (f). The solid square represents the real source location. The diamonds are the reconstructed source locations by the three reconstruction algorithms.

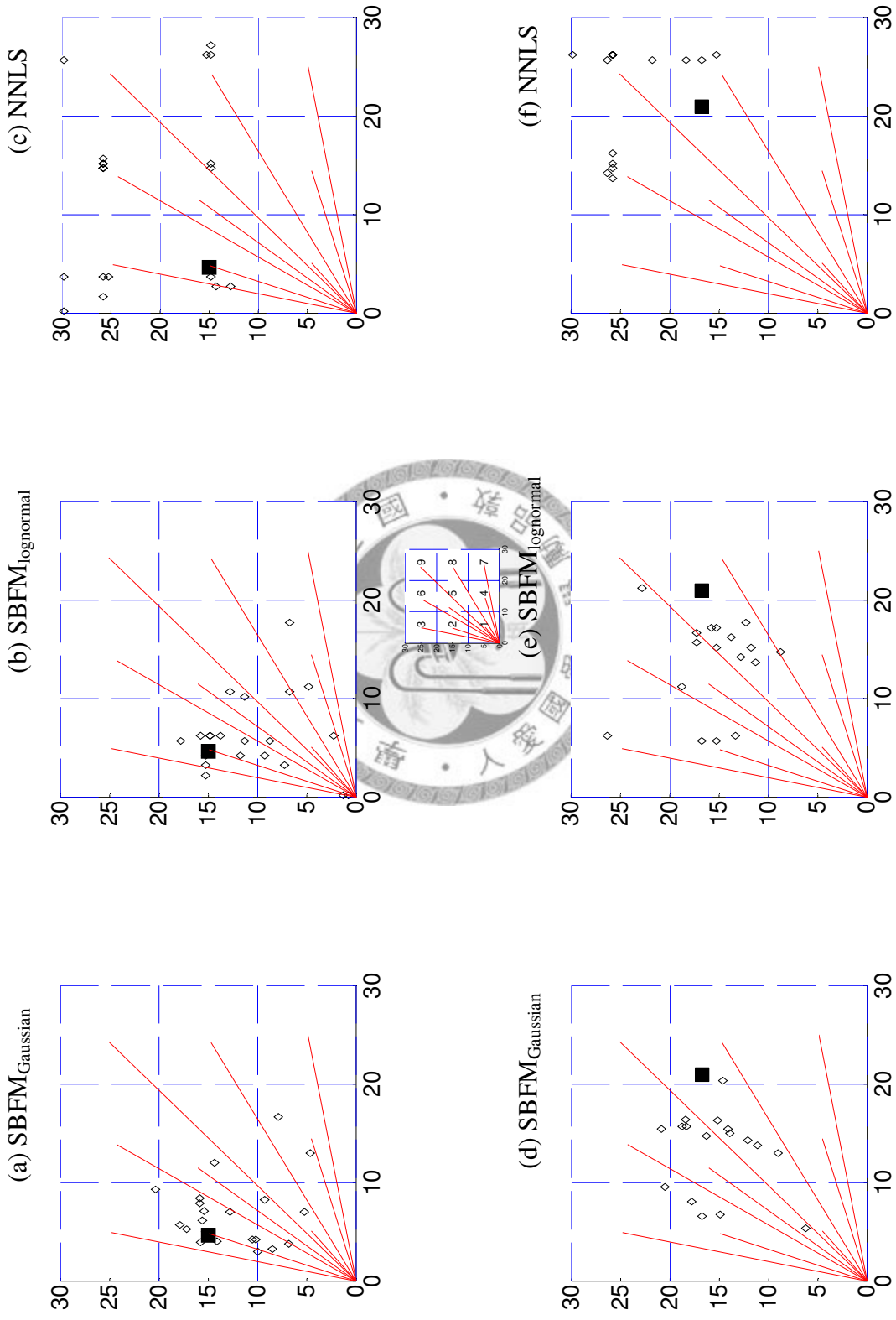


Figure 3.16 The reconstruction result of  $N_2O$  in experiment 3 (a) (b) (c) and 4 (d) (e) (f). The solid square represents the real source location. The diamonds are the reconstructed source locations by the three reconstruction algorithms.



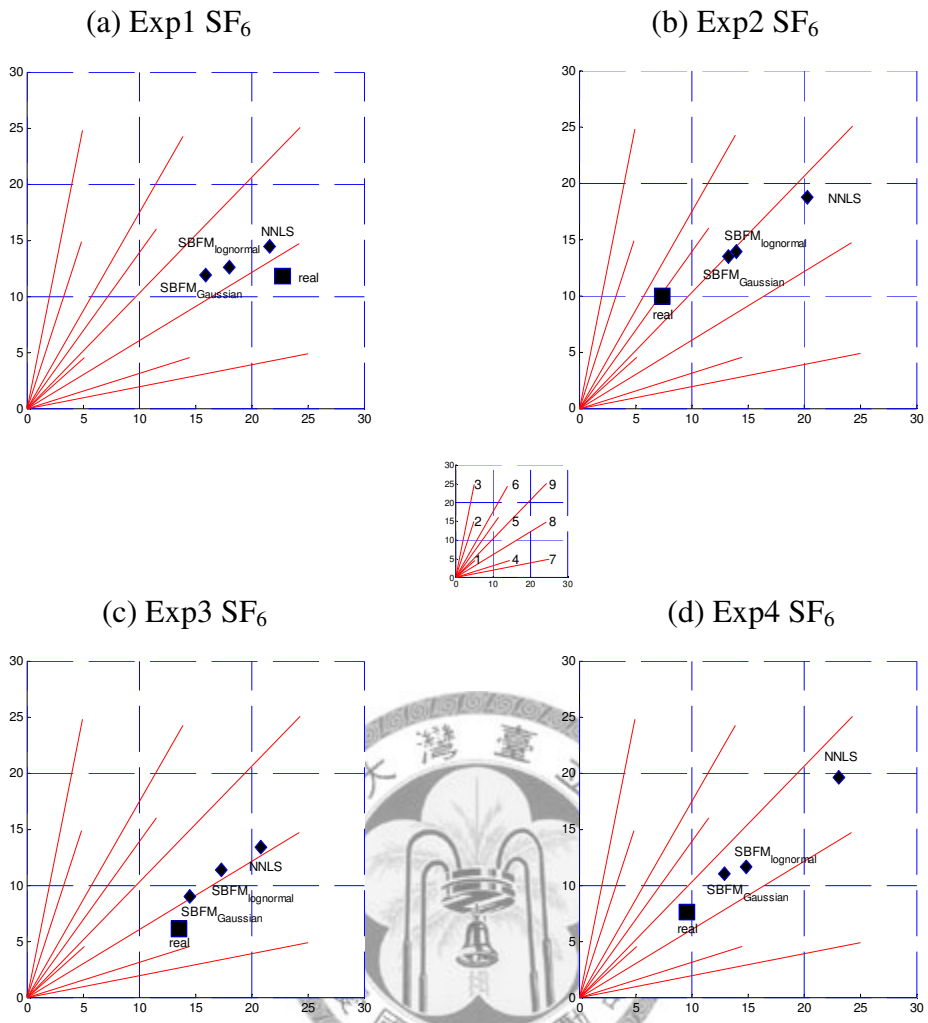


Figure 3.17 The average result of SF<sub>6</sub> in (a) Experiment1 (b) Experiment2 (c) Experiment3 (d) Experiment4. NNLS represents the average result using NNLS reconstruction. SBFM<sub>Gaussian</sub> represents the average result using SBFM reconstruction with bivariate Gaussian distribution as basis function. SBFM<sub>lognormal</sub> represents the average result using SBFM reconstruction with bivariate lognormal distribution as basis function.

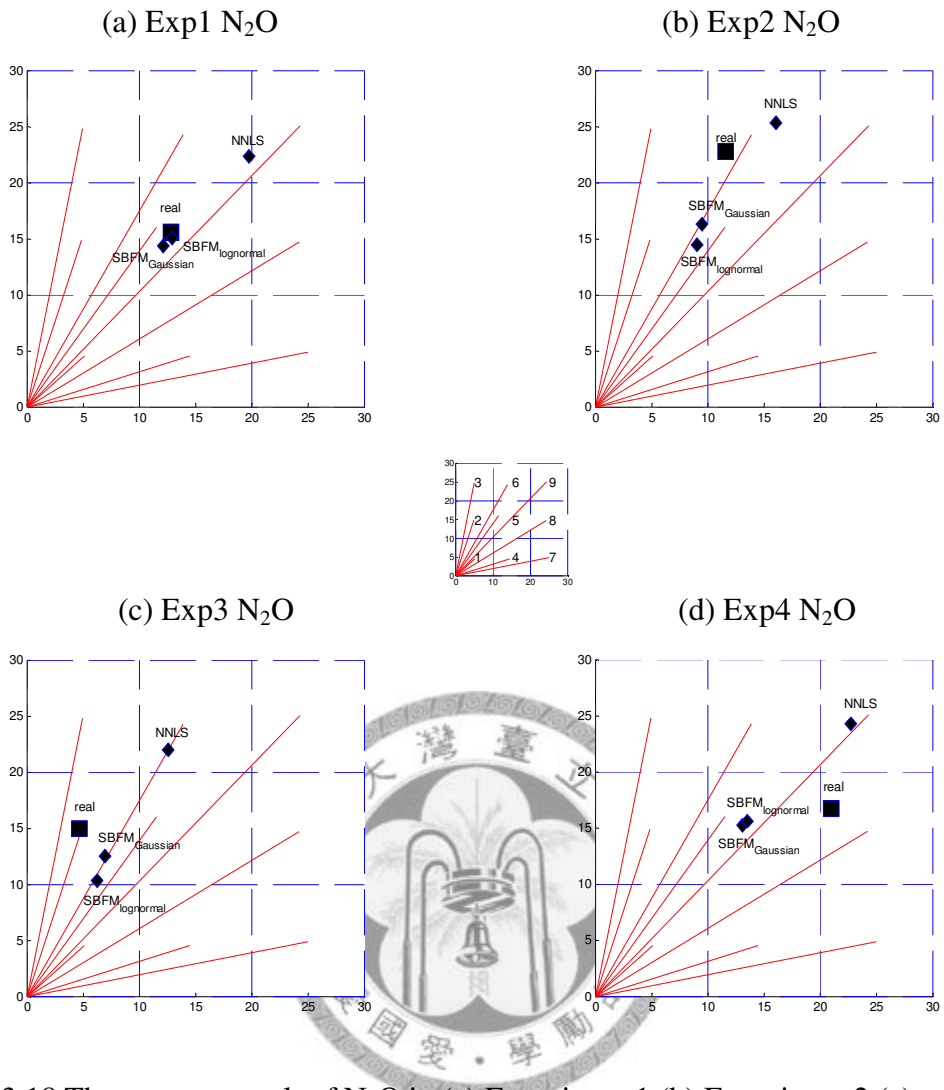
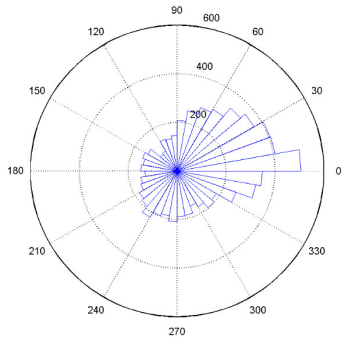
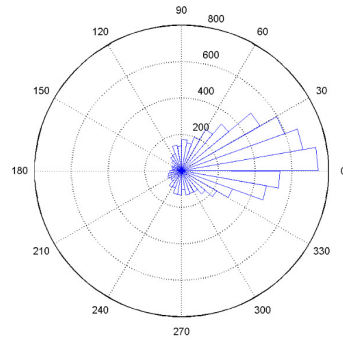


Figure 3.18 The average result of N<sub>2</sub>O in (a) Experiment1 (b) Experiment2 (c) Experiment3 (d) Experiment4. NNLS represents the average result using NNLS reconstruction. SBFM<sub>Gaussian</sub> represents the average result using SBFM reconstruction with bivariate Gaussian distribution as basis function. SBFM<sub>lognormal</sub> represents the average result using SBFM reconstruction with bivariate lognormal distribution as basis function.

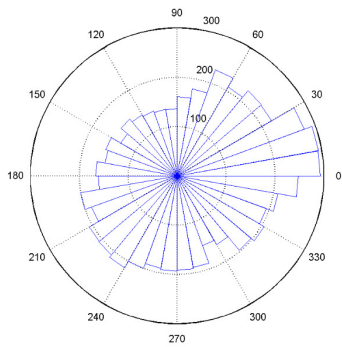
(a)



(b)



(c)



(d)

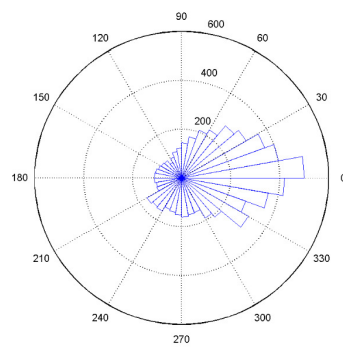


Figure 3.19 The wind rose of the four field experiments. (a) Exp1 (b) Exp2 (c) Exp3 and (d) Exp4

Table 3.1 The overall statistics of the reconstruction quality and source localization accuracy for different reconstruction methods.

Type of reconstruction	# of test distributions	Ray fit		reconstruction quality		Source localization accuracy	
		CCF <sub>PIC</sub> (Mean± std.)	CCF <sub>RPM</sub> (Mean ± std.)	Nearness <sub>RPM</sub> (Mean ± std.)	Peak error (Mean ± std.)	Hit	
RPM <sub>e_Gauss</sub>	450	0.90±0.22	0.29±0.32*	14.12±43.26*	7.94±5.94	162 (36.00%)	
RPM <sub>c_Gauss</sub>	450	0.94±0.10	0.27±0.31*	17.23±54.94*	8.02±5.61	140 (31.11%)	
RPM <sub>e_log</sub>	450	1.00±0.05	0.92±0.19*	0.27±0.74	1.53±2.82	406 (90.22%)	
RPM <sub>c_log</sub>	450	1.00±0.07	0.89±0.23*	0.32±0.66	2.04±3.18	377 (83.78%)	
RPM <sub>e&gt;NNLS</sub>	450	0.98±0.03	0.67±0.28*	0.71±0.32*	6.67±4.62	237 (52.67%)	
RPM <sub>c&gt;NNLS</sub>	450	0.95±0.06	0.56±0.30*	1.03±0.46*	7.70±5.10	197 (43.78%)	

\* represents significant difference to any other methods under paired t-test



Table 3.2 The t-test result for peak error between different reconstruction methods

	RPM <sub>e_log</sub>	RPM <sub>c_log</sub>	RPM <sub>e&gt;NNLS</sub>	RPM <sub>c&gt;NNLS</sub>	RPM <sub>e_Gauss</sub>	RPM <sub>c_Gauss</sub>
RPM <sub>e_log</sub>		○	○	○	○	○
RPM <sub>c_log</sub>	○		○	○	○	○
RPM <sub>e&gt;NNLS</sub>	○	○		○	○	○
RPM <sub>c&gt;NNLS</sub>	○	○	○		×	×
RPM <sub>e_Gauss</sub>	○	○	○	×		×
RPM <sub>c_Gauss</sub>	○	○	○	×	×	

○ represents significant difference under t-test, × represents non-significant difference under paired t-test

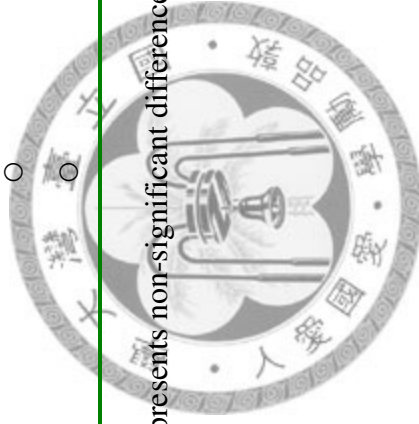


Table 3.3 The summary statistics of the reconstruction quality and source localization accuracy implemented by  $\text{RPM}_{\text{c\_Gauss}}$  classified by the pixel index.

Pixel index	# of test distributions	Ray fit		Reconstruction quality		Source localization accuracy	
		CCF <sub>PIC</sub> (Mean± std.)	CCF <sub>RPM</sub> (Mean ± std.)	Nearness <sub>RPM</sub> (Mean ± std.)	Peak error (Mean ± std.)	HIT	
1	50	0.99±0.01	0.57±0.24	1.31±0.69	3.11±1.74	26 (52.00%)	
2	50	0.94±0.15	0.50±0.33	12.81±60.83	3.44±2.88	32 (64.00%)	
3	50	0.89±0.15	0.22±0.27	15.59±28.16	10.99±5.51	4 (8.00%)	
4	50	0.94±0.12	0.34±0.31	25.73±82.34	3.26±1.49	31 (62.00%)	
5	50	0.96±0.05	0.33±0.31	9.07±25.07	3.90±2.80	43 (86.00%)	
6	50	0.90±0.09	0.12±0.26	8.92±27.12	10.96±4.35	2 (4.00%)	
7	50	0.90±0.12	0.14±0.23	13.96±24.56	11.15±3.70	1 (2.00%)	
8	50	0.97±0.06	0.14±0.20	22.76±59.79	9.85±3.94	1 (2.00%)	
9	50	0.96±0.05	0.05±0.15	44.95±97.58	15.54±3.78	0 (0.00%)	
total	450	0.94±0.10	0.27±0.31*	17.23±54.94*	8.02±5.61	140 (31.11%)	

Table 3.4 The summary statistics of the reconstruction quality and source localization accuracy implemented by  $\text{RPM}_{\text{c\_NNLS}}$  classified by the pixel index.

Pixel index	# of test distributions	Ray fit		Reconstruction quality		Source localization accuracy	
		CCF <sub>PIC</sub> (Mean± std.)	CCF <sub>RPM</sub> (Mean ± std.)	Nearness <sub>RPM</sub> (Mean ± std.)	Peak error (Mean ± std.)	at correct pixel	
1	50	0.91±0.05	0.22±0.23	1.39±0.30	15.35±6.73	1 (2.00%)	
2	50	0.94±0.08	0.51±0.23	1.15±0.38	7.85±5.54	21 (42.00%)	
3	50	0.91±0.08	0.59±0.27	0.93±0.26	5.49±2.53	21 (42.00%)	
4	50	0.94±0.03	0.30±0.25	1.52±0.46	10.98±4.26	8 (16.00%)	
5	50	0.96±0.03	0.51±0.26	1.23±0.44	8.54±3.99	10 (20.00%)	
6	50	0.97±0.05	0.79±0.15	0.67±0.29	5.27±2.57	36 (72.00%)	
7	50	0.94±0.04	0.55±0.26	0.98±0.28	5.98±1.97	18 (36.00%)	
8	50	0.96±0.04	0.75±0.15	0.85±0.39	5.29±2.45	38 (76.00%)	
9	50	0.98±0.01	0.80±0.24	0.55±0.27	4.58±2.08	44 (88.00%)	
total	450	0.95±0.06	0.56±0.30	1.03±0.46	7.70±5.10	197 (43.78%)	

Table 3.5 Sensitivity analysis of the center beam geometry

(a)  $RPM_{c\_Gauss}$

pixel	Sensitivity rate	Specificity rate	False negative rate	False positive rate
1	52.0%	94.3%	48.0%	5.8%
2	64.0%	88.0%	36.0%	12.0%
3	8.0%	99.3%	92.0%	0.7%
4	62.0%	87.0%	38.0%	13.0%
5	86.0%	54.5%	14.0%	45.5%
6	4.0%	99.8%	96.0%	0.2%
7	2.0%	100.0%	98.0%	0.0%
8	2.0%	99.8%	98.0%	0.2%
9	0.0%	100.0%	100%	0.0%

(b)  $RPM_{c\_log}$

pixel	Sensitivity rate	Specificity rate	False negative rate	False positive rate
1	100%	98.8%	0%	1.3%
2	94.0%	98.0%	6.0%	2.0%
3	86.0%	97.3%	4.0%	2.8%
4	86.0%	99.3%	4.0%	0.7%
5	86.0%	95.8%	4.0%	4.3%
6	62.0%	97.8%	38.0%	2.3%
7	86.0%	98.5%	14.0%	1.5%
8	74.0%	97.5%	26.0%	2.5%
9	80.0%	99.0%	20.0%	1.0%

(c)  $RPM_{c\_NNLS}$

pixel	Sensitivity rate	Specificity rate	False negative rate	False positive rate
1	2.0%	100.0%	98.0%	0.0%
2	42.0%	98.0%	58.0%	2.0%
3	42.0%	99.0%	58.0%	1.0%
4	16.0%	97.8%	84.0%	2.3%
5	20.0%	93.5%	80.0%	6.5%
6	72.0%	81.8%	28.0%	18.3%
7	36.0%	98.8%	64.0%	1.3%
8	76.0%	73.5%	24.0%	26.5%
9	88.0%	94.5%	12.0%	5.5%



Table 3.6 Sensitivity analysis of the extend beam geometry

(a)  $RPM_{e\_Gauss}$

pixel	Sensitivity rate	Specificity rate	False negative rate	False positive rate
1	60.0%	92.3%	40.0%	7.8%
2	84.0%	84.3%	16.0%	15.8%
3	0.0%	98.5%	100%	1.5%
4	48.0%	89.3%	52.0%	10.8%
5	76.0%	69.8%	24%	30.3%
6	16.0%	99.0%	84.0%	1.0%
7	8.0%	98.5%	92.0%	1.5%
8	32.0%	97.0%	68.0%	3.0%
9	0.0%	99.5%	100%	0.5%

(b)  $RPM_{e\_log}$

pixel	Sensitivity rate	Specificity rate	False negative rate	False positive rate
1	94.0%	99.0%	6.0%	1.0%
2	94.0%	98.3%	6.0%	1.8%
3	86.0%	98.3%	14.0%	1.8%
4	90.0%	99.3%	10.0%	0.7%
5	98.0%	98.0%	2.0%	2.0%
6	72.0%	97.8%	18.0%	2.3%
7	92.0%	99.8%	8.0%	0.2%
8	96.0%	99.0%	4.0%	1.0%
9	90.0%	99.8%	10.0%	0.2%

(c)  $RPM_{e\_NNLS}$

pixel	Sensitivity rate	Specificity rate	False negative rate	False positive rate
1	10.0%	100%	90.0%	0.0%
2	56.0%	97.8%	44.0%	2.3%
3	46.0%	99.3%	54.0%	0.7%
4	18.0%	97.3%	82.0%	2.8%
5	24.0%	94.5%	76.0%	5.5%
6	96.0%	81.8%	4.0%	18.3%
7	52.0%	98.5%	48.0%	1.5%
8	92.0%	78.8%	8.0%	21.3%
9	80.0%	99.0%	20.0%	1.0%

Table 3.7 The overall statistics of reconstruction quality and source localization accuracy classified by whether the shortest monitoring is able to detect the plume

Reconstruction type/shortestD	# of test distributions	Ray fit		Reconstruction quality		Source localization accuracy	
		CCF <sub>PIC</sub> (Mean± std.)	CCF <sub>RPM</sub> (Mean ± std.)	NearnessRPM (Mean ± std.)	Peak error (Mean ± std.)	at correct pixel	
RPM <sub>c_Gauss</sub> /Y	263	0.94±0.11	0.36±0.31*	14.66±50.61	6.38±4.60*	101 (38.40%)	
RPM <sub>c_Gauss</sub> /N	187	0.93±0.10	0.13±0.26	20.86±60.46	10.33±6.10	39 (20.86%)	
RPM <sub>e_Gauss</sub> /Y	368	0.91±0.22	0.34±0.32*	13.43±45.40	7.21±5.83*	148 (40.00%)	
RPM <sub>e_Gauss</sub> /N	82	0.88±0.22	0.07±0.21	17.21±32.00	11.20±5.40	14 (17.07%)	
RPM <sub>c_log</sub> /Y	263	1.00±0.00	0.96±0.14*	0.15±0.45*	1.17±2.04*	239 (90.87%)	
RPM <sub>c_log</sub> /N	187	0.99±0.10	0.79±0.29	0.57±0.82	3.27±4.00	138 (73.80%)	
RPM <sub>e_log</sub> /Y	368	1.00±0.00	0.94±0.17*	0.19±0.41*	1.25±2.10*	335 (91.03%)	
RPM <sub>e_log</sub> /N	82	0.99±0.11	0.83±0.23	0.63±1.45	2.81±4.69	71 (86.59%)	
RPM <sub>c_NNLS</sub> /Y	263	0.95±0.05	0.53±0.29	1.10±0.41	8.39±5.82	97 (36.88%)	
RPM <sub>c_NNLS</sub> /N	187	0.94±0.06	0.59±0.31*	0.92±0.50*	6.73±3.66*	100 (53.48%)	
RPM <sub>e_NNLS</sub> /Y	368	0.98±0.03	0.67±0.29	0.72±0.33	7.08±4.89	177 (48.10%)	
RPM <sub>e_NNLS</sub> /N	82	0.98±0.02	0.67±0.23	0.66±0.23	4.80±2.35*	60 (73.17%)	

The shortestD marked as Y represents the plume that is able to be detected by the shortest monitoring line. The shortestD marked as N represents the plume that is not able to be detected by the shortest monitoring line. The \* sign represents the significantly better performance

Table 3.8 The regression analysis result using dPIC, maxCON and shortestD as predictor variables.

Method/indicator	R <sup>2</sup>	dPIC	maxCON										shortestD	
			1	2	3	4	5	6	7	8	9	Y	N	
<b>RPM<sub>e_Gauss</sub></b>														
CCFRPM	0.32	0.08	0.05	0.40	0.13	0.28	0.22	-0.03	0.21	0.38	-	-	-0.06	-
NearnessRPM	0.10	-1.95	5.73	-15.93	-15.53	-15.65	-15.60	18.86	-6.12	-15.62	-	-	8.23	-
Peak error	0.36	-1.56	-1.10	-7.91	-5.08	-6.54	-8.19	-0.20	-3.91	-7.12	-	-	2.53	-
<b>RPM<sub>c_Gauss</sub></b>														
CCFRPM	0.38	0.03	-0.07	0.37	0.28	0.32	-0.04	0.44	0.10	0.37	-	-	0.06	-
NearnessRPM	0.05	3.47	21.77	-15.61	-14.39	-15.09	6.43	-14.77	7.78	-14.62	-	-	-13.54	-
Peak error	0.41	-0.98	3.42	-4.79	-3.04	-4.85	2.56	-5.59	-1.57	-4.81	-	-	0.56	-
<b>RPM<sub>e_log</sub></b>														
CCFRPM	0.14	0.05	-0.04	-0.12	-0.02	-0.03	-0.03	0.03	0.00	0.02	-	-	-0.04	-
NearnessRPM	0.11	-0.10	0.46	0.44	0.05	0.02	0.08	-0.03	0.00	-0.03	-	-	-0.16	-
Peak error	0.12	-0.52	0.08	-1.62	-0.57	-1.77	-1.72	-0.95	-1.31	-1.79	-	-	0.34	-
<b>RPM<sub>c_log</sub></b>														
CCFRPM	0.29	0.07	-0.08	-0.05	0.04	-0.11	0.05	0.02	-0.02	0.02	-	-	-0.03	-
NearnessRPM	0.23	-0.14	0.65	0.13	-0.02	0.18	-0.08	-0.05	0.08	-0.05	-	-	-0.03	-
Peak error	0.19	-0.61	-0.25	-2.22	-1.53	-1.99	-1.43	-2.44	-1.14	-2.44	-	-	-0.11	-
<b>RPM<sub>e&gt;NNLS</sub></b>														
CCFRPM	0.37	0.06	0.09	-0.22	-0.04	-0.46	-0.30	0.19	-0.07	-0.12	-	-	-0.1	-
NearnessRPM	0.42	-0.04	-0.08	0.21	-0.03	0.38	0.26	-0.36	0.03	0.15	-	-	0.09	-
Peak error	0.31	0.01	-0.97	3.68	2.33	8.60	5.18	-0.72	2.74	0.70	-	-	1	-
<b>RPM<sub>c&gt;NNLS</sub></b>														
CCFRPM	0.44	0.08	0.10	-0.16	-0.07	-0.47	0.17	-0.36	-0.15	-0.12	-	-	-0.2	-
NearnessRPM	0.32	-0.04	-0.05	0.24	0.04	0.45	-0.26	0.25	0.31	-0.01	-	-	0.19	-
Peak error	0.47	0.21	-1.62	-0.43	2.30	10.73	0.04	12.41	3.56	0.89	-	-	-0.43	-

The values in the columns of dPIC, maxCON and shortestD are the parameters of each variable.

Table 3.9 ANOVA analysis between different dPIC with test distributions without ShortestD

Beam Geometry	P-value	Ranking of dPIC
Geometry <sub>extend</sub>	0.19	6→5→4→3→7→2
Geometry <sub>center</sub>	0.01	8→6→7→4→5→3→2→1

The ranking represents the result of Bonferroni t-test from the best performance (shortest peak error) to the worst (longest peak error).

Table 3.10 The detection limit of SF<sub>6</sub> and N<sub>2</sub>O in each monitoring line

	Path1	Path2	Path3	Path4	Path5	Path6	Path7	Path8	Path9
SF <sub>6</sub>	0.069	0.088	0.075	0.118	0.075	0.192	0.064	0.060	0.107
N <sub>2</sub> O	0.643	0.847	0.416	0.659	0.397	1.037	0.518	0.458	0.618



Table 3.11 The summary statistics of PIC data in each experiment (a) SF<sub>6</sub> (b) N<sub>2</sub>O

(a) SF <sub>6</sub>									
Experiment	Path1 (Mean±SD)	Path2 (Mean±SD)	Path3 (Mean±SD)	Path4 (Mean±SD)	Path5 (Mean±SD)	Path6 (Mean±SD)	Path7 (Mean±SD)	Path8 (Mean±SD)	Path9 (Mean±SD)
1	0.41±1.13	0.43±1.18	0.68±1.10	0.51±1.12	2.11±3.12	0.09±0.26	4.06±5.56	0.30±1.06	1.83±3.23
2	0.81±1.54	0.51±1.27	2.36±3.50	4.13±5.33	8.90±7.71	0.48±1.25	2.77±3.17	0.67±2.14	0.64±1.55
3	0.22±0.40	0.11±0.30	0.47±0.90	0.26±0.50	0.86±1.14	0.13±0.25	1.61±1.72	1.11±1.87	1.15±2.01
4	0.78±1.11	0.44±0.86	1.28±1.81	1.26±1.56	4.04±3.46	0.43±1.04	3.43±2.73	0.92±1.18	1.33±1.19
total	0.55±1.14	0.38±0.99	1.20±2.21	1.56±3.25	4.02±5.49	0.28±0.84	2.99±3.72	0.73±1.64	1.24±2.19
(b) N <sub>2</sub> O									
Experiment	Path1 (Mean±SD)	Path2 (Mean±SD)	Path3 (Mean±SD)	Path4 (Mean±SD)	Path5 (Mean±SD)	Path6 (Mean±SD)	Path7 (Mean±SD)	Path8 (Mean±SD)	Path9 (Mean±SD)
1	2.06±4.62	0.94±3.15	4.78±6.98	6.99±10.55	12.07±10.51	0.67±2.30	4.63±5.16	0.59±1.82	0.78±1.47
2	3.90±9.73	0.45±1.33	71.71±80.99	2.50±6.08	18.24±31.14	0.37±1.93	5.33±15.82	1.66±6.22	2.18±6.50
3	14.93±21.66	17.11±28.71	26.66±34.94	8.68±10.49	10.75±19.90	0.82±2.61	6.27±10.73	4.15±10.07	3.44±6.50
4	1.74±3.44	0.95±2.38	4.26±8.25	0.78±1.29	14.28±16.64	0.08±0.37	6.04±9.60	0.71±2.63	0.97±2.54
total	5.56±13.04	4.70±15.67	27.30±52.41	4.80±8.64	13.86±20.88	0.49±2.01	5.53±10.86	1.75±6.15	1.83±4.86

Table 3.12 The reconstruction result of field experiment of two tracer gases (a) SF<sub>6</sub> (b) N<sub>2</sub>O

(a) SF<sub>6</sub>

	Valid PIC set	CCF <sub>PIC</sub> (Mean±SD)	Peak error (Mean±SD)	HIT
Exp1 (pixel 8)				
SBFM <sub>Gaussian</sub>	24	0.96±0.09	9.00±6.38	4(16.67%)
SBFM <sub>lognormal</sub>	24	0.89±0.27	10.55±7.35	6(25.00%)
NNLS	24	0.92±0.21	8.88±5.41	11(45.83%)
Exp2 (pixel 1)				
SBFM <sub>Gaussian</sub>	22	0.96±0.07	9.15±4.51*	4(17.39%)
SBFM <sub>lognormal</sub>	22	0.94±0.11	9.48±5.40	3(13.04%)
NNLS	22	0.80±0.29	17.44±7.02 <sup>#</sup>	1(4.35%)
Exp3 (pixel 4)				
SBFM <sub>Gaussian</sub>	21	0.93±0.14	9.07±5.72*	7(33.33%)
SBFM <sub>lognormal</sub>	21	0.83±0.31	12.79±7.29	3(14.29%)
NNLS	21	0.83±0.26	13.94±6.80	4(19.05%)
Exp4 (pixel 1)				
SBFM <sub>Gaussian</sub>	22	0.87±0.16	7.93±4.43*	3(13.64%)
SBFM <sub>lognormal</sub>	22	0.86±0.16	10.28±6.43	4(18.18%)
NNLS	22	0.87±0.25	19.93±4.41 <sup>#</sup>	1(4.55%)

(b) N<sub>2</sub>O

	Valid PIC set	CCF <sub>PIC</sub> (Mean±SD)	Peak error (Mean±SD)	HIT
Exp1 (pixel 5)				
SBFM <sub>Gaussian</sub>	25	0.95±0.10	6.44±3.12	12(50.00%)
SBFM <sub>lognormal</sub>	25	0.93±0.13	6.31±4.06	14(56.00%)
NNLS	25	0.82±0.26	13.46±4.83 <sup>#</sup>	2(8.00%)
Exp2 (pixel 6)				
SBFM <sub>Gaussian</sub>	21	0.97±0.07	8.08±4.46	1(4.35%)
SBFM <sub>lognormal</sub>	21	0.95±0.13	11.10±4.99	0(0.00%)
NNLS	21	0.99±0.01	6.93±4.18*	16(72.73%)
Exp3 (pixel 2)				
SBFM <sub>Gaussian</sub>	21	0.93±0.11	5.26±3.78	12(57.14%)
SBFM <sub>lognormal</sub>	21	0.87±0.18	6.39±4.99	10(47.62%)
NNLS	21	0.77±0.26	13.61±6.43 <sup>#</sup>	3(14.29%)
Exp4 (pixel 8)				
SBFM <sub>Gaussian</sub>	17	0.99±0.02	8.82±4.43	1(0.06%)
SBFM <sub>lognormal</sub>	17	0.99±0.02	8.74±4.50	0(0.00%)
NNLS	17	0.98±0.04	9.70±2.59	3(17.65%)

The \* represents significantly better than the other two methods.

The <sup>#</sup> represents significantly worse than the other two methods.

Table 3.13 The mean reconstruction result of field experiment of two tracer gases (a) SF<sub>6</sub> (b) N<sub>2</sub>O

(a) SF<sub>6</sub>

	Valid PIC set	Peak error	HIT
Exp1 (pixel 8)			
SBFM <sub>Gaussian</sub>	24	6.81	No
SBFM <sub>lognormal</sub>	24	4.79	No
NNLS	24	2.90	Yes
Exp2 (pixel 1)			
SBFM <sub>Gaussian</sub>	22	6.84	No
SBFM <sub>lognormal</sub>	22	7.70	No
NNLS	22	15.61	No
Exp3 (pixel 4)			
SBFM <sub>Gaussian</sub>	21	3.09	Yes
SBFM <sub>lognormal</sub>	21	6.52	No
NNLS	21	10.34	No
Exp4 (pixel 1)			
SBFM <sub>Gaussian</sub>	22	4.74	No
SBFM <sub>lognormal</sub>	22	6.56	No
NNLS	22	18.03	No

(b) N<sub>2</sub>O

	Valid PIC set	Peak error	HIT
Exp1 (pixel 5)			
SBFM <sub>Gaussian</sub>	25	1.38	Yes
SBFM <sub>lognormal</sub>	25	0.49	Yes
NNLS	25	9.70	No
Exp2 (pixel 6)			
SBFM <sub>Gaussian</sub>	21	6.81	No
SBFM <sub>lognormal</sub>	21	8.68	No
NNLS	21	5.14	Yes
Exp3 (pixel 2)			
SBFM <sub>Gaussian</sub>	21	3.33	Yes
SBFM <sub>lognormal</sub>	21	4.86	Yes
NNLS	21	10.59	No
Exp4 (pixel 8)			
SBFM <sub>Gaussian</sub>	17	8.03	No
SBFM <sub>lognormal</sub>	17	7.53	No
NNLS	17	7.74	No

Table 3.14 The chosen reconstructed result by the nearby beam paths

	Nearby paths	Nearby short paths	result
Exp1			
SF <sub>6</sub>	5, 6, 7	4, 6, 8	NNLS
N <sub>2</sub> O	4, 5	4*, 6	SBFM <sub>Gaussian</sub>
Exp2			
SF <sub>6</sub>	5, 6, 7	4*, 6, 8*	SBFM <sub>Gaussian</sub>
N <sub>2</sub> O	3, 4	2, 4	NNLS
Exp3			
SF <sub>6</sub>	5, 7	4, 6, 8*	SBFM <sub>Gaussian</sub>
N <sub>2</sub> O	2, 3, 4	2*, 4*	SBFM <sub>Gaussian</sub>
Exp4			
SF <sub>6</sub>	5, 6, 7	4*, 6, 8*	SBFM <sub>Gaussian</sub>
N <sub>2</sub> O	4, 5	4, 6	NNLS

The \* represents the monitoring line that detects the plume





Table 3.15 The wind data of the field experiment

	Number of data sets	Wind speed	Wind direction	
		Mean±SD	Mean±SD	Mode
Exp1	7865	0.43±0.41*	164.63±118.72*	0
Exp2	7383	0.46±0.38*	151.52±133.46*	359.1
Exp3	6900	0.86±0.70*	175.82±110.59	0
Exp4	6960	0.83±0.63*	175.02±125.62	356.3

The \* sign represents significant difference comparing to other experiments under t-test. There is no significant difference between the wind direction of Exp3 and Exp4.



## Chapter4. Conclusions and suggestions

### 4.1 The simulation experiment

In the simulation study, a series of test distributions are generated and be reconstructed by the three reconstruction algorithms used in the RPM technique. The result shows that the SBFM using bivariate lognormal distribution gives the best result than the other methods. From the overall results (Table 3.1), in the aspect of the plume reconstruction, the SBFM using bivariate lognormal distribution as basis function may give the best result followed by NNLS and SBFM reconstruction using bivariate Gaussian distribution as basis function. A slightly different result is observed in the aspect of the source localization ability. The SBFM using bivariate lognormal distribution still gives the best result while the SBFM reconstruction using bivariate Gaussian distribution as basis function may gives a comparable result to that by the NNLS reconstruction. If we further look at the performance of each reconstruction algorithms in each pixel the NNLS reconstruction is not able to localize the source in the proximity of OP-FTIR while the SBFM using bivariate Gaussian as basis function is not able to localize the source that is far from the OP-FTIR. The reasons are discussed in Chapter 3.

Another simulation study is conducted to investigate the uncertainty for the source localization ability for the SBFM algorithm. The larger the plume yields the smaller

uncertainty when applying the SBFM reconstruction. Furthermore, the source location in the proximity of the OP-FTIR yields smaller uncertainty. The source location far from the FTIR may still be localized if multiple monitoring lines are able to detect the plume.

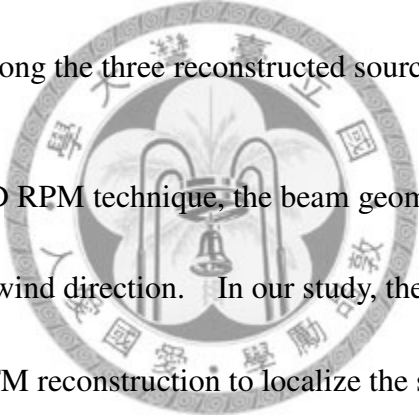
#### ***4.2 The field study***

In the field study, four experiments are conducted with four pairs of artificially released emission sources. The results shows that the NNLS reconstruction is able to localize the source that is far from the OP-FTIR while the SBFM reconstruction using bivariate Gaussian as basis function is able to localize the source that is in the proximity of the OP-FTIR. With adequate segmenting information, the SBFM using bivariate Gaussian distribution as basis function may still localize the emission source that is far from the OP-FTIR. Furthermore, the reconstructed sources by the three methods are able to point out the correct direction of the real source location. Judging by the measured PIC data, one of the reconstructed source locations among the three methods can be chosen as the real source location. If the peripheral short paths are able to detect the plume, the real source location might be close to the reconstructed source by the SBFM using Gaussian distribution as basis function.

#### ***4.3 Suggestions***

In the OTM-10 by US EPA, this guideline suggests that when using the HRPM, the

NNLS algorithm, rather than the SBFM algorithm, is used as the reconstruction algorithm to localize the “Hot spot” of the plume. However in our study, we have found out that the NNLS reconstruction may only reconstruct the plume that is far from the OP-FTIR. Thus we may suggest that when using the HRPm described in the OTM-10, not only the NNLS reconstruction should be implemented but also the SBFM reconstruction is needed. The reconstructed source locations may be further screened by the PIC that is detected by peripheral short paths. The 1-D RPM may also be applied after the 2-D reconstruction is implemented. This is to further verify the possible source location among the three reconstructed source locations.



When applying the 2-D RPM technique, the beam geometry must be set up by previously considering the wind direction. In our study, the segmenting information is needed when using the SBFM reconstruction to localize the source. If the wind direction is due the direction towards the OP-FTIR, all of the monitoring lines will be able to detect the plume. Under this circumstance, the SBFM algorithm is likely not able to localize the source. Thus before setting up the beam geometry, the wind direction must be first considered to be perpendicular to the monitoring lines to prevent the situation that described above.

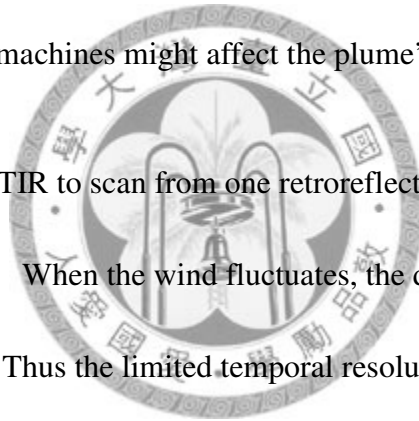
When applying the SBFM reconstruction, a more flexible bivariate distribution might be used as the basis function. Although the lognormal distribution used in our

study is a skewed distribution, it may still not be flexible enough to characterize the plume because it is only able to be skewed to certain directions. For example, the bivariate skewed distribution<sup>46</sup>.

#### ***4.4 Limitations***

The experiment is conducted on a lawn without obstacles. However with the obstacles' present, the shape of the plume may be affected. This will limit the application of the SBFM reconstruction. For example, for the application in the factory, the manufacturing machines might affect the plume's distribution.

The time for the OP-FTIR to scan from one retroreflector to the other may cause error to the reconstruction. When the wind fluctuates, the distribution of the plume also fluctuates with time. Thus the limited temporal resolution of the PIC data may come from the different distributions. This will cause error when applying the reconstruction algorithm.



## References

1. Smith, B.C., *Fundamentals of Fourier Transform Infrared Spectroscopy*. Boca Raton: CRC Press, 1996.
2. Kristine, C. and J.D. Tate, *Monitor ambient air with optical sensing system*. Chemical Engineering, 1997. **104**(6): p. 110.
3. Freddie, E.H., Jr., *Case study: Environmental monitoring using remote optical sensing (OP-FTIR) technology at the Oklahoma City Air Logistics Center industrial wastewater treatment facility*. Federal Facilities Environmental Journal, 2004. **15**(1): p. 21-37.
4. Michel, G. and F.-J. Edgar, *Air pollution monitoring with two optical remote sensing techniques in Mexico City*. 2004: SPIE.
5. Lin, C., L. Naiwei, C. Pao-Erh, Y. Jen-Chin, and S. Endy, *Fugitive coke oven gas emission profile by continuous line averaged open-path Fourier transform infrared monitoring*. Journal of the Air and Waste Management Association (1995), 2007. **57**(4): p. 472-9.
6. Wu, C.F., M.G. Yost, J. Varr, and R.A. Hashmonay, *Applying open-path FTIR with a bi-beam strategy to evaluate personal exposure in indoor environments: experimental results of a validation study*. AIHA J (Fairfax, Va), 2003. **64**(2): p. 181-8.
7. Wu, C.F., M.G. Yost, R.A. Hashmonay, and T.V. Larson, *Applying open-path FTIR with computed tomography to evaluate personal exposures. Part 2: experimental studies*. Ann Occup Hyg, 2005. **49**(1): p. 73-83.
8. Chan, C.-C., S. Ruei-Hao, C. Ta-Yuan, and T. Dai-Hua, *Workers' exposures and potential health risks to air toxics in a petrochemical complex assessed by improved methodology*. International Archives of Occupational and Environmental Health, 2006. **79**(2): p. 135-42.
9. Hashmonay, R.A., D.F. Natschke, K. Wagoner, D.B. Harris, E.L. Thompson, and M.G. Yost, *Field evaluation of a method for estimating gaseous fluxes from area sources using open-path Fourier transform infrared*. Environmental Science & Technology, 2001. **35**(11): p. 2309-13.
10. Todd, L.A., M. Ramanathan, K. Mottus, R. Katz, A. Dodson, and G. Mihlan, *Measuring chemical emissions using open-path Fourier transform infrared (OP-FTIR) spectroscopy and computer-assisted tomography*. Atmospheric

- Environment, 2001. **35**(11): p. 1937-1947.
11. Todd, L. and D. Leith, *Remote Sensing and Computed Tomography in Industrial Hygiene*. American Industrial Hygiene Association Journal, 1990. **51**(4): p. 224-233.
  12. Todd, L. and G. Ramachandran, *Evaluation of algorithms for tomographic reconstruction of chemical concentrations in indoor air*. American Industrial Hygiene Association Journal (AIHAJ), 1994. **55**(5): p. 403-17.
  13. Yost, M.G., A.J. Gadgil, A.C. Drescher, Y. Zhou, M.A. Simonds, and S.P. Levine, *Imaging indoor tracer-gas concentrations with computed tomography: experimental results with a remote sensing FTIR system*. American Industrial Hygiene Association Journal (AIHAJ), 1994. **55**(5): p. 395-402.
  14. Park, D.Y. and M.G. Yost, *Evaluation of virtual source beam configurations for rapid tomographic reconstruction of gas and*. Journal of the Air & Waste Management Association (1995), 1997. **47**(5): p. 582.
  15. Hashmonay, R.A. and M.G. Yost, *Localizing gaseous fugitive emission sources by combining real-time optical remote sensing and wind data*. J Air Waste Manag Assoc, 1999. **49**(11): p. 1374-9.
  16. Hashmonay, R.A., M.G. Yost, and C.F. Wu, *Computed tomography of air pollutants using radial scanning path-integrated optical remote sensing*. Atmospheric Environment, 1999. **33**: p. 267-274.
  17. Piper, A.R., L.A. Todd, and K. Mottus, *A field study using open-path FTIR spectroscopy to measure and map air emissions from volume sources*. Field Analytical Chemistry & Technology, 1999. **3**(2): p. 69.
  18. Wu, C.F., M.G. Yost, A.H. R, and D.Y. Park, *Experimental evaluation of a radial beam geometry for mapping air pollutants using optical remote sensing and computed tomography*. Atmospheric Environment, 1999. **33**: p. 4709-4716.
  19. Hashmonay, R.A., K. Wagoner, D.F. Natschke, D.B. Harris, and E. Thompson, *Radial Computed Tomography Of Air Contaminants Using Optical Remote sensing*. The Air & Waste Management Association's 95th Annual Conference & Exhibition, 2002.
  20. *Compendium Method TO-16: Long-Path Open-Path Fourier Transform Infrared Monitoring of Atmospheric Gases, 2nd edition*. . U.S EPA, 1999.
  21. Chen, C.-L., F. Hung Yuan, and S. Chi-Min, *Source location and*

- characterization of volatile organic compound emissions at a petrochemical plant in Kaohsiung, Taiwan.* Journal of the Air and Waste Management Association (1995), 2005. **55**(10): p. 1487-97.
22. Byer, R.L., *Two-dimensional remote air-pollution monitoring via tomography.* Optics Letters, 1979. **4**: p. 75.
  23. Todd, L.A. and R. Bhattacharyya, *Tomographic reconstruction of air pollutants: evaluation of measurement geometries.* Applied Optics-OT, 1997. **36**(30): p. 7678.
  24. Verkruyse, W., *Improved method "grid translation" for mapping environmental pollutants using a two-dimensional CAT scanning system.* Atmospheric Environment, 2004. **38**(12): p. 1801.
  25. Verkruyse, W. and A.T. Lori, *Novel algorithm for tomographic reconstruction of atmospheric chemicals with sparse sampling.* Environmental Science & Technology, 2005. **39**(7): p. 2247-54.
  26. Samanta, A. and L.A. Todd, *Mapping chemicals in air using an environmental CAT scanning system: evaluation of algorithms.* Atmospheric Environment, 2000. **34**(5): p. 699-709.
  27. Gordon, R., R. Bender, and G.T. Herman, *Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography.* Journal of Theoretical Biology, 1970. **29**(3): p. 471-81.
  28. Tsui, B.M.W., *Comparison between ML-EM and WLS-CG algorithms for SPECT image reconstruction.* IEEE Transactions on Nuclear Science, 1991. **38**(6): p. 1766.
  29. Reis, M.L., *Maximum entropy algorithms for image reconstruction from projections.* Inverse Problems, 1992. **8**(4): p. 623.
  30. Drescher, A.C., A.J. Gadgil, P.N. Price, and W.W. Nazaroff, *Novel approach for tomographic reconstruction of gas concentration distributions in air: Use of smooth basis functions and simulated annealing.* Atmospheric Environment, 1996. **30**(6): p. 929-940.
  31. Tsai, M.Y., M.G. Yost, C.F. Wu, R.A. Hashmonay, and T.V. Larson, *Line profile reconstruction: validation and comparison of reconstruction methods.* Atmospheric Environment, 2001. **35**: p. 4791-4799.
  32. 陳靜慧, *運用開徑式傅立葉轉換紅外光光譜儀定位大氣環境中之逸散源.*



- 台灣大學, 2008.
33. Price, P.N., *Pollutant tomography using integrated concentration data from non-intersecting optical paths*. Atmospheric Environment, 1999. **33**(2): p. 275.
  34. Ram, A.H., G.Y. Michael, D.B. Harris, and Edgar L. Thompson, Jr. *Simulation study for gaseous fluxes from an area source using computed tomography and optical remote sensing*. 1999: SPIE.
  35. *Other Test Method OTM10: Optical Remote Sensing for Emission Characterization from Non-Point Sources*. U.S EPA, 2006.
  36. Hashmonay, R.A., R.M. Varma, M.T. Modrak, R.H. Kagann, R.R. Segall, and P.D. Sullivan, *Radial Plume Mapping: A US EPA Test Method for Area and Fugitive Source Emission Monitoring Using Optical Remote Sensing*. Advanced Environmental Monitoring, 2007. **Springer**: p. 21-36.
  37. Hashmonay, R.A., *Theoretical evaluation of a method for locating gaseous emission hot spots*. Journal of the Air and Waste Management Association (1995), 2008. **58**(8): p. 1100-6.
  38. Aitchison J, B.J., *The Lognormal Distribution*. 1957.
  39. Lin, L.I., *A concordance correlation coefficient to evaluate reproducibility*. Biometrics, 1989. **45**(1): p. 255-68.
  40. Herman, G.T., A. Lent, and S.W. Rowland, *ART: mathematics and applications. A report on the mathematical foundations and on the applicability to real data of the algebraic reconstruction techniques*. Journal of Theoretical Biology, 1973. **42**(1): p. 1-32.
  41. Lawson, C.L. and R.J. Janson, *Solving Least Squares Problems*. Society for Industrial and Applied Mathematics: Philadelphia, 1995. **Chapter 23**: p. 158-165.
  42. Nielsen, H.B., S.N. Lophaven, and J. Søndergaard, *The Kriging Toolbox for MATLAB*. <http://www2.imm.dtu.dk/~hbn/dace/>, 2002.
  43. Stockwell, W.R., P. Middleton, J.S. Chang, and X. Tang, *The second generation regional acid deposition model chemical mechanism for regional air quality modeling*. Journal of Geophysical Research ; Vol/Issue: 95:D10, 1990: p. Pages: 16,343-16,367.
  44. Tsoukala, V.K. and C.I. Moutzouris, *Gas transfer under breaking waves: experiments and an improved vorticity-based model*. Annales Geophysicae,

2008. **26**(8): p. 2131-2142.
45. Misra, P.K. and A. Chtcherbakov, *Study of atmospheric dispersion of pollutant plumes from elevated stacks assuming a finite limit to the rate of vertical dispersion*. Atmospheric Environment, 2008. **42**(19): p. 4601-4610.
46. Gupta, A.K., G. Gonzalez-Farias, and J.A. Dominguez-Molina, *A multivariate skew normal distribution*. Journal of Multivariate Analysis, 2004. **89**(1): p. 181-190.



# Appendix

Appendix 1. The standard operation procedure of the quality assurance.

## A. The procedure of determining the method noise and result

1. Collect two spectra simultaneously in each beam path. Do not allow any time to elapse between these two spectra.
2. Generate an absorbance spectrum with either of these two spectra as background.
3. Analyze this absorbance spectrum for the RMS deviation in the three wave number regions  $968\text{--}1008\text{ cm}^{-1}$ ,  $2480\text{--}2520\text{ cm}^{-1}$ , and  $4380\text{--}4420\text{ cm}^{-1}$ .

	Path1	Path2	Path3	Path4	Path5	Path6	Path7	Path8	Path9
Region 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5	7	4	5	6	8	5	5	5
Region 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	6	4	4	4	7	3	3	4
Region 3	0.000	0.001	0.001	0.000	0.001	0.001	0.001	0.000	0.000
	9	3	1	8	3	8	0	7	7

## B. The procedure of remaining proper beam intensity.

1. Display the single beam spectrum when alignment. Visually examine the spectral region of  $650\text{ cm}^{-1}$ .
2. If there is a dip in this region, move the telescope of the OP-FTIR away slightly from the current position until the dip disappears.
3. Once the dip disappears, record the current coordinate of the scanner and the current beam intensity.
4. Move on to the next retroreflector.

## C. The procedure for determine detection limit

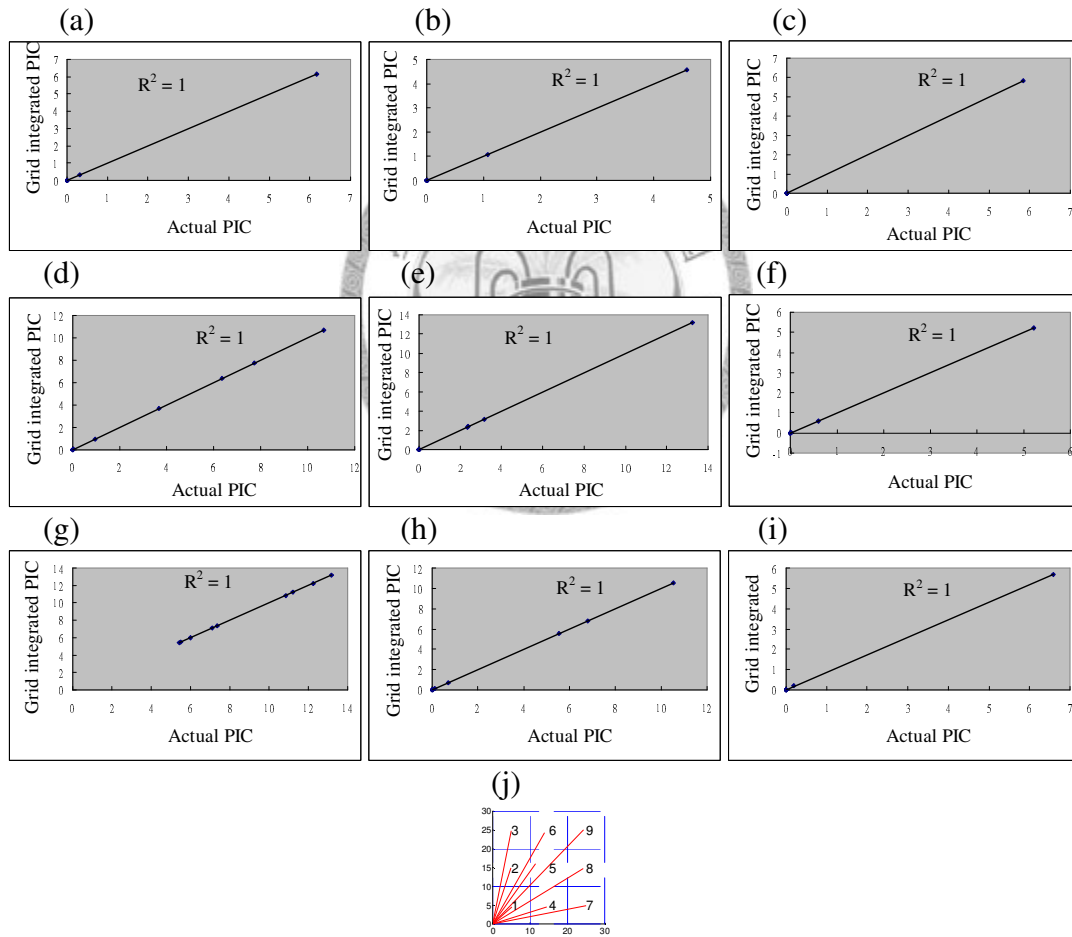
1. Make sure there is no back ground concentration among the experimental domain.
2. Collect 16 single beam spectra continuously for each monitoring line.

3. Generate the absorbance spectrum from the secondly collected spectrum using the first collected spectrum as background.
4. Generate the absorbance spectrum from the thirdly collected spectrum using the secondly collected spectrum as background.
5. Continue the previous step until 15 absorbance spectra are generated.
6. Quantify the target compound from the 15 absorbance spectra.
7. Calculate the standard deviation of the target compound's concentration among the 15 spectra.
8. Three times the standard deviation is the detection limit of current beam path.



Appendix 2. The scatter plot between real and Grid integrated PIC.

We generate a bivariate Gaussian distribution with peak location in the middle of each pixel. For each distribution, both the grid integrated method (Equation 2.3) and the analytical method (Equation 2-10) are applied to calculate the derived PIC. The Figure and the corresponding Table shows the actual PIC (i.e. calculated by analytical method) and the grid integrated PIC. It can be seen that in all of the cases, the error between the actual and grid integrated PIC is less than 3%.



(a)

	Path1	Path2	Path3	Path4	Path5	Path6	Path7	Path8	Path9
Actual PIC	6.17	0.01	0.31	0.00	0.00	0.00	0.00	0.00	0.00
Grid PIC	6.17	0.01	0.31	0.00	0.00	0.00	0.00	0.00	0.00
Error(%)	-0.07	1.36	0.77	0	0	0	0	0	0

(b)

	Path1	Path2	Path3	Path4	Path5	Path6	Path7	Path8	Path9
Actual PIC	0.01	0.00	4.58	0.01	1.08	0.00	0.00	0.00	0.00
Grid PIC	0.01	0.00	4.58	0.01	1.08	0.00	0.00	0.00	0.00
Error(%)	-0.10	0	-0.07	-2.78	0.17	0	0	0	0

(c)

	Path1	Path2	Path3	Path4	Path5	Path6	Path7	Path8	Path9
Actual PIC	0.00	0.00	0.00	0.00	5.83	0.00	0.01	0.00	0.00
Grid PIC	0.00	0.00	0.00	0.00	5.83	0.00	0.01	0.00	0.00
Error(%)	0.10	0	0.09	0	-0.04	0	2.93	0	0

(d)

	Path1	Path2	Path3	Path4	Path5	Path6	Path7	Path8	Path9
Actual PIC	10.68	6.35	7.73	3.66	0.95	0.00	0.04	0.00	0.00
Grid PIC	10.68	6.35	7.73	3.67	0.96	0.00	0.04	0.00	0.00
Error(%)	-0.06	-0.09	0.01	0.26	0.72	-1.63	1.27	1.95	2.50

(e)

	Path1	Path2	Path3	Path4	Path5	Path6	Path7	Path8	Path9
Actual PIC	0.006	0.011	2.388	2.356	13.24	0	3.158	0.007	0.005
Grid PIC	0.006	0.011	2.394	2.348	13.23	0	3.165	0.007	0.005
Error(%)	1.70	2.39	0.27	0.31	0.10	0	0.20	2.22	1.41

(f)

	Path1	Path2	Path3	Path4	Path5	Path6	Path7	Path8	Path9
Actual PIC	0.00	0.00	0.00	0.00	0.61	0.00	5.22	0.00	0.01
Grid PIC	0.00	0.00	0.00	0.00	0.61	0.00	5.22	0.00	0.01
Error(%)	0	0	0	0	0.26	1.66	-0.04	0	0

(g)

	Path1	Path2	Path3	Path4	Path5	Path6	Path7	Path8	Path9
Actual PIC	5.50	7.35	10.87	12.25	13.17	6.02	11.22	7.10	5.44
Grid PIC	5.51	7.36	10.87	12.24	13.16	6.01	11.21	7.11	5.45
Error(%)	0.21	0.10	-0.04	-0.08	-0.10	-0.12	-0.05	0.09	0.23

(h)

	Path1	Path2	Path3	Path4	Path5	Path6	Path7	Path8	Path9
Actual PIC	0.00	0.00	0.03	0.11	0.71	0.01	6.79	5.54	10.53
Grid PIC	0.00	0.00	0.03	0.11	0.72	0.01	6.79	5.53	10.51
Error(%)	2.35	2.19	1.33	1.00	0.82	1.60	0.04	-0.21	-0.11

(i)

	Path1	Path2	Path3	Path4	Path5	Path6	Path7	Path8	Path9
Actual PIC	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.00	6.59
Grid PIC	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.00	5.69
Error(%)	0	0	0	0	0	0	0.76	-0.98	-13.6

