

THESIS

SOURCE APPORTIONMENT OF AEROSOL MEASURED IN THE NORTHERN  
SOUTH CHINA SEA DURING SPRINGTIME

Submitted by

Samuel A. Atwood

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2012

Master's Committee:

Advisor: Sonia M. Kreidenweis

Susan C. van den Heever

Jennifer L. Peel

## ABSTRACT

### SOURCE APPORTIONMENT OF AEROSOL MEASURED IN THE NORTHERN SOUTH CHINA SEA DURING SPRINGTIME

Large sources of aerosol are known to exist in Asia, but the nature of these sources and their impacts on surface particulate matter concentrations are presently not well understood, due in part to the complex meteorology in the region and the lack of speciated aerosol observations. This work presents findings from a pilot study that was aimed at improving knowledge in these areas. Aerosol was collected at a sea-level surface site using an 8-stage DRUM cascade impactor during an approximately six week study at Dongsha Island in the northern South China Sea in the Spring of 2010. The samples were analyzed by X-ray fluorescence (XRF) for selected elemental concentrations, and factor analysis was performed on the results using principal component analysis (PCA). The six factors extracted by PCA were identified as various dust, pollution, and sea salt aerosol types. A refined coarse mode only factor analysis yielded three coarse factors identified as dust, pollution laden dust, and sea salt. Backtrajectory analysis with the HYSPLIT trajectory model indicated likely source regions for dust factors to be in western and northern China and Mongolia, consistent with the known dust sources in the Gobi and Taklimakan Deserts. Pollution factors tended to be associated with transport from coastal China where large population and industrial centers exist, while sea salt sources indicated more diffuse marine regions. The

results were generally consistent with observations from a co-located three-wavelength nephelometer and AERONET radiometer, along with model predictions from the Navy Aerosol Analysis and Prediction System (NAAPS).

Backtrajectories indicated that transport of aerosol to the surface at Dongsha was occurring primarily within the boundary layer from regions generally to the north; an observation consistent with the dominance of pollution and dust aerosol in the ground-based data set. In contrast, more westerly flow aloft transported air from regions to the south and west, where biomass burning was a more significant aerosol source; however, this particle type was not clearly identified in the surface aerosol composition, consistent with it remaining primarily aloft and not mixing strongly to the surface during the study. Significant vertical wind shear and temperature inversions in the region support this conceptual understanding and suggest the potential for considerable vertical inhomogeneity in the SCS aerosol environment.

## ACKNOWLEDGEMENTS

I would like to thank my advisor, Sonia Kreidenweis, and acknowledge the tremendous role she has played in guiding me through the research and preparation of this thesis. The ultimate path this project took, like many scientific endeavours, was neither straightforward nor predictable. Despite my propensity to get carried away on a veritable search for Atlantis, she was able to keep me on track and focused while at the same time teaching me the intricacies of a new field. For that, I am immensely thankful. Thanks also to my committee members Sue van den Heever and Jennifer Peel.

I also gratefully acknowledge the efforts of Jeff Reid to assist me with this project. In serving as both mentor for part of my research, and challenger of my ideas on science, he has pushed me to develop further as a scientist and engineer. I could not have completed this work, nor fully understood the complexity involved in analysis of large datasets without his many hours of teaching and mentoring during my time working with him at the Naval Research Lab.

I would also like to thank members of the Aerosol and Radiation Section at NRL for many useful conversations and help with data processing, including James Campbell, Ed Hyer, Cynthia Curtis, Walter Sessions, Betsy Reid, and Peng Xian-Lynch; along with the efforts of the ALS program including Kevin Perry. The CSU CG/AR staff were of great help to me throughout my work. In particular, I would like to thank Andy Jones, Loretta Wilson, and Tom Vonder Haar for their enthusiasm and support. Thanks also to the faculty and staff of the Climate and Society program at Columbia University for first introducing me to the study of aerosol science.

This work was funded by the NRL Base Research Program and the Colorado State University Center for Geosciences/Atmospheric Research (CG/AR). The Dongsha Experiment, as part of the international 7SEAS activities, was funded by the Taiwanese Environmental Protection Administration and National Science Council for the participation of Taiwanese scientists, and by the NASA Radiation Sciences Program, managed by Dr. Hal B. Maring, for the deployment of the COMMIT mobile laboratory. Part of my research was conducted as part of the Naval Research Enterprise Internship Program (NREIP) internship at the U.S. Naval Research Lab in Monterey, CA. I am grateful for the support of the staff and students of the National Central University of Taiwan. The support of Taiwan EPA, for the development of the Dongsha site and its maintenance, is also gratefully recognized.

Finally, I would like to thank my parents and brothers for their continuous support over the years during my constant wanderings and pursuit of discovery. I would also like to thank Matt Garcia for his insight, discussion, humor, and efforts to keep me sane during the last few years.

## TABLE OF CONTENTS

<b>1. Introduction.....</b>	<b>1</b>
1.1. Receptor Modeling .....	1
1.1.1. Factor Analysis Methods .....	2
1.1.2. Backtrajectory Methods .....	3
1.1.3. Measurements of Aerosol Properties .....	4
1.1.4. Limitations of Source Apportionment .....	4
1.2. Specific DoD Needs.....	5
1.3. Locations of Interest .....	6
1.4. Primary Study Goals .....	7
<b>2. Methods.....</b>	<b>9</b>
2.1. DRUM Cascade Impactor and X-Ray Fluorescence Analysis .....	10
2.1.1. Timestamp Correction of DRUM data .....	13
2.2. Factor Analysis of DRUM Data .....	15
2.3. Trajectory Modeling .....	17
2.3.1. The HYSPLIT Model .....	18
2.3.2. Baghdad Example Case Study .....	20
2.3.3. Individual Backtrajectories .....	20
2.3.4. Residence Time Analysis.....	24
2.3.5. Source Contribution Functions .....	25
2.3.6. Use of Measured In-Situ Data.....	27
2.3.7. Comparison of the Normal and Special Cases.....	31
2.4. Additional Supporting Measurements .....	39
2.4.1. NAAPS Model .....	39
2.4.2. Additional Dongsha In-situ Measurements .....	40
<b>3. Results and Discussion.....</b>	<b>41</b>
3.1. Site Description and Background .....	42
3.1.1. Meteorology .....	43
3.1.2. Expected Source Regions .....	44
3.1.3. Low-Level Inversion and Atmospheric Stability.....	46
3.2. DRUM/XRF Results.....	49
3.2.1. Time Stamp Correction.....	49
3.2.2. Size Fraction Correction .....	53
3.3. Factor Analysis .....	54
3.3.1. Three-Mode Factor Analysis .....	56
3.3.2. Coarse-Mode Factor Analysis.....	59

3.4. Backtrajectory Analysis .....	62
3.4.1. Dongsha Normal Case .....	62
3.4.2. Dongsha Three-Mode Factor Analysis Special Cases .....	66
3.4.3. Dongsha Coarse-Mode Factor Special Cases .....	77
3.5. Discussion and Additional Analyses .....	80
3.5.1. Study Period – Regional Average .....	81
3.5.2. Study Period – Dongsha Time Series .....	82
<b>4. Summary and Conclusions.....</b>	<b>86</b>
4.1. Future Work.....	90
<b>5. References.....</b>	<b>93</b>
Appendix A <b>Matrix Effects Correction .....</b>	<b>99</b>
Appendix B <b>Three-Mode Factor Analysis.....</b>	<b>101</b>
Appendix C <b>Coarse-Mode Factor Analysis .....</b>	<b>106</b>
Appendix D <b>Additional Special Case Backtrajectory Analysis .....</b>	<b>108</b>

LIST OF TABLES

Table 1 Elemental species measured by XRF and average uncertainties ( $\mu\text{g}/\text{m}^3$ ) across all measurements and size fractions..... 12

Table 2 Dongsha DRUM size fraction cross-correlation. The peak cross-correlation between each size fraction (e.g. S1 for size fraction 1) for selected elemental and summed mass concentrations are shown. .... 51

Table 3 Dongsha Three-Mode Factor Source Types and Associated Elemental Composition..... 58

Table 4 Dongsha Coarse-Mode Factor Source Types and Associated Elemental Composition..... 61

Table 5 KMO and Bartlett's Test Statistics for the Dongsha Three Mode Factors 101

Table 6 Total Variance Explained for Dongsha Three Mode Factors ..... 101

Table 7 Scree Plot for Dongsha Three Mode Factors..... 101

Table 8 Rotated Component Matrix for Dongsha Three Mode Factors ..... 102

Table 9 KMO and Bartlett's Test Statistics for the Dongsha Coarse Mode Factors ..... 106

Table 10 Total Variance Explained for Dongsha Coarse Mode Factors ..... 106

Table 11 Scree Plot for Dongsha Coarse Mode Factors ..... 106

Table 12 Rotated Component Matrix for Dongsha Coarse Mode Factors ..... 107

## LIST OF FIGURES

Figure 1 Backtrajectories from 31 May to 21 June for a Baghdad receptor (33.26 N, 44.27 E), colored by date of arrival. In this and subsequent similar plots, the receptor site is indicated by a black star. ....	21
Figure 2 Heights of backtrajectories shown in Figure 1. Trajectories that reach the top of the model domain are truncated and are indicated by stars for the entire trajectory. ....	21
Figure 3 Backtrajectories in Figure 1, color-coded by height (m).....	22
Figure 4 Backtrajectories in Figure 1, color-coded by age (hours). ....	22
Figure 5 Backtrajectories in Figure 1, shown as a function of latitude and height (m), and color-coded by age (hours).....	23
Figure 6 Backtrajectories in Figure 1, shown as a function of longitude and height (m), and color-coded by age (hours).....	23
Figure 7 The Relative Residence Time Analysis plot for Baghdad during the study period. ....	25
Figure 8 Source Contribution Function for Baghdad during the study period. ....	27
Figure 9 Baghdad DRUM timeline for Coarse and Fine Lead fractions in June 2008. ....	28
Figure 10 Special Case Residence Time Analysis for elevated coarse mode Pb. ...	29
Figure 11 Specific Special Case Residence Time Analysis for elevated Coarse mode Pb. ....	30
Figure 12 Specific Special Case Source Contribution Function for elevated Coarse mode Pb. ....	31
Figure 13 Potential Source Contribution Function (%) for elevated Coarse mode Pb. ....	34
Figure 14 Potential Source Contribution Function (%) with a 10 endpoint minimum criterion for elevated coarse mode Pb.....	35
Figure 15 Residence Time Difference Function for elevated coarse mode Pb.....	37
Figure 16 Source Contribution Difference Function for elevated coarse mode Pb. ....	37
Figure 17 Residence Time Difference Function plotted on a gridded domain. ....	39
Figure 18 Dongsha Island location. ....	43
Figure 19 Dongsha seasonal SCF plots for 2009 to 2010 at 100m receptor heights. ....	45
Figure 20 Dongsha seasonal SCF plots for 2009 to 2010 at 1500m receptor heights. ....	46
Figure 21 NCEP Reanalysis temperature profile for lowest 3000 m of only the grid box containing Dongsha Island. Dry Adiabatic Lapse Rate (solid gray) and Saturated Adiabatic Lapse Rate (dashed gray) are shown for comparison.....	48

Figure 22 NCEP Reanalysis Lapse Rate profile at standard pressure levels for grid boxes in the northern South China Sea area during the Dongsha Experiment. DALR and SALR shown in gray. ....	48
Figure 23 Dongsha Size Fraction Time Series. Selected plots for concentration ( $\text{ng}/\text{m}^3$ ) of iron, total mass, and lead show alignment between size fractions. ....	52
Figure 24 Dongsha DRUM summed mass concentration for various size fractions and nephelometer scattering. ....	53
Figure 25 Dongsha mass size distributions ( $\text{ng} / \text{m}^3$ ) for summed mass and selected elements. ....	54
Figure 26 Timeline of Dongsha Three-Mode Factor Analysis scores for (a) the raw factor scores, and (b) factors smoothed by 24-hour boxcar average. ....	59
Figure 27 Timeline of Dongsha Coarse-Mode Factor Analysis scores. ....	61
Figure 28 Dongsha coarse-mode mass fractions of indicated elements for all time stamps, and for time stamps when each factor score was greater than 1. ....	62
Figure 29 Dongsha Normal Case backtrajectories and SCF at 100m, 500m, 1000m, and 1500m. ....	64
Figure 30 Dongsha Normal Case backtrajectory heights at 100m, 500m, 1000m, and 1500m. ....	65
Figure 31 Dongsha Three Mode Special Case backtrajectory plots for 100m receptor heights. ....	67
Figure 32 Dongsha Three Mode Special Case RTA plots for 100m receptor heights. ....	68
Figure 33 Dongsha Three Mode Special Case PSCF plots for 100m Receptor Heights. ....	69
Figure 34 Dongsha Three Mode Special Case PSCF plots for 100m Receptor Heights with 10 endpoint cutoff. ....	71
Figure 35 Atmospheric sounding from the China coast just to the Northwest of the Pearl River Delta showing a low-level temperature inversion. ....	72
Figure 36 Dongsha Three Mode Special Case PSCF plots for 100m receptor heights, filtered for endpoints within 2000m above ground height. ....	73
Figure 37 Dongsha Three Mode Special Case RTDF plots for 100m receptor heights. ....	74
Figure 38 Dongsha Coarse-Mode Special Case backtrajectory plots for 100m Receptor Heights. Shown are trajectory (a) arrival times, (b) altitudes (m), and (c) ages (hours). ....	79
Figure 39 Dongsha Coarse-Mode Special Case backtrajectory plots for 100m Receptor Heights. Shown are (a) special case SCF, (b) PSCF, and (c) RTDF plots. ....	80
Figure 40 Regional average AODs during the Dongsha Experiment from (a) MODIS and (b) NAAPS. ....	82
Figure 41 Dongsha receptor (a) daily averaged column AOD from AERONET, (b) backscattering from TSI three-wavelength nephelometer, (c) column AOD for	

various aerosol constituents from NAAPS, and (d) surface concentrations ( $\mu\text{g}/\text{m}^3$ ) for various aerosol constituents from NAAPS. ....	85
Figure 42 Concentration ratios ( $\text{ng} / \text{m}^3$ ) against Fe for light elements in Dongsha DRUM data. ....	100
Figure 43 Three-Mode Factor 1 time series with coarse mode (top), and accumulation and ultrafine mode (bottom) concentrations ( $\text{ng} / \text{m}^3$ ) for elements with high factor 1 component scores. ....	103
Figure 44 Three-Mode Factor 2 time series (similar to Figure B-1) for elements with high factor 2 component scores. Note that accumulation mode Zn concentration is often greater than scale shown, with peaks reaching as high as $500 \text{ ng} / \text{m}^3$ . ....	103
Figure 45 Three-Mode Factor 3 time series (similar to Figure B-1) for elements with high factor 3 component scores. ....	104
Figure 46 Three-Mode Factor 4 time series (similar to Figure B-1) for elements with high factor 4 component scores. ....	104
Figure 47 Three-Mode Factor 5 time series (similar to Figure B-1) for elements with high factor 5 component scores. ....	104
Figure 48 Three-Mode Factor 6 time series (similar to Figure B-1) for elements with high factor 6 component scores. ....	105
Figure 49 Dongsha Three Mode Special Case PSCF plots for 500m Receptor Heights. ....	108
Figure 50 Dongsha Three Mode Special Case PSCF plots for 1000m Receptor Heights. ....	109
Figure 51 Dongsha Three Mode Special Case PSCF plots for 1500m Receptor Heights. ....	110

## **1. INTRODUCTION**

Atmospheric aerosol is a complex amalgam of gaseous, aqueous, and solid trace constituents known to have impacts on air quality, visibility, human health, climate, and cloud and precipitation formation. Decisions on emissions control strategies aimed at improving air quality or health impacts require knowledge of the impacts specific changes are likely to have. Two common methods for assessing these likely impacts are dispersion modeling to investigate possible fates of specified emissions, and receptor modeling and source apportionment to examine likely sources and types of aerosol measured at specific locations. This research utilizes several methods of receptor modeling to support source apportionment efforts, and to allow for investigation of the nature of atmospheric aerosol that is measured at a location.

### **1.1. RECEPTOR MODELING**

A variety of receptor models and source apportionment methods have been in common use for many years to identify the types of emission sources potentially impacting a receptor. Optical and chemical measurements of the atmospheric environment are often analyzed for likely source type, while meteorological and transport information can provide indications of the potential location of emission sources. Information on the specific potential impact these sources may have on air quality at a

receptor can then be used to better inform decision making and target resources to efforts that are more likely to improve air quality and achieve desired results.

Linking a source to a receptor is often conducted via chemical mass balance (CMB) when sources are known, or through factor analysis when sources are to be estimated based on measured results [Hopke, 2003; Watson *et al.*, 2008]. The U.S. Environmental Protection Agency (EPA) supports several computer receptor models for source apportionment studies [Coutant *et al.*, 2003], while other authors give source profiles typical of various types of sources [Cohen *et al.*, 2010a; VanCuren *et al.*, 2005]. In addition, transport models such as HYSPLIT and FLEXPART provide means to estimate the link between measured values at a receptor and potential source regions using meteorological information [Ashbaugh *et al.*, 1985; Cohen *et al.*, 2010b; Draxler and Hess, 1998; Hopke, 2003; Stohl, 1996]. Hopke (2003) provided an overview of recent developments in receptor modeling methods, which served as a reference for the descriptions in the following subsections.

#### **1.1.1. FACTOR ANALYSIS METHODS**

Determination of source type using measured or observed data from a receptor is often conducted by matching measured aerosol properties to typical source profiles of known aerosol sources. When measured datasets are complex, various data reduction methods, known collectively as factor analysis, simplify the data into manageable components that can be compared to the source profiles. These factors allow for both identification of types of sources impacting the receptor, and determination of when each source was prevalent in measured results.

Several factor analysis methods have been developed and applied to analysis of measured aerosol properties [Watson *et al.*, 2008], the most common of which is known as Principal Component Analysis (PCA). This statistical technique uses eigenvector analysis to extract factors based on the degree to which variables within the data covary with each other [Henry, 1991; Hopke, 1985]. Henry (1987) noted a weakness in PCA in that it did not take into account the uncertainty in individual measurements, leading to improper scaling of the factors. Paatero and Tapper (1993; 1994) then described an alternate method, known as Positive Matrix Factorization (PMF) that accounts for this bias through weighting of each data point using an explicit least squares approach to factor analysis. Differences exist between the methods, with PMF involving a more complex analysis and requiring more information on the uncertainty associated with measured data. Huang *et al.* (1999) reported that the methods yielded generally similar results, with PMF providing moderately better quantification at the expense of a more complex analysis.

### **1.1.2. BACKTRAJECTORY METHODS**

Meteorological information and atmospheric models are often used to estimate transport from source to receptor. In the case of receptor modeling, estimations of the trajectory of air parcels arriving at the receptor are known as backward trajectories, or simply backtrajectories. A backtrajectory is calculated by starting an air parcel at some height above the receptor and running the model backwards to estimate its path to the receptor. Ashbaugh *et al.* (1985) described a statistical method to analyze backtrajectories arriving at time periods of interest in order to assess likely source areas. Further development included numerous methods for utilizing measured properties of aerosol and

factors to indicate likely source areas [Poirot *et al.*, 2001]. Ultimately, two methods known generally as residence time analysis (RTA) and potential source contribution function (PSCF) analysis are typically used for source apportionment of measured aerosol at a receptor [Hopke, 2003].

### **1.1.3. MEASUREMENTS OF AEROSOL PROPERTIES**

Apportioning aerosol to specific sources is conducted using numerous techniques for measuring aerosol properties. Aerosol measurements subjected to factor analysis have included ambient particle size distribution, black carbon mass, aerosol optical properties, ion concentration in precipitation, and elemental concentration, among many others [Juntto and Paatero, 1994; E Kim *et al.*, 2004; Polissar *et al.*, 1999]. Size resolved elemental concentrations of aerosol collected on a DRUM cascade impaction sampler are used for source apportionment in this research. Similar methods were used by VanCuren *et al.* (2005), who used PCA of aerosol sampled by a DRUM instrument to investigate intercontinental transport of Asian dust and pollution to California, and to differentiate between local dust, Asian dust, and marine aerosol impacts. Han *et al.* (2005) likewise used the combination of DRUM impactor and PCA to generate factors associated with known emission sources in Korea.

### **1.1.4. LIMITATIONS OF SOURCE APPORTIONMENT**

While source apportionment methods allow for the estimation of sources impacting a receptor using a variety of tools, their limitations should also be noted prior to their use. Source apportionment, by its nature, is based on modeled and statistical estimates of actual physical processes. As a result, it is subject to any uncertainties in the methods it utilizes, and cannot achieve confidence greater than limitations of these methods.

In the case of this research, extracted factors explain only a fraction of the variability in aerosol measurements. Source apportionment will therefore explain only the potential sources associated with this fraction of variance, leaving the remainder to be either from an unexplained source or associated only with noise in the measurement. Small or less frequent sources could therefore be missed by this method. Backtrajectories used here likewise explain only potential source regions based on a synoptic modeling of the atmosphere. Physical processes not accounted for in the model, but which may nevertheless contribute to transport of aerosol to the receptor cannot be resolved and can therefore limit source apportionment efforts.

## **1.2. SPECIFIC DoD NEEDS**

The Department of Defense (DoD) maintains interests in aerosol source apportionment that are similar to the wider aerosol community, however several needs specific to their activities arise. High particulate matter loadings at DoD locations, particularly those in South and Southwest Asia, have been recorded at levels that may create adverse health effects for personnel located in these regions [Engelbrecht *et al.*, 2008]. Burn pits, dust storms, and pollution are among contributors to periods of potentially hazardous air quality, but determination of the specific impacts from each possible source is difficult. Source apportionment of measured particulate matter data represents a potential method for differentiating between local and regional or long-range transport aerosol sources, as well as to determine what specific types of sources may have been impacting various receptors.

In addition to health effects, aerosol can potentially obstruct communications, degrade visibility, and interfere with the propagation of electromagnetic signals through

the atmosphere. The chemical and physical properties of aerosol, along with their size and concentration, will determine the degree to which any of these adverse effects may impact a given region. Source apportionment studies provide information that can potentially improve knowledge of these properties through better understanding of the role each source or source type plays in the aerosol environment.

Several atmospheric general circulation and chemical models are supported by the DoD and used for prediction of aerosol at various locations around the Earth [*Hogan and Rosmond, 1991; Reid et al., 2009*]. Source apportionment studies and a better understanding of aerosol at a receptor allow for validation of model results and improvement in predictive abilities. As DoD strives to maintain operational capability throughout the world, improving these aerosol models will support its efforts.

### **1.3. LOCATIONS OF INTEREST**

In the region encompassing Southeast and East Asia, increasing sources of aerosol from biomass burning, population and industrial centers, and deserts have generated significant interest. A number of research partners, including those from DoD, are collaborating in the multi-year 7SEAS campaign to investigate atmospheric aerosol and its interactions with, and effects on, air quality, climate, clouds and precipitation in Southeast Asia and the South China Sea (SCS) [*7SEAS-Whitepaper, March 4, 2009*].

Major aerosol sources for the region include biomass burning from slash and burn conversion of forests to agricultural land, yearly burning of waste agricultural products, smoldering peat fires, and highly variable natural and anthropogenically caused wildfires. Anthropogenic sources of pollution are varied and include industrial emissions, road dust, and combustion byproducts, among others from major population and industrial centers

throughout Asia. Crustal dust sources from the large Taklimakan and Gobi deserts additionally contribute to measured aerosol in the region. Source apportionment methods applied to data gathered during this campaign will be analyzed in light of these aerosol sources in order to support ongoing research efforts.

In addition to East Asia and the SCS, initial aerosol source apportionment techniques for Baghdad, Iraq were considered as part of ongoing DoD efforts to study this area. Future work will likely involve additional data collection and analysis, and efforts to differentiate between local and regional or distant sources that may be impacting the receptor.

#### **1.4. PRIMARY STUDY GOALS**

In support of the needs of the DoD, along with the wider aerosol community, the goals of this research are twofold. First, a method was to be developed that utilized existing techniques for source apportionment through the use of both chemical and meteorological information. These efforts are intended to support research in areas important to the DoD and others.

Second, the method is to be applied to support an approximately six week study in Spring 2010 at Dongsha Island as part of an initial investigation into aerosol in the SCS. In support of this goal, measured in-situ data from a DRUM instrument are presented and analyzed to better understand the chemistry and source types of aerosol that were impacting the SCS, while meteorological information is utilized to investigate potential aerosol source regions. As research in this region continues, the results of this source apportionment study will allow for better understanding of the potential aerosol sources

impacting the SCS in springtime, and can be used for model validation and planning purposes for future study.

## 2. METHODS

A dataset of measured aerosol chemical and physical properties typically provides information about the constituents of the atmosphere at one (or several) discrete receptor points over some period of time. In addition to information about the state of atmospheric aerosol at these receptor points, it is often useful to know what the sources of aerosol are, and how each source has contributed to the observation at a particular time. Receptor modeling therefore serves as a tool for establishing source/receptor relationships [Hopke; 1985; 2003]. Involved in the method is the interpretation of measured in-situ aerosol properties, along with additional meteorological, transport, and source characteristics information, to formulate a better understanding of the aerosol sources and environmental conditions which lead to a given measured atmospheric state.

Methods of receptor modeling are varied, but may include analysis of chemical and physical data to identify aerosol type, use of meteorological data to identify source areas, or evaluation of variability to identify trends or specific events. Using a variety of approaches can provide the user with a certain degree of independence between estimates and improve confidence in the results if several methods indicate a common aerosol source. Thus, methods that allow for interpretations of available data from a variety of perspectives may result in a more complete picture of the atmospheric aerosol than is possible by simply analyzing in-situ data from a single source that is considered independent of the wider environment in which it exists.

The methods described here represent the measurements and receptor modeling methods used to conduct an analysis of the aerosol environment at Dongsha Atoll (discussed further in Methods section) in the Spring of 2010. The methods described may be applied to similar measurements from a Baghdad receptor or other locations of interest, although each application requires an individual assessment for data accuracy and the suitability of the technique for the specific location and time period.

## **2.1. DRUM CASCADE IMPACTOR AND X-RAY FLUORESCENCE ANALYSIS**

The eight-stage Davis Rotating-drum Uniform size-cut Monitor (DRUM) sampler is an aerosol cascade impactor [Raabe *et al.*, 1988] that segregates aerosol particles based on their aerodynamic diameters. A series of orifices and rotating substrates are configured to consecutively remove aerosol of decreasing size from the airflow by inertial impaction. As the sample air laden with aerosol progresses through each stage, the inertia of the largest aerosol particles causes them to impact on the substrate rather than passing through the stage along fluid streamlines. Each stage is configured so as to remove a specific size fraction of aerosol (aerodynamic diameter,  $D_p$ , larger than the cut size) which can later be subjected to a variety of analyses.

The DRUM used in this study was a version of the DRUM sampler originally described by Cahill *et al.* (1985), modified to utilize slit orifices and configured to run at 16 lpm as described in Reid *et al.* (2008). A PM10 sample inlet was used, followed by collection stages with nominal 50% aerodynamic diameter cut sizes of 5  $\mu\text{m}$ , 2.5  $\mu\text{m}$ , 1.15  $\mu\text{m}$ , 0.75  $\mu\text{m}$ , 0.56  $\mu\text{m}$ , 0.34  $\mu\text{m}$ , 0.26  $\mu\text{m}$ , and 0.07  $\mu\text{m}$ . Aerosol particles are collected on Mylar strips coated with Apiezon grease and wrapped around each rotating drum. The drums are then rotated at a consistent rate such that nominal three-hour

resolution timestamps can be assigned to data points of measured species concentration based on location along the strip.

DRUM samples were subjected to X-Ray Fluorescence (XRF) Analysis at the Advanced Light Source (ALS) of Lawrence Berkeley National Laboratory to provide measurements of selected elements shown in Table 1 [*Haller and Knochel, 1996; Perry et al., 2004*]. These elements, with the exception of Sulfur, are generally associated with primary emissions from a variety of typical aerosol sources, including crustal and road dust, anthropogenic pollution, and sea salt. Sulfur may be emitted directly, but is also associated with secondary aerosol formation during transport in the atmosphere. Only elements heavy enough to be detected by XRF and with sufficient variability in typical aerosol are reported, leaving some potentially useful elements such as carbon and oxygen unable to be analyzed by this method.

Mylar strips from the DRUM are bombarded with X-rays thereby causing certain elements within the particles to fluoresce. This fluorescence involves the emission of radiation characteristic to each element being measured. Therefore, the ambient concentrations of elements in the sample can be established by detection of the magnitude of this characteristic radiation, after suitable calibration with standards and after correction for the sample flow rate. The X-ray beam is directed at sequential points along the length of the sample strip to measure elemental concentration at selected timestamps. As each Mylar strip contains aerosol of a specific size range, size and time resolved aerosol elemental concentrations are established.

Reported uncertainty values in DRUM measurements are calculated based on the propagation of uncertainties from several inputs. A flow rate uncertainty and standard

measurement value uncertainty of 5% each is used for this sampling, along with uncertainty from the XRF data reduction algorithm which varies for each element and sample. The result is an uncertainty value for each reported data point that represents the standard deviation of expected measurements of that data point. If the uncertainty is greater than the measured value, a zero is given in the data set. Average uncertainties for each elements for all size fractions are given in Table 1. The minimum detection limit (MDL) for the method also varies depending on experimental setup and the spectral analysis of each element. An MDL can generally be calculated for this XRF method by averaging the reported uncertainty for zero values for each element in the sample and multiplying by three. The MDL is therefore three times the standard deviation of measured values, giving reported results a three sigma, or greater than 99% chance, of being statistically different from a blank (zero concentration) value [Cliff, 2012].

Table 1 Elemental species measured by XRF and average uncertainties (ng/m<sup>3</sup>) across all measurements and size fractions.

Mg	24.96	V	0.04	As	0.04
Al	6.35	Cr	0.01	Se	0.03
Si	6.18	Mn	0.05	Br	0.17
P	0.56	Fe	1.58	Rb	0.07
S	10.00	Co	0.01	Sr	0.11
Cl	11.88	Ni	0.02	Y	0.13
K	1.48	Cu	0.07	Zr	0.17
Ca	2.16	Zn	4.39	Mo	0.39
Ti	0.17	Ga	0.02	Pb	0.22

DRUM instruments, and cascade impactors in general, are known to be subject to several sampling artifacts, most notably the potential for bounce-off and particle shattering, thereby shifting collected mass from upper to lower stages (undersizing) [Reid *et al.*, 2008; J S Reid *et al.*, 2003]. High relative humidities during sampling and humidification due to the pressure drop through the jets may also cause hygroscopic

growth of aerosol particles, resulting in collection of the particle on too high a stage (oversizing). Therefore, confidence in the nominal size resolution obtained from the DRUM stages may be low in certain circumstances.

XRF analysis of these samples may additionally be subject to various sources of potential bias not adequately represented by the nominal reported uncertainty value. Elemental concentrations can be skewed due to matrix effects, particularly in the larger size fractions [E A Reid *et al.*, 2003]. For large particles and cases with significant mass loading, elements lighter than calcium are known to self-absorb photons emitted by XRF excitation. The resulting data points generated under these circumstances will be skewed towards lower concentrations. To correct for this effect, light elements in the larger size fractions are compared to elements not subject to these matrix effects. For instance, in the case of aluminum, concentrations of aluminum ([Al]) are compared against concentrations of iron ([Fe]). [Al] vs [Fe] scatter plots are first created to test for linearity. The relative concentrations of these crustal materials from a specific source are expected to remain roughly constant at any overall concentration of aerosol at a given size fraction. If a mass deficit is discovered at high concentrations where matrix effects would predominate, a linear correction is applied to the concentration data points as a function of [Fe].

### **2.1.1. TIMESTAMP CORRECTION OF DRUM DATA**

XRF results for each sample strip were provided at three hour time resolution based on linear distance along the sample strip. These results are based on start and end times noted in a field log, along with the assumption of fixed times between data points based on constant rotation speed. In order to ensure these time stamps are correct for each

sample strip, the various elemental concentration time series from each strip (that is, each size cut) were first compared to each other to ensure the relative alignment between strips was correct. Next, timelines of elemental concentrations were compared with timelines of light scattering data for a co-located nephelometer. Light scattering measurements are closely related to aerosol mass concentrations, and can be taken and recorded at high time resolution. Detection of similar high mass concentration / high scattering events can thus provide confidence in the timestamps of those events.

A three-wavelength TSI nephelometer (450, 550, 700 nm) instrument suitable for measuring light scattering from a sample of atmospheric aerosol was deployed to Dongsha Island as part of the 7SEAS campaign during the time when the DRUM was operating, although we use only the green channel (550 nm) for timing and estimation of the relative duration of events. Truncation/non-lambertian light source corrections were made to the nephelometer data using the parameterizations of *Anderson and Ogren* (1998).

Nephelometer data were provided at five minute intervals. A three-hour average time series was created to match the initial estimated DRUM timestamp. This time series was then cross-correlated to DRUM total mass and specific DRUM elemental time series at a variety of lagged time steps. To provide a further refinement to this correction, multiple rolling three-hour average time series, offset in 30 minute increments from the DRUM timestamp, were created and likewise cross-correlated to the DRUM time series.

A further “hand analysis” of the data was conducted wherein plumes or characteristic events in both the DRUM and nephelometer data were compared against each other. The timestamps of large peaks or variability in the data were compared in the two datasets to

provide a further means of aligning the DRUM data in time. After any required adjustments, the final result of these combined analyses is high confidence in the time stamps assigned to the DRUM data.

## **2.2. FACTOR ANALYSIS OF DRUM DATA**

The raw time series of the DRUM data represent in-depth measurements of the nature of aerosol impacting the receptor. Although providing useful information, the sheer amount of data can be difficult to analyze by visual observation alone. A more objective and manageable interpretation of aerosol variability can be created by extracting only the most important components of the aerosol that change with time. For instance, examination of the time series of concentrations of typical crustal constituents such as iron can indicate when dust impacts are occurring, while peaks in certain heavy metals may signal industrial pollution impacts. Thus, a complex series of variables is reduced to a simple series that still retains effectiveness. However, in order to make full use of the available measurements, more advanced statistical techniques can be used in order to both include all measures of variability while simultaneously reducing the data to meaningful quantities.

Factor Analysis (FA) provides a means of data reduction through examination of the relationships between measured variables [Hopke, 2003]. Here, each of the DRUM data points in the time series includes a measurement for each elemental concentration at each size fraction. These elemental concentration variables are all considered to be a separate dimension of the measurement at any point in time. Each data point therefore has a magnitude in each of these dimensions. The variability within this dataset consequently arises as a result of the variability within each individual dimension. In the case described

above of iron variability as a proxy for dust impacts, we note that other crustal materials such as silicon and aluminum would be expected to vary in a similar manner as well. These common variations among dimensions represent the data reduction strategy employed by FA. Common patterns of variability are explained by a new set of fewer dimensions, called factors.

While PMF has the potential for better resolution and better quantification of factors, conventional factor analysis using PCA was deemed sufficient for the initial analysis and identification of sources in this research. As the Dongsha dataset was limited to only six weeks as part of an initial study, the more complex PMF method can be utilized in future work with the goal of better quantification of factors and sources.

Factor analysis using PCA aims to reduce the covariability in the dataset variables to a set of factors that explain some fraction of the variance in the larger dataset. The PCA was performed using the IBM SPSS statistical package (IBM SPSS Statistics, Version 19; <http://www-01.ibm.com/software/analytics/spss/>). Explanations will reference specific terminology used in this software package, however, many computer packages exist which use the same methodology.

Linear combinations of variables were first generated with PCA, which are then rotated in order to maximize the variance explained by a selected number of factors using the SPSS Varimax rotation option. Data suitability for factor analysis was verified by first testing to ensure Bartlett's test of sphericity to be significant ( $p > 0.05$ ) and the Kaiser-Meyer-Olkin (KMO) statistic to be greater than 0.6 as suggested by *Pallant*, (2007). An exploratory approach was then used to determine the number of factors to extract. First, Kaiser's criterion, wherein factors with an eigenvalue greater than 1 are extracted, was

utilized, followed by a refined analysis in which fewer factors were extracted based on a scree test [Pallant, 2007; Tabachnick and Fidell, 2008]. The final consideration in determining the number of factors to extract was based on a physical assessment of the generated factors using elemental concentrations and the rotated component matrix. The number of factors to be extracted for analysis was selected to ensure a coherent physical and chemical understanding of the factors was possible. In addition, FA was performed on subsets of the data (for example, only lumped coarse mode data) expected to have unique sources compared to the full spectrum of particle sizes.

### **2.3. TRAJECTORY MODELING**

Trajectory modeling and analysis involves the use of an atmospheric transport model to estimate the path of an air parcel as it travels through the atmosphere. Information output from a meteorological model is used to advect an (massless) air parcel through the atmosphere from a starting position. Chemical or physical changes in the aerosol, dispersion and diffusion, wet and dry deposition processes, and cloud processing have no effect on the calculated trajectory. These effects can sometimes be significant (as in secondary formation of sulfur aerosol) and must be considered during analysis to determine if trajectory modeling is appropriate for each source type or environment.

Furthermore, the resolution of the meteorological fields used to drive the model directly relates to the scale of transport processes that can be resolved. Therefore, if output from the meteorological model is at a synoptic scale, on the order of  $1^\circ \times 1^\circ$  spatial resolution and six-hour temporal resolution, mesoscale and microscale processes such as gust fronts, convection, and turbulence will not be accounted for. It is therefore the responsibility of the user to ensure that proper consideration of the uncertainties

generated by these drawbacks are taken into account to ensure the trajectory model retains its usefulness [Draxler and Hess, 1998; Stohl *et al.*, 2005]. In the case of this research, the trajectory model was used only to attempt to explain the variability in the measured results associated with synoptic processes, and to provide general indications of potential source areas.

### **2.3.1. THE HYSPLIT MODEL**

The National Oceanic and Atmospheric Administration's Air Resources Lab (NOAA-ARL) developed and currently supports the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model for the purpose of trajectory and dispersion modeling. Version 4.9 of the model [Draxler, 2004; Draxler and Hess, 1997; 1998] supports single particle trajectory calculations in three dimensions using several different supported meteorological datasets. Additional options include ensemble methods, forward-backward runs, and built-in trajectory frequency analysis methods. While puff and plume modes also exist within the model, which can include dispersion and deposition processes, these options are not used here. Decisions on which internal HYSPLIT methods and options to use, additional external analysis methods applied, and the reasons for such decisions are outlined below and in the following sections.

In addition to the HYSPLIT model, NOAA-ARL provides meteorological datasets which have been pre-processed into a format able to drive HYSPLIT. The meteorological dataset used in this manuscript is GDAS1, a 1° x 1° global meteorological dataset, generated for HYSPLIT from the Global Data Assimilation System (associated with the GFS meteorological forecast model). GDAS is the only dataset provided by NOAA-ARL, aside from the NCEP Reanalysis, which has global coverage needed for the

Baghdad and Dongsha locations. The model vertical velocity option, which uses vertical velocity from the meteorological dataset, was used to drive vertical motion as opposed to constraining air parcels to follow other paths in the vertical such as constant pressure or potential temperature surfaces. Due to the relatively coarse resolution, calculated trajectories were expected to be sensitive only to transport due to synoptic scale atmospheric processes. As HYSPLIT was used here only to provide general indications of source areas associated with measured aerosol concentrations, the NCEP Reanalysis and various model options were not extensively used.

Trajectories can be run in both the forward and backward directions in time, thereby estimating the probable destination and source regions, respectively. Backward trajectories, also known simply as backtrajectories, are particularly useful for receptor analysis in that they can help identify the source regions for in-situ measurements at specific times [Ashbaugh *et al.*, 1985]. Once a number of trajectories have been generated, further statistical analyses can be used to establish estimated transport pathways, seasonal and annual patterns, and climatologies. HYSPLIT trajectories can also be combined with measured in-situ data to layer transport information on top of analyses of the variability of chemical and physical aerosol properties.

Individual backtrajectories are composed of “endpoints” which, in this research, were output from HYSPLIT every 60 minutes backwards from the arrival time for a total of 120 hours. Each recorded endpoint contains information that includes parcel location in three dimensions and associated meteorological data including solar insolation, precipitation, temperature, potential temperature, and relative humidity. The model setup described in this section was used for all analyses shown in this research. After individual

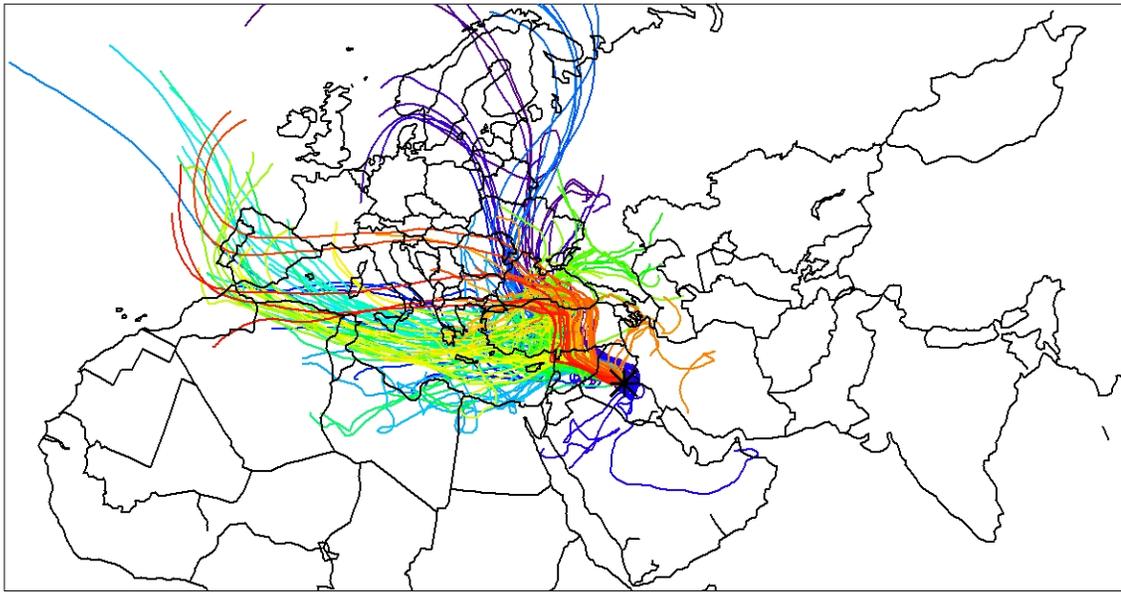
trajectories have been generated, further statistical analysis was conducted using these endpoints.

### **2.3.2. BAGHDAD EXAMPLE CASE STUDY**

To demonstrate the application and interpretation of backtrajectories as used in this work, we used data from a receptor in Baghdad, Iraq. An approximately three week long sample time series of size segregated lead (Pb) concentrations as measured by a DRUM sampler located in Baghdad (33.26 N, 44.27 E) was available for June 2008. Since the data were not quality checked or complete, they were used only to demonstrate the methodology. Application to the more complete Dongsha case study is shown in Chapter 3 (Results).

### **2.3.3. INDIVIDUAL BACKTRAJECTORIES**

Individual backtrajectories were first generated at the receptor for the period of interest. Backtrajectories were generated from the receptor site every three hours beginning at four different heights above the ground (100m, 500m, 1000m, 1500m), to examine potential source areas both within the boundary layer where the in-situ measurements take place and above the boundary layer where air parcels would not be expected to be as closely related to surface concentrations. The backtrajectories were run backwards for five days, a timeframe which includes more distant sources but limits the increasing uncertainty in location as backtrajectories are advected further from the receptor. Each backtrajectory for the Baghdad case study is shown in Figure 1 and colored according to the time the trajectory arrived at the receptor. The associated heights for each backtrajectory are shown in Figure 2. Trajectories that reach the top of the model domain are automatically truncated by HYSPLIT and are plotted as stars in Figure 2.



2008 531      2008 6 5      2008 610      2008 616      2008 621

Figure 1 Backtrajectories from 31 May to 21 June for a Baghdad receptor (33.26 N, 44.27 E), colored by date of arrival. In this and subsequent similar plots, the receptor site is indicated by a black star.

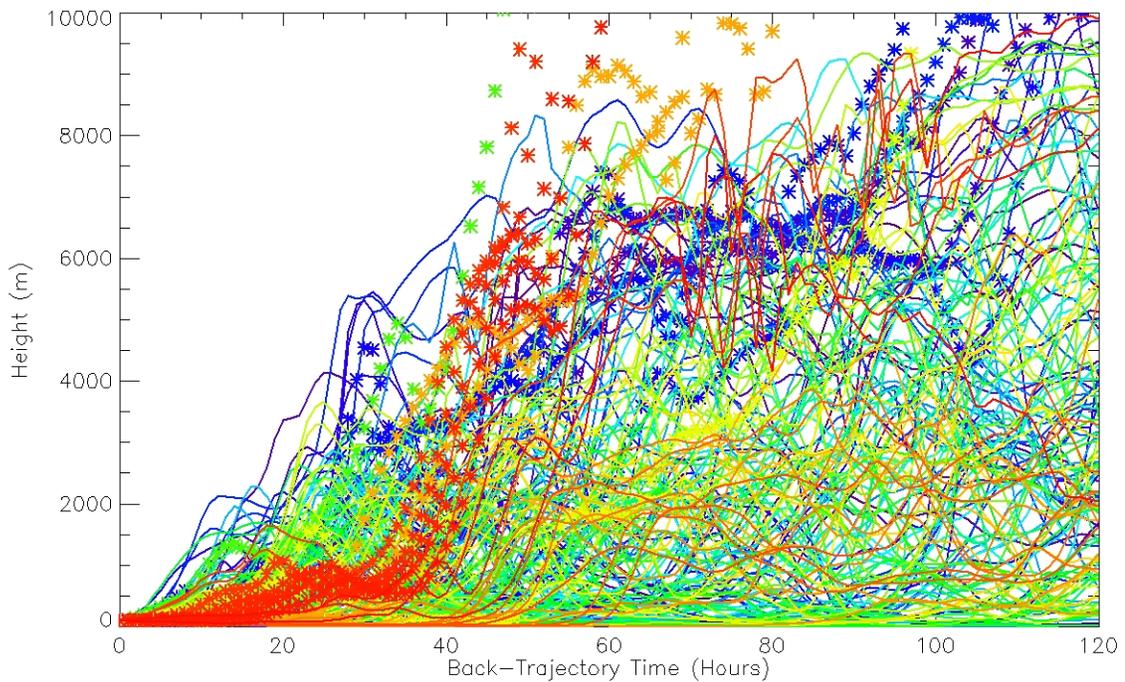


Figure 2 Heights of backtrajectories shown in Figure 1. Trajectories that reach the top of the model domain are truncated and are indicated by stars for the entire trajectory.

Trajectories can alternatively be colored to indicate trajectory height (Figure 3), age (Figure 4), latitude (Figure 5), and longitude (Figure 6).

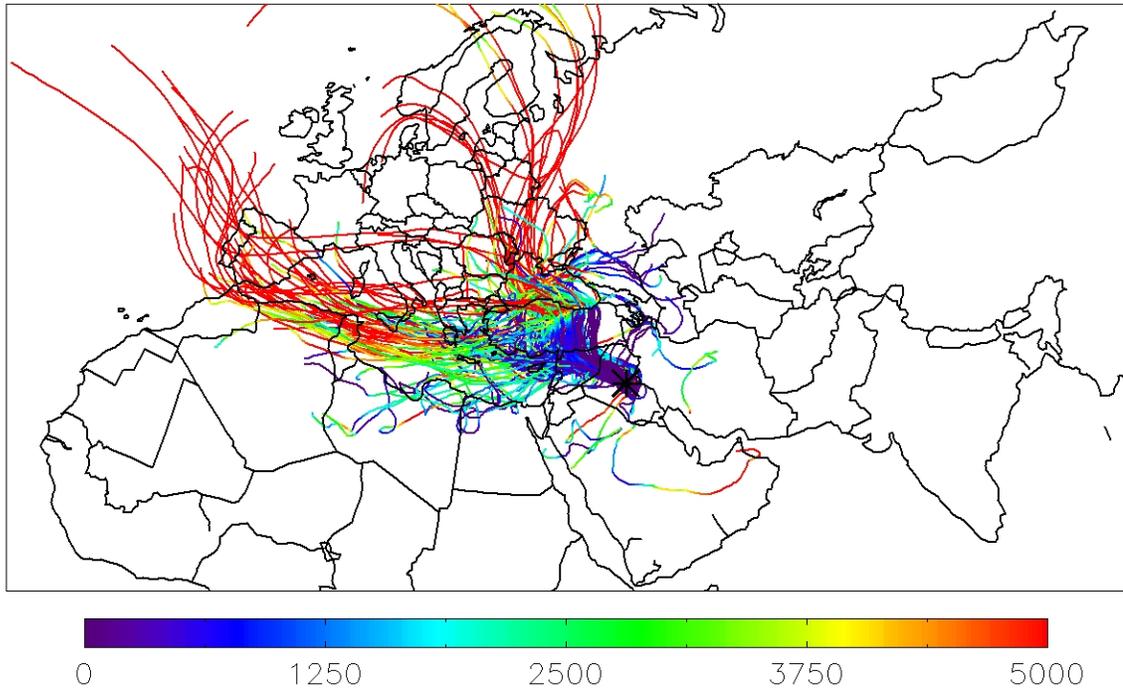


Figure 3 Backtrajectories in Figure 1, color-coded by height (m).

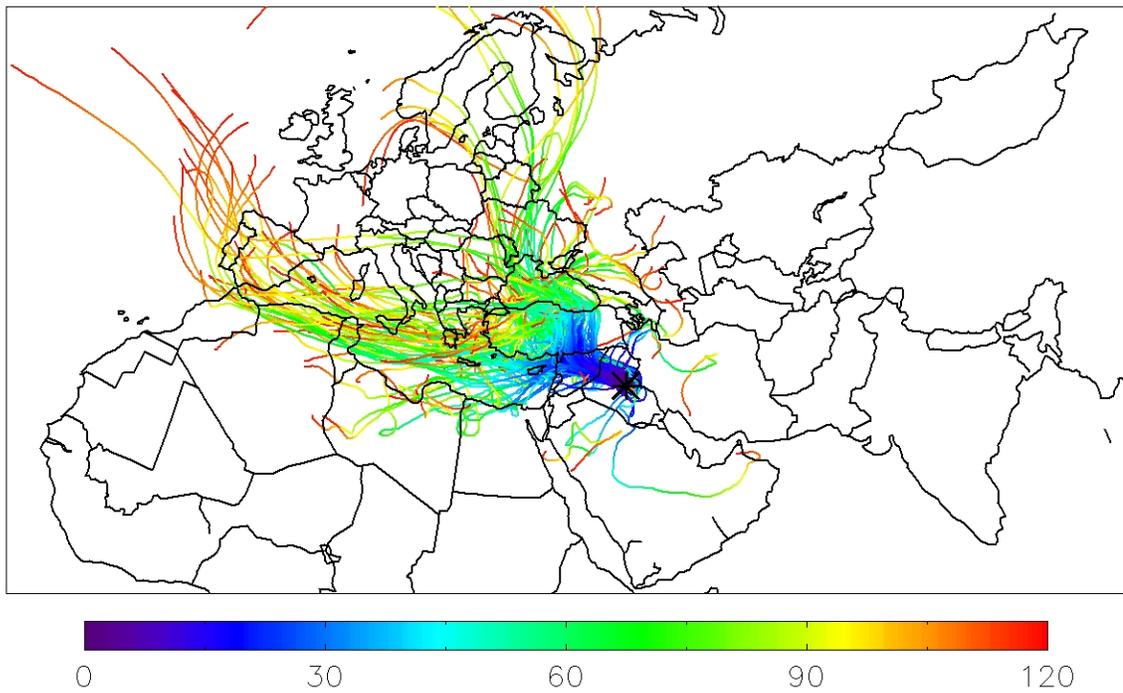


Figure 4 Backtrajectories in Figure 1, color-coded by age (hours).

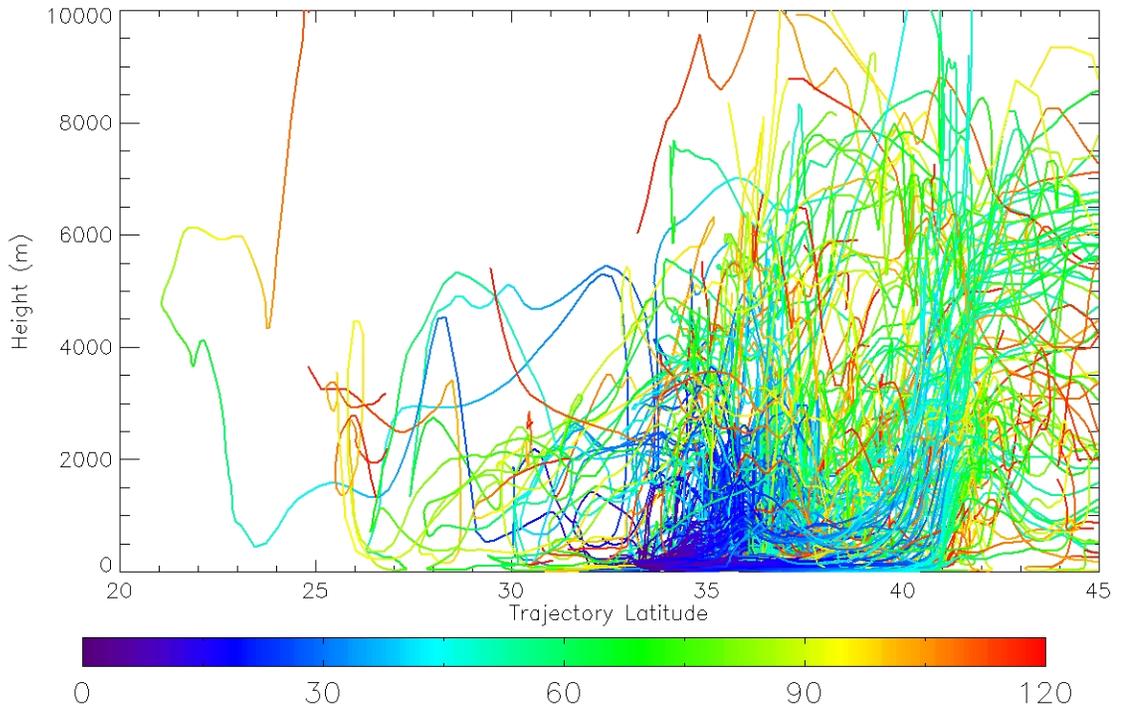


Figure 5 Backtrajectories in Figure 1, shown as a function of latitude and height (m), and color-coded by age (hours).

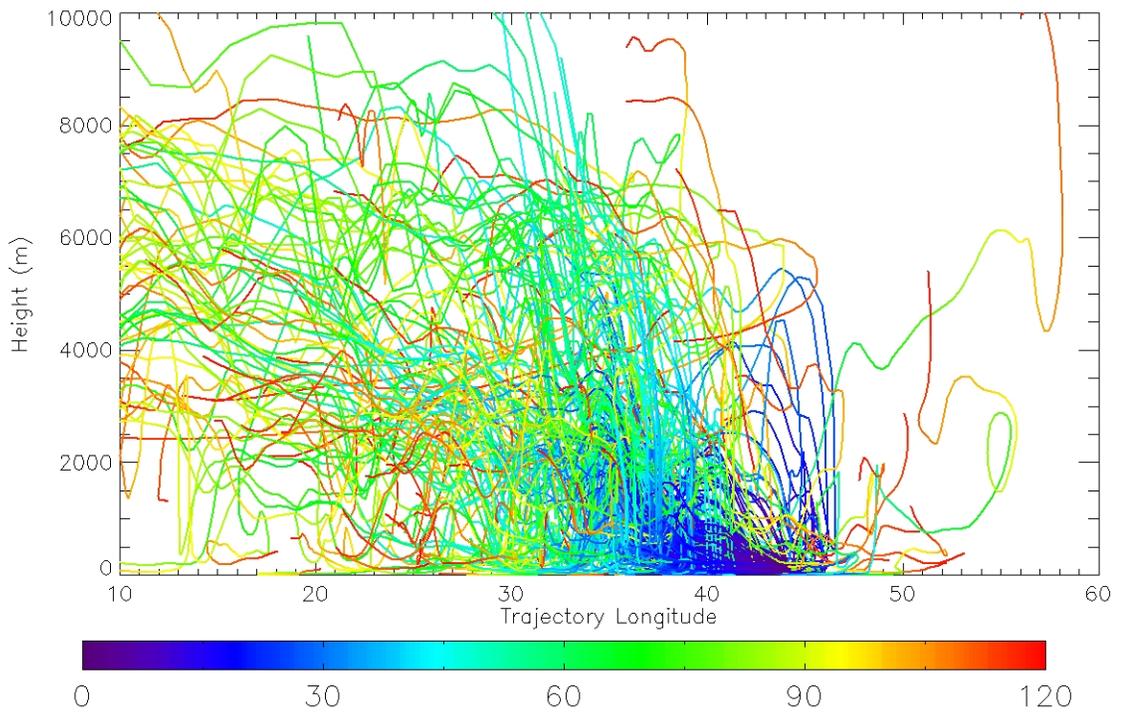


Figure 6 Backtrajectories in Figure 1, shown as a function of longitude and height (m), and color-coded by age (hours).

#### 2.3.4. RESIDENCE TIME ANALYSIS

Probabilistic estimates of the transport pathways of air masses arriving at the receptor over some time period are generated using Residence Time Analysis (RTA) following the methods proposed by *Ashbaugh et al.* (1985). The Residence Time is an indication of the relative amount of time air masses arriving at the receptor spent over all possible source areas. It is determined by first dividing the globe up into  $1^\circ \times 1^\circ$  grid boxes and then counting the number of backtrajectory endpoints in each box. The Residence Time is then calculated by:

$$P[A_{ij}] \cong \frac{n_{ij}}{N} \quad (1)$$

$n_{ij}$ : Number of endpoints in  $ij^{th}$  cell  
 $N$ : Total number of endpoints

The probability of event  $A_{ij}$ , that the air mass was bounded by the box  $ij$  at some point before its arrival at the receptor during the time period, is assumed to be well represented by the fraction of endpoints within that box. The Residence Time is therefore the fraction of end points in each grid box, and can be interpreted as the probability that an air mass reaching the receptor passed through the grid box during the time period in question. In this work, contour plots of the Residence Times are valued based on the percentage of end point counts in a given grid box relative to the maximum number of end points counts in any grid box on that plot. This is referred to as the Relative Residence Time Analysis:

$$RTA[A_{ij}] = \frac{P[A_{ij}]}{\max(P[A_{**}])} \quad (2)$$

Figure 7 shows the Relative Residence Times for Baghdad during the study period.

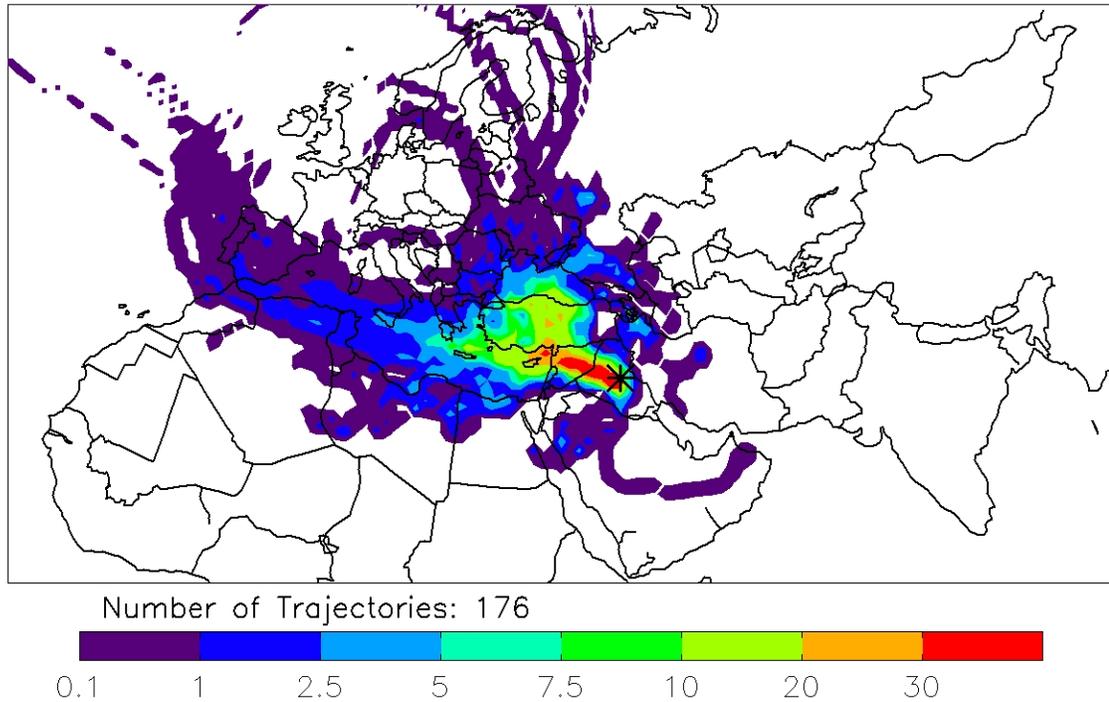


Figure 7 The Relative Residence Time Analysis plot for Baghdad during the study period.

### 2.3.5. SOURCE CONTRIBUTION FUNCTIONS

Grid boxes near the receptor feature prominently in the RTA plots due to the higher density of endpoints occurring near the site, where trajectories naturally converge, resulting in a peak in the plots near the receptor which can obscure the importance of source regions further away. Removing this peak allows for important source regions to be highlighted, ideally without regard for distance from the receptor site. *Ashbaugh et al.* (1985) describe a hypothetical probability distribution function consisting of a set of random trajectories that could come from any direction equally, and used to normalize the RTA to remove the “bullseye” around the receptor location. The Hypothetical Probability Distribution, shown in Equation 3, allows for the random event  $H_{ij}$  to be similar to the event  $A_{ij}$ , but created solely as a function of the hypothetical trajectory’s velocity and distance from the receptor. The choice of velocity and distance in a particular application are guided by average velocity of the trajectories and the modeled

time period. When incorporating the area of the grid box, which varies with latitude, the equation becomes:

$$P[H_{ij}] = \frac{\text{grid box area}}{2\pi Rr} \quad (3)$$

$$R = vT$$

*r*: mean distance from grid box to receptor

*v*: Trajectory velocity (assumed constant along trajectory path)

*T*: Maximum trajectory time (i. e. trajectory duration)

A Source Contribution Function (SCF), which indicates the likelihood that a region is a source for the receptor during the time period, independent of distance from the site, is defined by *Ashbaugh et al.* (1985) as:

$$SCF[A_{ij}] = \frac{P[A_{ij}]}{P[H_{ij}]} \quad (4)$$

For grid boxes where SCF[A<sub>ij</sub>] is greater than 1, the Residence Time is greater than would be expected by the chance hypothetical function, indicating a larger than random influence on the air mass reaching the receptor. The SCF for Baghdad is shown in Figure 8.

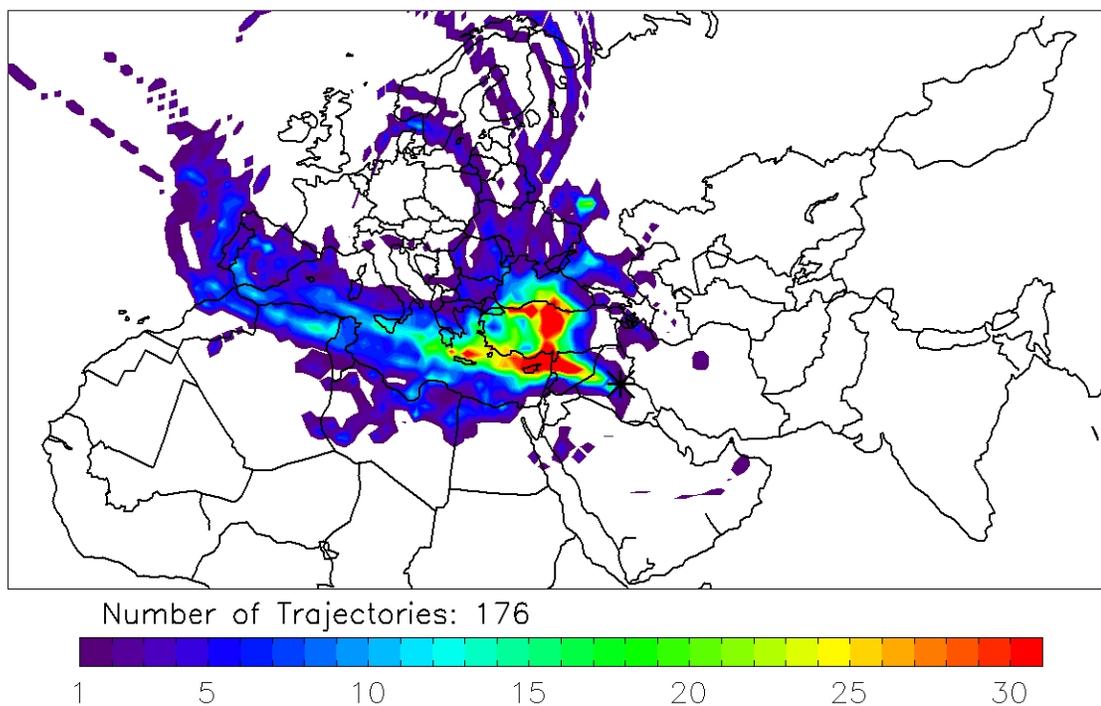


Figure 8 Source Contribution Function for Baghdad during the study period.

### 2.3.6. USE OF MEASURED IN-SITU DATA

The RTA and SCF are both representations of expected source regions for an air parcel that arrives at the receptor during the study period. In order to convey additional information about expected source regions for specific sources, in-situ measurements at the receptor are used to differentiate between individual backtrajectories. Following the terminology of *Ashbaugh et al.* (1985), the Normal Case is considered to be these analyses (RTA or SCF) applied to every backtrajectory within the study period, while the Special Case is the subset of the backtrajectories which conform to some criterion that is to be investigated. The criterion for the special case can be any characteristic which can be applied to each backtrajectory, such as a measured aerosol concentration above some cutoff value, or an observed atmospheric condition. For the purposes of this research, the DRUM elemental and factor timelines are used to form special cases.

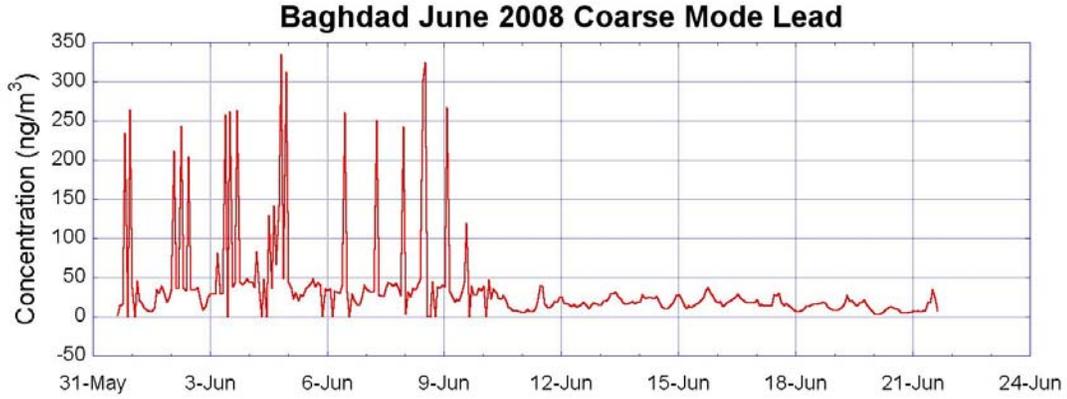


Figure 9 Baghdad DRUM timeline for Coarse and Fine Lead fractions in June 2008.

The sample Baghdad elemental lead (Pb) concentration for the summed coarse size fractions is shown in Figure 9. Using an arbitrary cutoff value of  $100 \text{ ng/m}^3$ , the backtrajectories that arrive at the receptor when a coarse Pb concentration is above this value are assigned to the Special Case. The event  $B_{ij}$  is defined as the instance of an endpoint existing within box  $ij$  from a trajectory that is part of the special case. The probability of event  $B_{ij}$  is therefore given by:

$$P[B_{ij}] \cong \frac{m_{ij}}{N} \quad (5)$$

$m_{ij}$ : Number of special case endpoints in  $ij^{\text{th}}$  cell

$N$ : Total number of endpoints

$$RTA[B_{ij}] = \frac{P[B_{ij}]}{\max(P[B_{**}])} \quad (6)$$

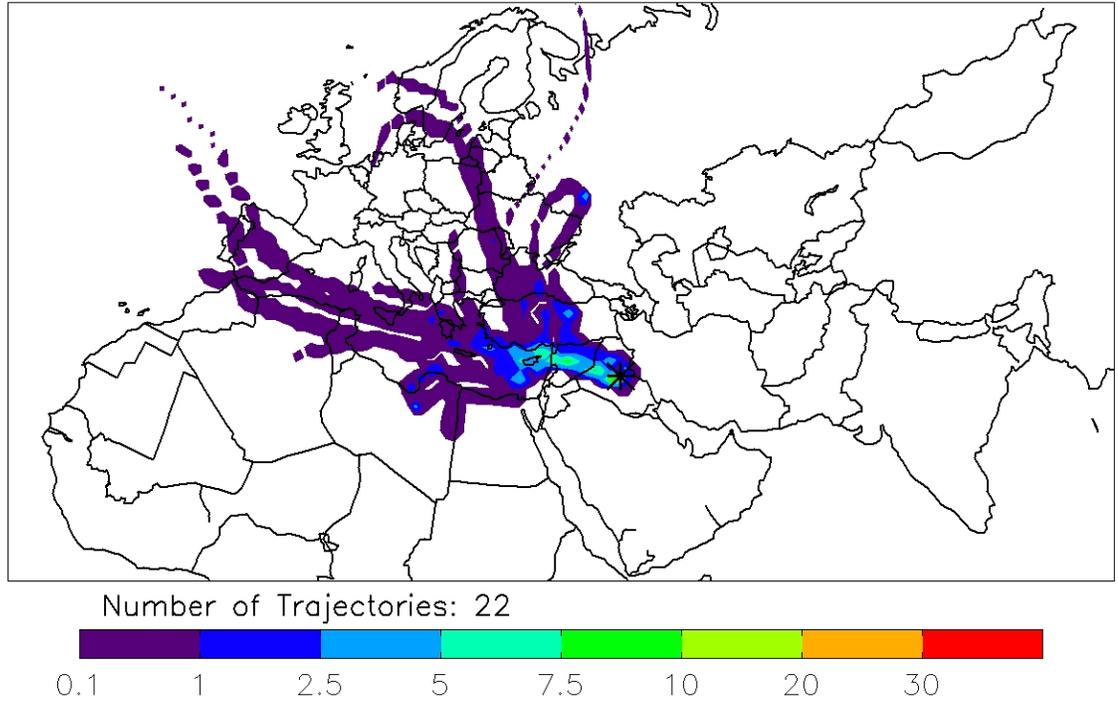


Figure 10 Special Case Residence Time Analysis for elevated coarse mode Pb.

This Special Case RTA is shown in Figure 10. The Normal and Special Case RTAs provide indications of, respectively, the Residence Times for all air parcels and for air parcels in which coarse lead concentrations are high. Since  $P[B_{ij}]$  is calculated by dividing  $m_{ij}$  by the total number of endpoints,  $N$ , the scale in both figures is the same. Therefore the difference between the value of the two functions for each grid box gives an indication of how the Special Case differs from the Normal Case. Alternatively, an additional analysis we refer to as the Specific Special Case RTA, shown in Figure 11, is calculated by:

$$P'[B_{ij}] \cong \frac{m_{ij}}{M} \quad (7)$$

$m_{ij}$ : Number of special case endpoints in  $ij^{th}$  cell

$M$ : Total number of special case endpoints

$$\text{Specific RTA}[B_{ij}] = \frac{P'[B_{ij}]}{\max(P'[B_{**}])} \quad (8)$$

The Specific Special Case RTA provides a representation of the special case trajectories that is calculated by the same technique as the normal case. The specific

residence time is calculated based on the fraction of endpoints in only the special case trajectories, giving a similar estimate of the probability that a special case air mass over a given region without regard for trajectories not included in this case.

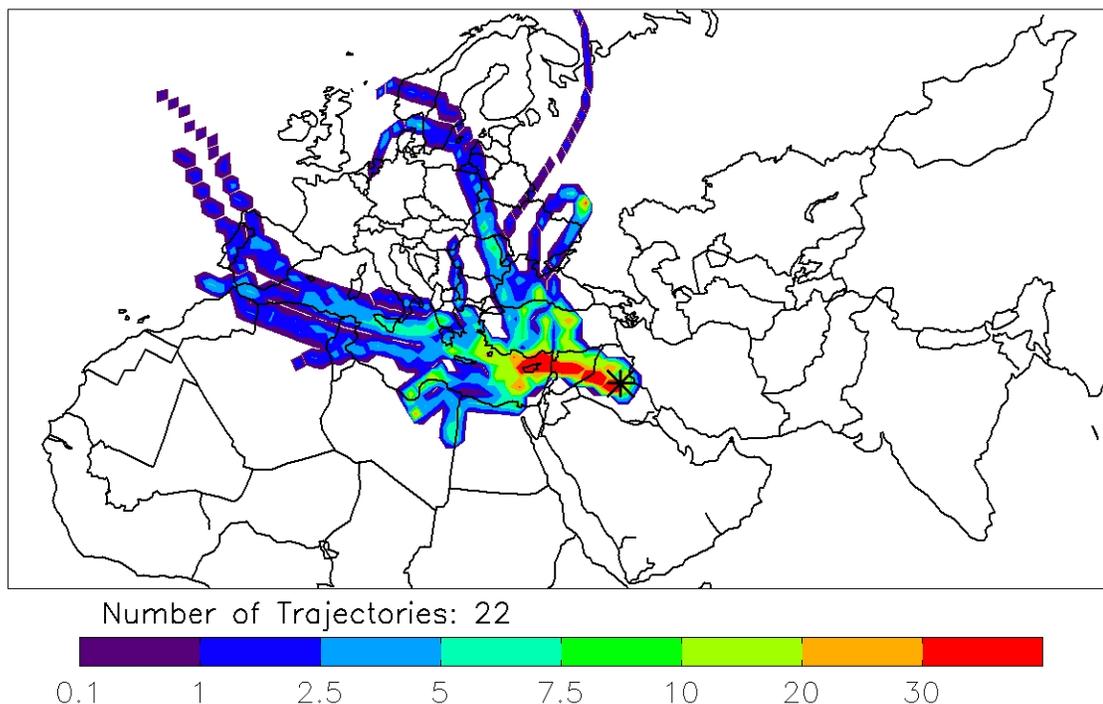


Figure 11 Specific Special Case Residence Time Analysis for elevated Coarse mode Pb.

The use of relative Residence Times in this work is intended to allow for reasonable scales and to emphasize that the difference in value between grid boxes is the characteristic which has the most utility. In comparing the Special Case RTA (Figure 10) to the Normal Case RTA (Figure 7), the difference between contour levels indicates how much of the Normal Case is accounted for in the Special Case. Similarly the Specific Special Case RTA (Figure 11) represents an equivalent analysis to the Normal Case RTA, so that the important features can be compared at the same relative scale.

Similarly, a Special Case SCF and Specific Special Case SCF (Figure 12) can be created to remove the same inherent peak surrounding the receptor in the RTA figures.

This is likewise calculated by:

$$SCF[B_{ij}] = \frac{P[B_{ij}]}{P[H_{ij}]} \quad (9)$$

or

$$SCF'[B_{ij}] = \frac{P'[B_{ij}]}{P[H_{ij}]} \quad (10)$$

Note that the only difference between these alternative calculations of the Special Case SCFs is in the scale of the result. In both cases, the difference between any two grid boxes within the figure carry the primary information, namely which areas are more important than others in terms of their contribution to high concentrations of coarse mode lead aerosol at the receptor.

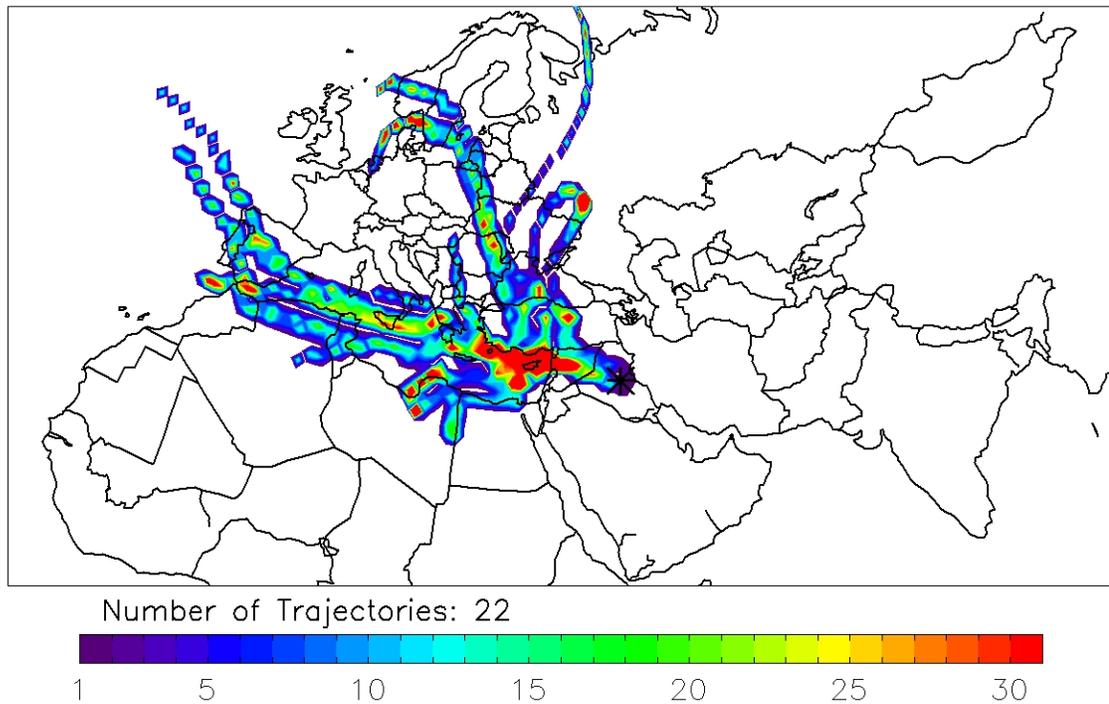


Figure 12 Specific Special Case Source Contribution Function for elevated Coarse mode Pb.

### 2.3.7. COMPARISON OF THE NORMAL AND SPECIAL CASES

Establishing the various RTA and SCF functions for the Normal and Special Cases allows for analysis of important source regions for the receptor. The Special Case,

however, was established in order to better understand which of these potential source regions is more likely to contribute to the environmental conditions that comprise the Special Case subset. In the Baghdad example case, the hypothesis is that some of the potential source regions identified in the Normal Case are more likely to be sources of coarse mode Pb. Testing this hypothesis requires differentiating between the Special and Normal Case RTA or SCF functions.

The null hypothesis in this case would be that the Special and Normal Cases are not, in fact, different from each other. Such a finding could result from the meteorological model being of too coarse spatial or temporal resolution to resolve the Pb variability. Alternatively, the source of Pb could be so close to the receptor that large-scale synoptic variability does not control the variability seen in the Pb concentrations. Pb emissions might also be variable in time. In any of these cases, the selection of Special Case backtrajectories had little to do with the meteorology that HYSPLIT utilizes. Under such conditions, HYSPLIT cannot explain any part of the variance in Pb concentrations, and the RTA or SCF function would therefore be expected to be highly similar for both the Normal and Special Cases.

In contrast, if it were true that there was only one source of coarse mode Pb that was emitted continuously from a location well removed from the receptor, it would be valid to assume that synoptic scale meteorology would have a significant influence on observed receptor concentrations. In this instance, the Special Case RTA and SCF would assign higher frequency of endpoint occurrence to grid boxes within the Pb source area, while other areas would have lower endpoint frequencies and a resulting lower special case residence time.

In order to test the hypothesis that HYSPLIT trajectories can explain a portion of the variance in the in-situ measured data, it is therefore necessary to calculate a function which analyzes the difference between Special and Normal Case functions. [Ashbaugh *et al.*, 1985] suggested such a function that they called the Conditional Probability Function, which looks at the fraction of normal case endpoints within each grid box that are part of the special case. This same function has more recently been referenced as the Potential Source Contribution Function (PSCF) [Hopke, 2003; Polissar *et al.*, 1999]; we will use this more common terminology within this manuscript.

$$P[B_{ij}|A_{ij}] = \frac{P[B_{ij} \cap A_{ij}]}{P[A_{ij}]} \quad (11)$$

Equation 11 gives the probability that the event  $B_{ij}$  (a special case endpoint exists in some box  $ij$ ) will occur given that the event  $A_{ij}$  (a normal case endpoint exists in that same box) occurs as well. Since  $B_{ij}$  is a subset of the set of events  $A_{ij}$ , the equation simply becomes:

$$PSCF_{ij} = P[B_{ij}|A_{ij}] = \frac{P[B_{ij}]}{P[A_{ij}]} = \frac{m_{ij}/N}{n_{ij}/N} = \frac{m_{ij}}{n_{ij}} \quad (12)$$

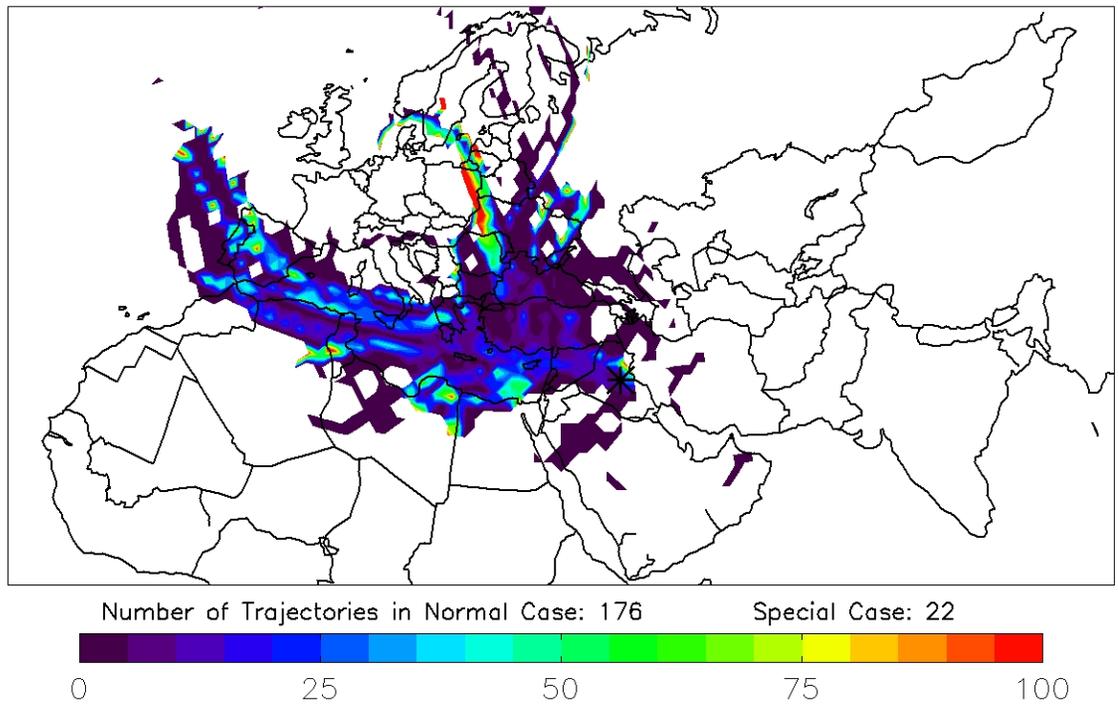


Figure 13 Potential Source Contribution Function (%) for elevated Coarse mode Pb.

The PSCF is therefore the ratio of Special Case to Normal Case endpoints within each grid box. As can be seen in the PSCF for the example Baghdad case study (Figure 13), several areas with the highest PSCF are at the extreme edges of the potential source areas. This result is likely due to the small number of endpoints (if  $n_{ij}$  is small, a small  $m_{ij}$  could still yield a high PSCF) in these areas rather than to actual sources which contribute to the high concentration events. A minimum endpoint criterion or weighting function can be utilized to reduce the impact of this effect on the plots. In Figure 14, a minimum endpoint criterion is applied, leaving only grid boxes with a minimum of 10 endpoints in the Normal Case to be plotted.

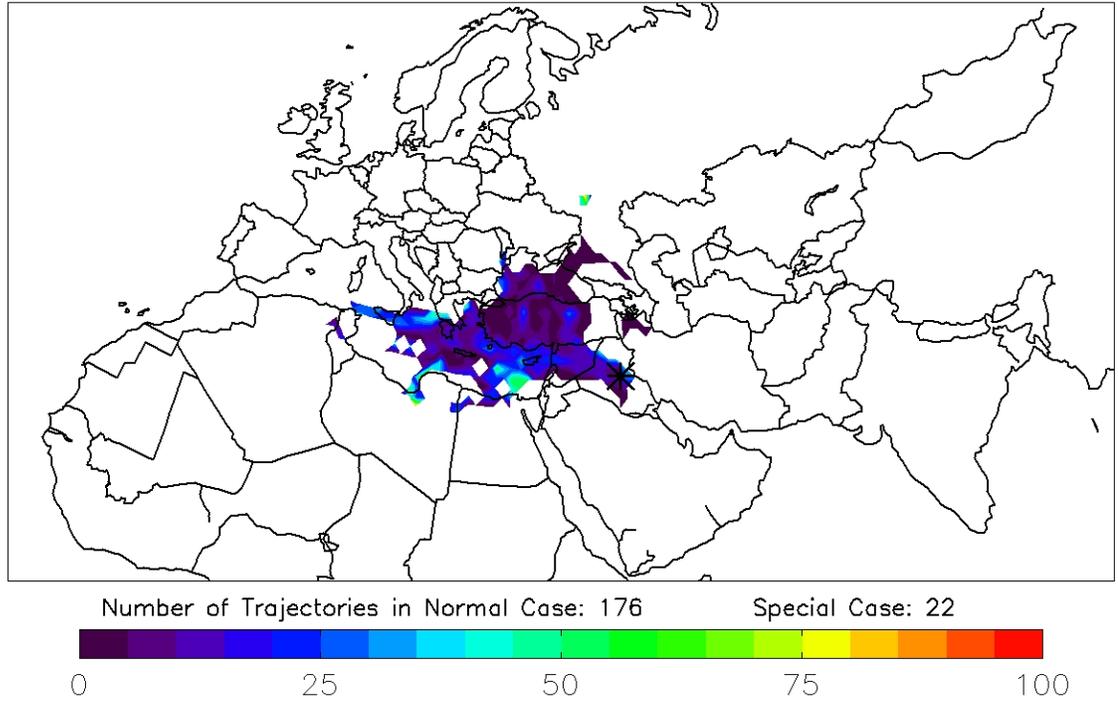


Figure 14 Potential Source Contribution Function (%) with a 10 endpoint minimum criterion for elevated coarse mode Pb.

A similar function which utilizes the difference between  $n_{ij}$  and  $m_{ij}$  rather than the ratio was referenced without specific by *Poirot et al.* (2001) as part of Residence Time Analysis. We propose referring to this refinement of RTA as the Residence Time Difference Function (RTDF), which identifies the differences between the Normal Case RTA and the Special Case RTA by:

$$RTDF_{ij} = P'[B_{ij}] - P[A_{ij}] = \frac{m_{ij}}{M} - \frac{n_{ij}}{N} \quad (13)$$

Likewise for the SCF, the Source Contribution Difference Function (SCDF) gives a similar analysis of a normalized difference between the two cases:

$$SCDF_{ij} = SCF'[B_{ij}] - SCF[A_{ij}] = \frac{P'[B_{ij}] - P[A_{ij}]}{P[H_{ij}]} = \frac{m_{ij}/M - n_{ij}/N}{P[H_{ij}]} \quad (14)$$

The RTDF function is effectively comparing the probability values between the normal and special cases on an absolute difference basis, rather than the ratio basis of the PSCF. The equation is conditional on the Specific Special Case RTA and the Normal

Case RTA both being valid representations of the important source regions for the Special and Normal Case sets of backtrajectories, respectively. This implies that each set must contain enough backtrajectories to adequately represent actual source region probabilities. The SCDF differs from the RTDF only in that it analyzes the difference between the normal and special cases on a distance normalized basis (i.e. the HDF is used to remove the bias of endpoints to pass through grid boxes near the receptor).

As can be seen in Figure 15, roughly the same areas are highlighted by both PSCF and RTDF analysis, however, some differences emerge. As noted by *Poirot et al.* (2001), the RTDF is based on absolute difference, so grid boxes with more overall endpoints are emphasized. The PSCF will rate grid boxes based solely on the ratio of the number of Special to Normal Case endpoints, so the absolute number of endpoints has no impact. A grid box far removed from the receptor with a value of (10 Special Case endpoints / 20 Normal Case endpoints) will be valued the same as a nearby grid box with a value of (500 Special Case endpoints / 1000 Normal Case endpoints). It can be imagined that there are some circumstances when this closer source, which impacts the receptor more often and at higher concentrations, should be emphasized. The RTDF allows for this type of analysis.

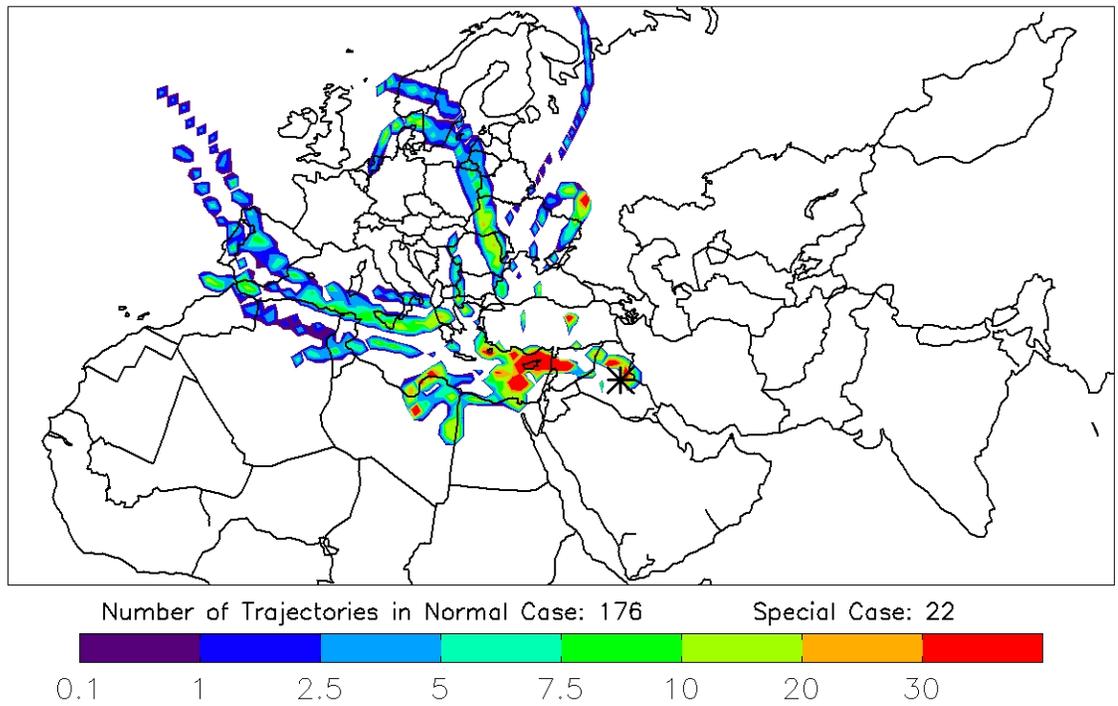


Figure 15 Residence Time Difference Function for elevated coarse mode Pb.

The SCDF is also shown in Figure 16 showing the de-emphasis of sources closer to the receptor as is characteristic of SCF plots.

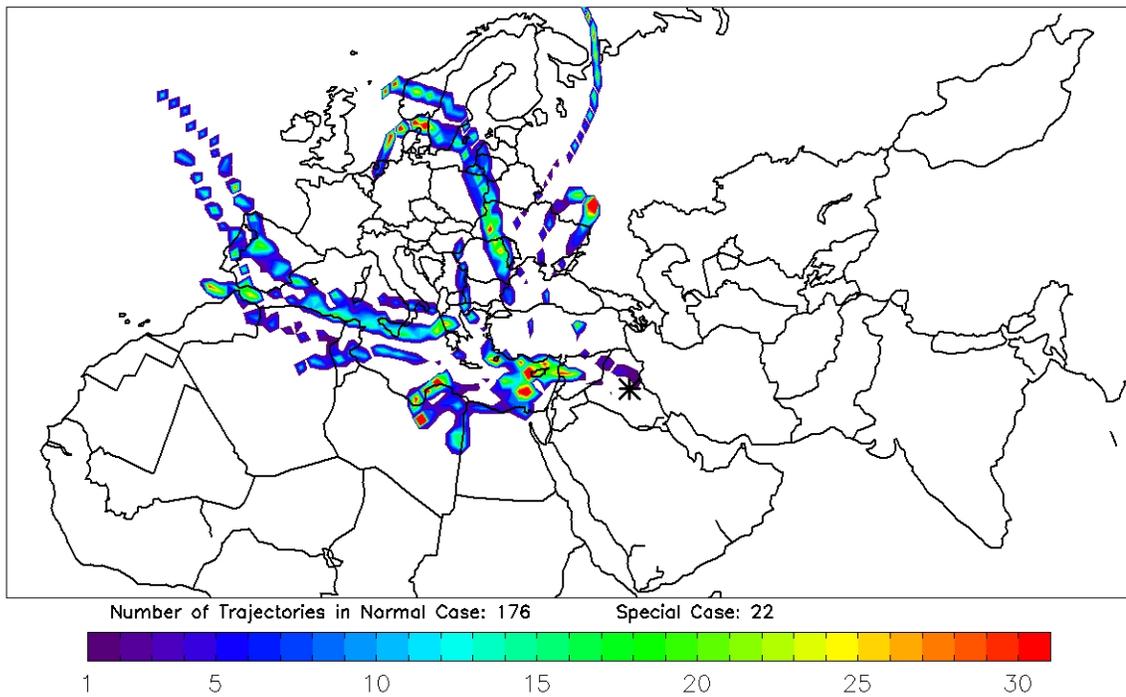


Figure 16 Source Contribution Difference Function for elevated coarse mode Pb.

So-called leading and trailing effects should be noted in the SCF, PSCF and RTDF/SCDF functions, identified by *Poirot et al. (2001)* and *Polissar et al. (2001b)*. The nature of synoptic scale meteorology is such that common patterns of long-range transport to a receptor often emerge. Thus, an air parcel which passes over some point, A, before reaching the receptor, will often also pass over some point, B, in between the two. If a major source being analyzed for the Special Case is located at B, many of the backtrajectories which pass over B will also have passed over A. This will have the effect of emphasizing point A in the Special Case despite there being no actual source there. Such a case is an example of a trailing effect, with a leading effect yielding the same result, but with the actual source at A rather than B. Long trails in the Special Case may therefore simply be the result of leading or trailing effects.

A final point to make deals with the representation of these functions when plotted. While smooth contours have been shown in Figures 7 - 16, in fact, the actual data are better represented by colored grid boxes. The same data from Figure 15 are plotted on a gridded domain in Figure 17. Contoured plots may incorrectly assign a gradient when in fact a single high value grid box is all that exists.

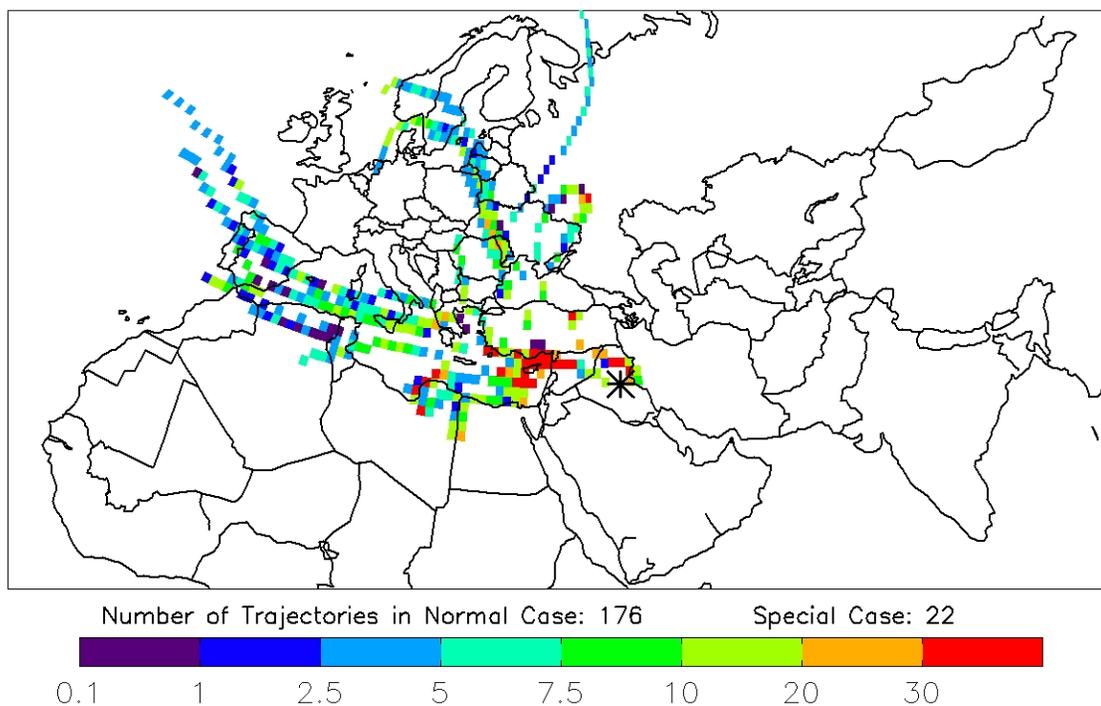


Figure 17 Residence Time Difference Function plotted on a gridded domain.

## 2.4. ADDITIONAL SUPPORTING MEASUREMENTS

Two methods have been described to this point that are used in the analysis of the source/receptor relationship: Factor Analysis and Trajectory Modeling. These methods represent relatively independent approaches for assessing sources of aerosols observed at a receptor site. The first method utilizes chemical and physical analysis to identify aerosol types, while the second uses meteorological information from a trajectory model to compare potential source areas dominant at different times. We used additional model output and observation data in this work to support the overall identification of aerosol types and sources.

### 2.4.1. NAAPS MODEL

Output from the Navy Aerosol Analysis and Prediction System (NAAPS) is used in this work to qualitatively compare observed aerosol plumes to modeled plumes. NAAPS

is a global  $1^\circ \times 1^\circ$  aerosol transport model maintained operationally by the U.S. Navy to support various operations and research [Reid *et al.*, 2009]. Aerosol emissions and sinks are simulated, along with global sulfate, smoke, and dust aerosol transport, allowing for predicted optical depths and concentrations at desired locations and heights to be extracted. Aerosol optical depths from satellite observations are assimilated into the analyses. Specific aerosol events, spikes, and variability seen in measured data are qualitatively compared (timing and magnitude) against NAAPS concentrations from a grid point containing the receptor site.

#### **2.4.2. ADDITIONAL DONGSHA IN-SITU MEASUREMENTS**

In situ measurements from the TSI three-wavelength nephelometer were described in Section 2.3.1. Data from an AERONET radiometer located at Dongsha Island (20.70 N, 116.73 E, 5 m ASL) (Dongsha Island AERONET PI George Lin, nhlin@cc.ncu.edu.tw). The daily average Level 2.0 Spectral Deconvolution Algorithm (SDA) Version 4.1 dataset was used to separate fine and coarse mode contributions to AOD at 500 nm, and is available at [http://aeronet.gsfc.nasa.gov/data\\_menu.html](http://aeronet.gsfc.nasa.gov/data_menu.html) [Holben *et al.*, 2001; O'Neill *et al.*, 2003].

### **3. RESULTS AND DISCUSSION**

The 7 SouthEast Asian Studies (7SEAS) Campaign [*7SEAS-Whitepaper*, March 4, 2009] is intended to characterize aerosol properties and transport in Southeast Asia. Dongsha Island serves as one of two supersites, along with Singapore, from which numerous measurements are to be taken throughout the study. During the Spring of 2010, a variety of instruments were deployed, including a DRUM sampler, TSI three-wavelength nephelometer, and meteorological instrumentation. The goal was to form an initial understanding, prior to the main 7SEAS campaign, of the nature, sources, and transport patterns for aerosol observed at Dongsha Atoll as they relate to the larger South China Sea environment during springtime.

Southeast Asia remains an area of ongoing research into atmospheric aerosol due to a relative lack of study compared to other regions in the world, and the significant impact aerosol has in the region. Health impacts from biomass burning ranging on the scale of small cook stoves to large agricultural and forest fires are known to exist, while the potential for changes in visibility, precipitation and cloud formation, and climate impacts are important in this region. The results of the study are therefore intended to support ongoing efforts to understand these impacts by a number of stakeholders in the region.

### **3.1. SITE DESCRIPTION AND BACKGROUND**

The Dongsha Islands (also known as the Pratas Islands) are a small group of low lying islands forming an atoll located in the Northern part the South China Sea (SCS) (Figure 18; Lat=20.7 N; Lon=116.7 E). There are no indigenous inhabitants located on the islands, however, a small population exists for research, fishing and military purposes. The island is powered by a diesel generator and an airstrip hosts flight operations once or twice per week. These potential pollutant sources are outside the prevailing wind directions of the receptor site. As a result, Dongsha is considered to be generally independent of local sources, with some possible exceptions when meteorological or other conditions allow for it, and therefore can be a suitable location for study of aerosol transport into the SCS.

The Dongsha Experiment field collection occurred from mid-March through mid-May, 2010. The DRUM aerosol impactor was operated continuously for approximately five weeks without interruption between March 30th and May 9th, 2010. The DRUM was collocated with the NASA COMMIT aerosol trailer which provided support. A permanent weather station provides information on local wind conditions.



Figure 18 Dongsha Island location.

### 3.1.1. METEOROLOGY

Transport patterns and winds in the South China Sea (SCS) are known to vary with height [Satake *et al.*, 2004; Takemi *et al.*, 2006], with periods of significant low-level vertical wind shear. Onset of the East Asian summer monsoon in the SCS typically occurs in early to mid-May, bringing low-level westerlies and air masses originating from Southeast Asia and the Maritime Continent over the measurement site. Sampling for this study occurred before the onset of the monsoon, with low-level winds predominately from the Northeast, originating from China and East Asia. Mid-level westerlies dominated above several thousand meters during the experiment, bringing air parcels from Southeast Asia and the southern SCS.

### **3.1.2. EXPECTED SOURCE REGIONS**

Backtrajectory analysis using HYSPLIT was conducted for 2009 and 2010. Source Contribution Functions (Figures 19 and 20) show the seasonal variation in transport patterns into the SCS at several levels. The monsoonal shift in summer is easily seen in Figure 19, showing expected source regions for low-level arrival heights. In contrast, Figure 20 shows significantly different expected source regions for 1500 m arrival heights, particularly in the winter and spring. This effect, a decoupling of transport in the boundary layer from the free troposphere, results from frequent low-level inversions and significant low-level vertical wind shear in the region, and helps explain the reported vertical inhomogeneity of the atmospheric environment in the SCS. In addition, the year to year variability in the back trajectory analysis between 2009 and 2010 was considered small, with much smaller impact on potential source areas than changes in the height or season. Consistent with expectations based typical winds noted in the previous section, the results of backtrajectory analysis therefore indicate that during the Dongsha Experiment transport from costal China is expected to dominate near the surface, while air masses above roughly 1500 m altitude should be primarily from Southeast Asia, the Maritime Continent, and marine regions.

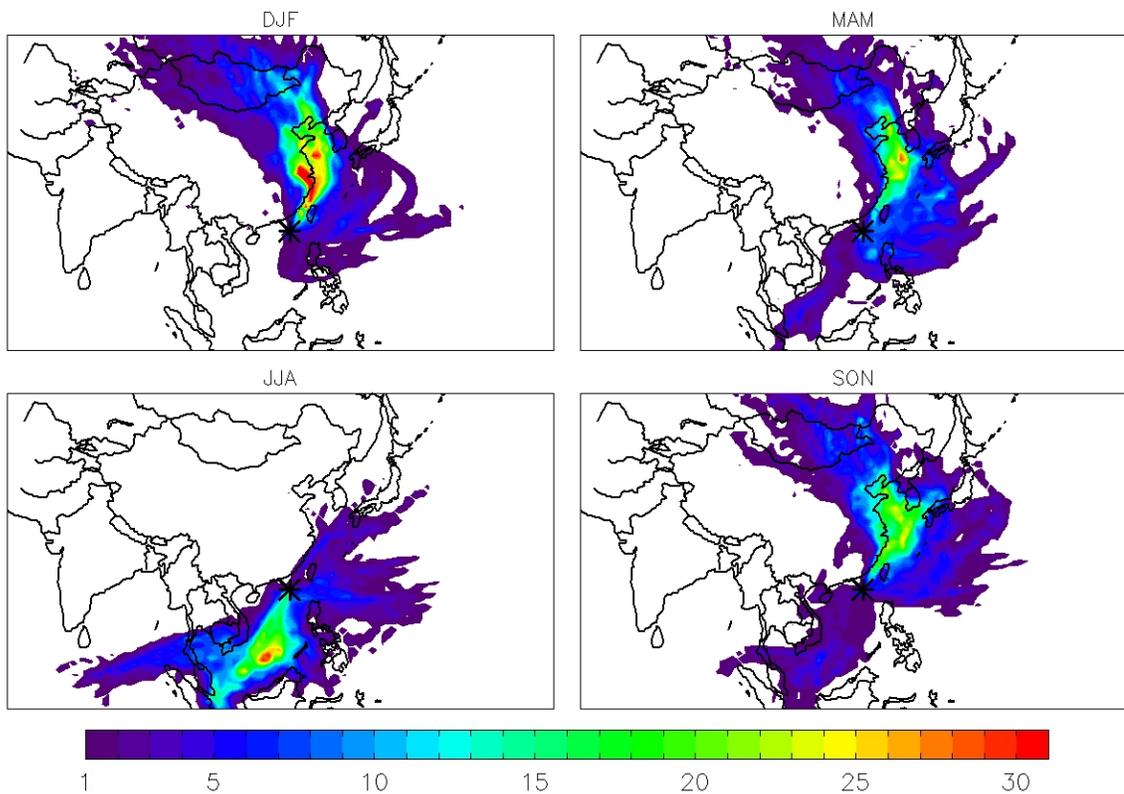


Figure 19 Dongsha seasonal SCF plots for 2009 to 2010 at 100m receptor heights.

*S H Wang et al. (2011)* investigated a significant low-level dust plume which impacted Dongsha shortly before the period of this study on 21 March 2010, which they attributed to low level transport over the SCS from the deserts of inland China and Mongolia. *Cohen et al. (2010a;c)* found similar source regions for dust blowing into Hanoi, Vietnam, along with pollution from local and Chinese sources, while *Han et al. (2005)* identified a variety of natural and anthropogenic sources impacts Gosan, Korea.

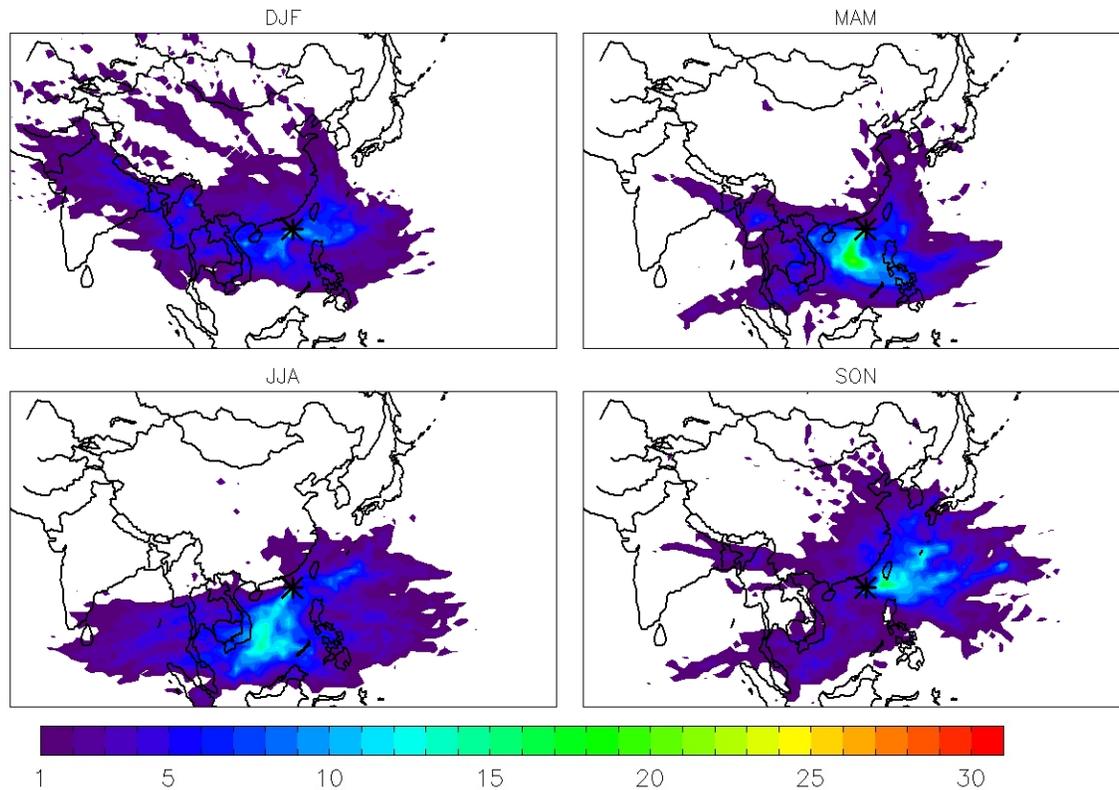


Figure 20 Dongsha seasonal SCF plots for 2009 to 2010 at 1500m receptor heights.

### 3.1.3. LOW-LEVEL INVERSION AND ATMOSPHERIC STABILITY

The low-level wind shear noted in the previous section indicates the potential for different source regions for air masses arriving at various heights over Dongsha island. A stratified atmosphere would have important implications for the representativeness of surface measurements to the wider environment in the region. In addition to aerosol from divergent source regions being advected into the column at different heights as a result of wind shear, the stability of the atmospheric column itself can be a useful indicator of the potential for stratification and heterogeneity by height. Figures 21 and 22 show the temperature and lapse rate profiles (six grid boxes between 17.5N – 21.5N; 115E – 120E) for the northern SCS from the NCEP Reanalysis meteorological dataset (provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site

at <http://www.esrl.noaa.gov/psd/>) [Kalnay *et al.*, 1996] for the time period of the Dongsha Experiment between 31 March and 9 May, 2010.

While actual temperature inversions are rare in this dataset, this may be due to the relatively sparse temperature data that is reported at only standard pressure levels on a 2.5° x 2.5° grid. Inversions occurring over a more narrow vertical range may still effectively cap mixing and lead to a more stratified atmospheric column. Soundings released from coastal mainland to the west were regularly stronger than is suggested by this dataset.

Despite the limitations of this dataset, it is evident that stable environments associated with low environmental lapse rates occurred often at low levels near the boundary layer in the northern SCS. Such stable atmospheric columns may develop from subsidence aloft, warming the air as it descends and limiting the amount of vertical mixing which takes place. These types of events would tend to support the development of stratified layers within the atmospheric column, and may impact the concentrations of aerosol measured at the surface.

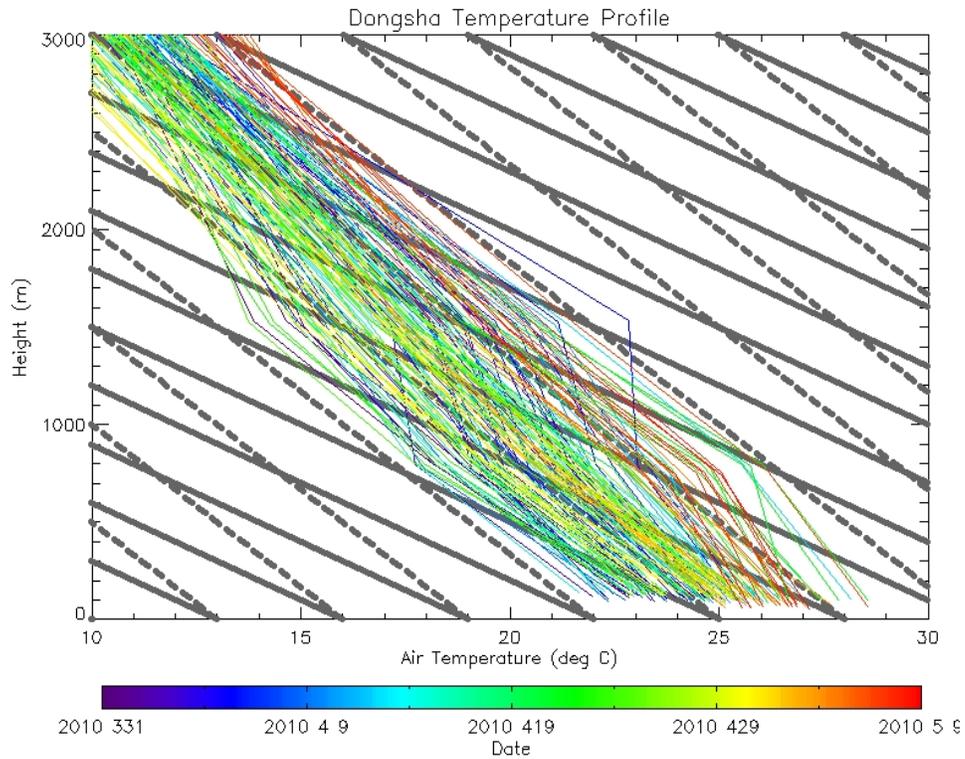


Figure 21 NCEP Reanalysis temperature profile for lowest 3000 m of only the grid box containing Dongsha Island. Dry Adiabatic Lapse Rate (solid gray) and Saturated Adiabatic Lapse Rate (dashed gray) are shown for comparison.

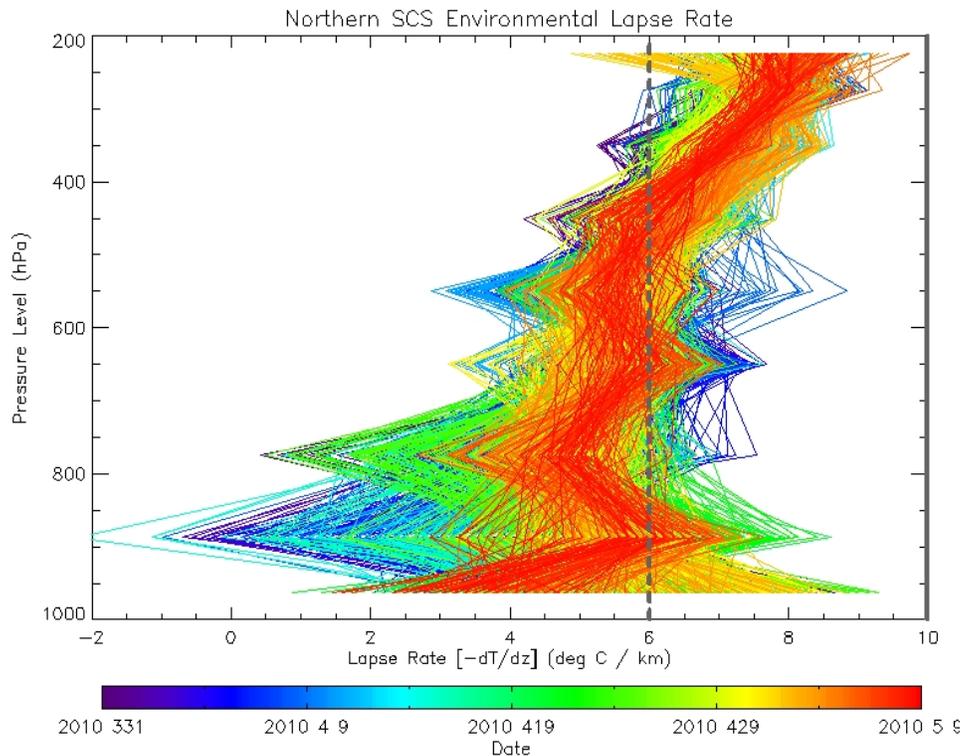


Figure 22 NCEP Reanalysis Lapse Rate profile at standard pressure levels for grid boxes in the northern South China Sea area during the Dongsha Experiment. DALR and SALR shown in gray.

## **3.2. DRUM/XRF RESULTS**

DRUM data collected at Dongsha during the experiment and analyzed by XRF are presented below. As the focus of this research is on use of the DRUM instrument for source apportionment at Dongsha, other measured data were primarily used to assist with interpretation and reduce uncertainty in the use and application of DRUM data. We describe methods used to sequester the sources of bias resulting from time stamp uncertainty, and particle size uncertainty. Matrix effects were considered and it was determined that no correction was required (See Appendix A).

### **3.2.1. TIME STAMP CORRECTION**

During XRF analysis, an estimated time stamp is assigned to each data point for the DRUM sample strips based on start and end times noted in the field log. As described in the analysis section, the raw DRUM time stamp for each size resolved sample strip was first cross-correlated to each of the other size fraction strips to correct any relative time stamp discrepancies between strips. Table 2 shows the peak cross-correlation between all size fractions for selected individual elements and a summed total elemental concentration. In this calculation, the cross-correlation is Pearson correlation coefficient between the size fraction timelines, one of which is offset in either direction by one or more time steps (the lag or lead). If the peak cross-correlation occurs when one time series is offset in time, it may be evidence that the sample strips are not aligned correctly.

The cross-correlation results show some indication that certain size fractions may be offset by several data points. In particular, size fraction 5 appears to lead other size fractions by somewhere between 2 and 6 data points, while size fraction 1 appears to lag other fractions by 1 to 3 data points. Figure 23 shows the time series of size fraction

concentrations. Major peaks for size fraction 5 appear to be well aligned with other size fractions in the iron and lead time series. In addition, the periods of minimal detectable concentrations for all size fractions, including size fraction 1, seem to coincide with each other.

The offset indicated by the cross-correlations therefore are not supported by a simple visual inspection of the time series for each size fraction. Since none of the cross-correlations for the size fractions showed a consistent lead or lag in all elements or against all size fractions, and since none of these offsets were evident in the visual inspection of the time series, it was determined that no adjustment of the DRUM time stamps was warranted.

Table 2 Dongsha DRUM size fraction cross-correlation. The peak cross-correlation between each size fraction (e.g. S1 for size fraction 1) for selected elemental and summed mass concentrations are shown.

Total Mass by Size Fraction									S by Size Fraction									
Lag/Lead for Maximum Cross Correlation									Lag/Lead for Maximum Cross Correlation									
Correlation Value at this Lag/Lead									Correlation Value at this Lag/Lead									
Var	S1	S2	S3	S4	S5	S6	S7	S8	Var	S1	S2	S3	S4	S5	S6	S7	S8	
S1	-	*	*	0	-1	**	-1	*	S1	-	-2	-1	-1	-1	-6	*	*	
				0.22	0.25		0.53				0.56	0.34	0.53	0.17	0.37			
S2		-	-1	-1	**	**	*	*	S2		-	0	-1	1	-3	0	0	
			0.54	0.62								0.48	0.48	0.22	0.36	0.23	0.18	
S3			-	0	**	0	*	*	S3			-	0	-1	0	0	-1	
				0.55		0.51							0.47	0.16	0.41	0.42	0.42	
S4				-	**	-1	*	*	S4				-	**	-4	0	-1	
						0.41									0.48	0.44	0.46	
S5					-	4	0	0	S5						-	**	0	*
						0.37	0.19	0.12									0.18	
S6						-	0	1	S6							-	-2	-6
							0.15	0.12									0.37	0.33
S7							-	-1	S7								-	-1
								0.65										0.66
S8								-	S8									-
Fe by Size Fraction									Pb by Size Fraction									
Lag/Lead for Maximum Cross Correlation									Lag/Lead for Maximum Cross Correlation									
Correlation Value at this Lag/Lead									Correlation Value at this Lag/Lead									
Var	S1	S2	S3	S4	S5	S6	S7	S8	Var	S1	S2	S3	S4	S5	S6	S7	S8	
S1	-	-1	-3	-2	**	-2	*	*	S1	-	*	*	*	*	*	*	*	*
		0.48	0.23	0.12		0.12												
S2		-	0	0	-2	-1	0	1	S2		-	0	1	-3	*	*	**	
			0.84	0.74	0.65	0.52	0.37	0.48				0.63	0.45	0.30				
S3			-	1	-2	0	0	1	S3			-	1	-4	*	0	**	
				0.95	0.57	0.76	0.65	0.66					0.82	0.27		0.17		
S4				-	-4	-1	0	0	S4				-	-4	-1	-1	0	
					0.48	0.82	0.69	0.72						0.39	0.19	0.33	0.30	
S5					-	3	5	6	S5					-	3	4	6	
						0.44	0.35	0.37							0.60	0.46	0.29	
S6						-	0	1	S6							-	1	1
							0.73	0.71									0.39	0.20
S7							-	1	S7								-	1
								0.65										0.21
S8								-	S8									-
*: Correlation not significant (two sigma : ~95%)																		
**: Multiple peak lag correlations																		

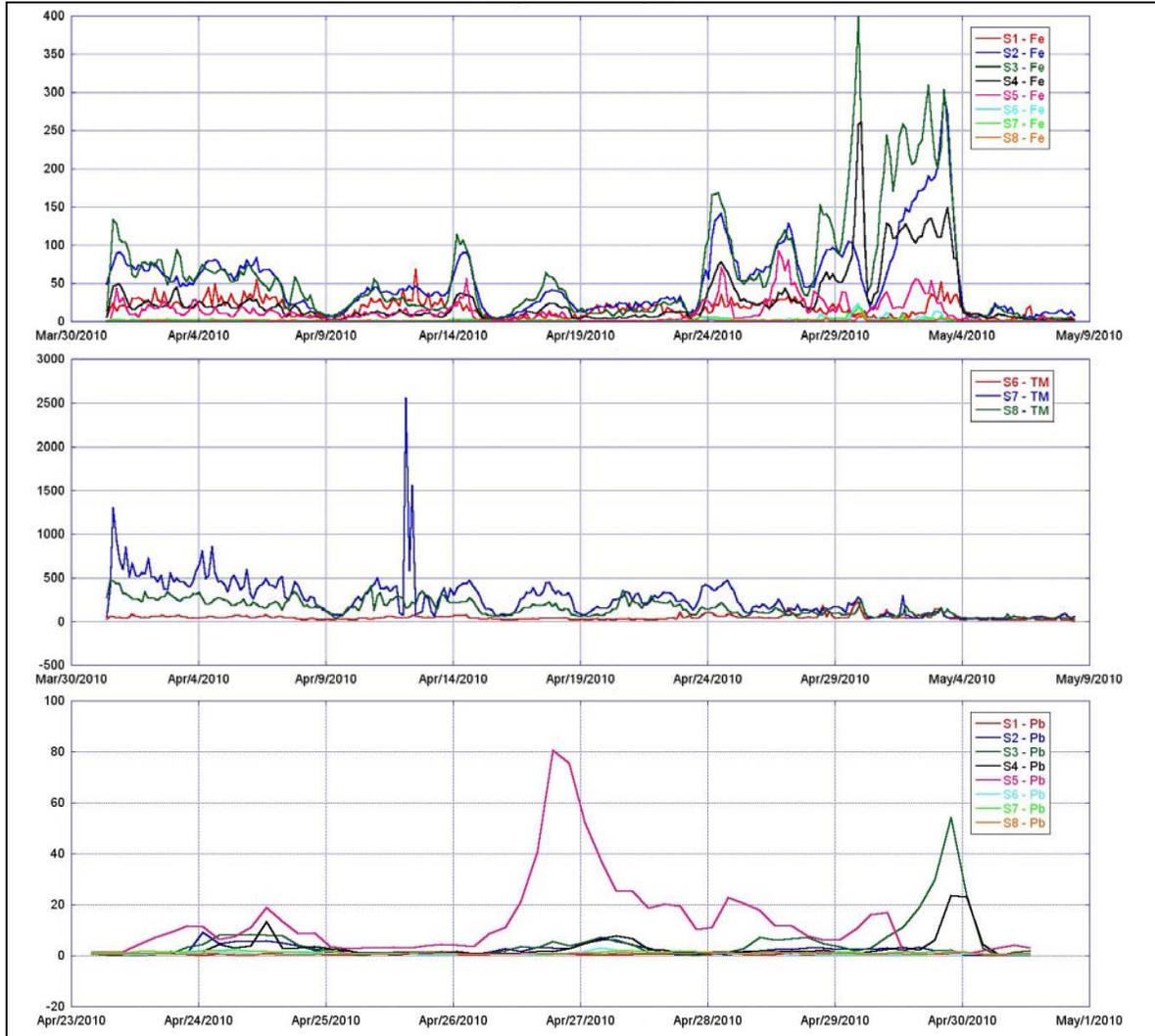


Figure 23 Dongsha Size Fraction Time Series. Selected plots for concentration ( $\text{ng}/\text{m}^3$ ) of iron, total mass, and lead show alignment between size fractions.

The DRUM data were then compared to the nephelometer data to verify the raw DRUM time stamps and correct them if necessary. Similar to the relative adjustment of individual size fractions, DRUM size fraction data were cross-correlated against three-hour averaged nephelometer data to determine peak correlations at various lagged time steps. Three-hour nephelometer data were then offset in  $\pm 30$  minute increments and likewise cross-correlated to further refine the correlation at various lagged time steps at a resolution finer than the three-hour DRUM time steps.

Figure 24 shows time series for DRUM summed elemental mass concentration and nephelometer scattering data. Correlations of summed and specific elemental concentrations to the scattering data generally yielded peak correlations at zero lag indicating that the best alignment between the two data sources was within  $\pm 1.5$  hours (half the three hour time step) of each other. The 30 minute offsets of the three-hour rolling averaged nephelometer data were used to further analyze this three hour window to determine if time stamp adjustments of less than one full time step were warranted. Again, no consistent lagged peak cross-correlation was found. In addition, a visual inspection of Figure 24 and the offset nephelometer time series yielded no indication that a time stamp correction was required. Thus, the raw time stamps for both nephelometer and DRUM data sets were retained for the remainder of the analysis.

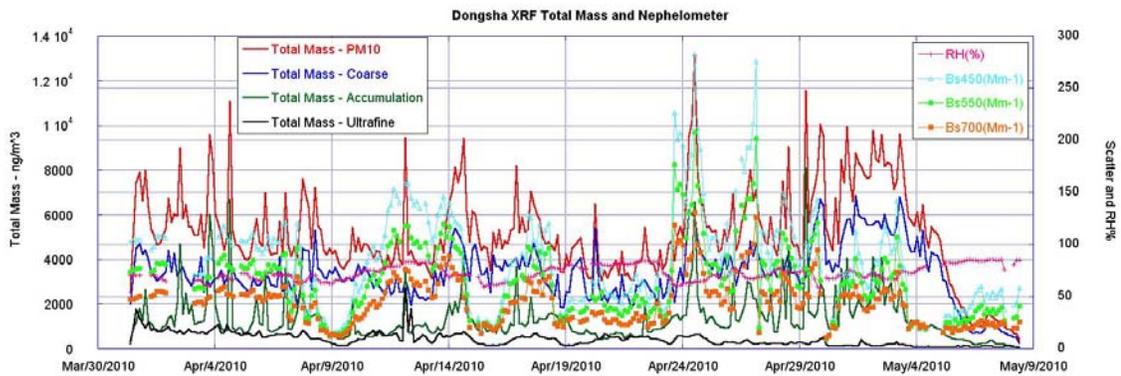


Figure 24 Dongsha DRUM summed mass concentration for various size fractions and nephelometer scattering.

### 3.2.2. SIZE FRACTION CORRECTION

DRUM measurements are nominally segregated into eight size fractions based on aerodynamic particle diameter. However, as discussed earlier, particle bounce-off, fracturing, and hygroscopic growth can cause sampled particles to be collected on an incorrect size fraction. As a result, the true degrees of freedom for particle size may be less than implied by raw DRUM data. In this work, elemental concentration data for the

eight raw size fractions were segregated into three combined size fractions based on typical aerosol size distributions [Seinfeld and Pandis, 2006]. Coarse, accumulation, and ultrafine modes were generated by combining stages 1, 2, and 3 ( $5\ \mu\text{m}$ ,  $2.5\ \mu\text{m}$ ,  $1.15\ \mu\text{m}$ ); stages 4, 5, and 6 ( $0.75\ \mu\text{m}$ ,  $0.56\ \mu\text{m}$ ,  $0.34\ \mu\text{m}$ ); and stages 7 and 8 ( $0.26\ \mu\text{m}$ ,  $0.07\ \mu\text{m}$ ); respectively.

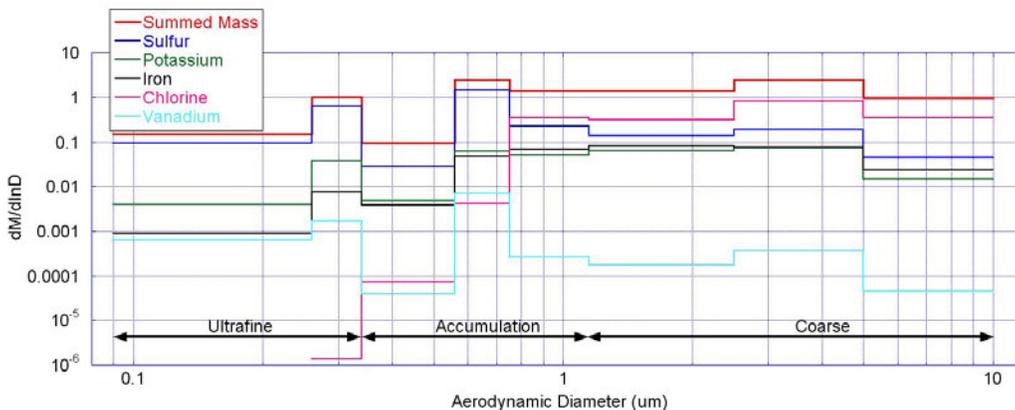


Figure 25 Dongsha mass size distributions ( $\text{ng} / \text{m}^3$ ) for summed mass and selected elements.

Summed mass and selected of elements are shown as mass distribution functions in Figure 25. Stages summed into the ultrafine, accumulation, and coarse modes are shown.

### 3.3. FACTOR ANALYSIS

Approximately 300 DRUM data points (time stamps) exist for the Dongsha experiment, each with 27 elemental concentration variables in each of the three size fractions, for a total of 81 variables. Factor analysis was performed to reduce the number of dimensions to describe the variability as described in the methods section. However, in order to better elucidate the nature of this variability, a number of different configurations of the data were subjected to FA.

The first configuration simply utilized all three lumped size fractions for all of the elements available. A second configuration, which is discussed below, used only the coarse mode variables. Other configurations involved combining the accumulation and ultrafine modes into one “fine” mode, limiting the number of elements subjected to analysis, and including other data sets such as the nephelometer in the factor analysis. Furthermore, different numbers of factors were extracted for each configuration that was tried. The ultimate decision for which configurations and number of factors extracted to use was made based on the physical interpretation of the factor composition and variability as described in the Methods section.

Ultimately, it was determined that the first two configurations, namely the use of all three size fractions (Three-Mode FA), and only the coarse size fraction (Coarse-Mode FA), captured the majority of the variability able to be explained by this FA method. Attempts to extract more factors in these configurations did not yield factors which had strong physical interpretations, while extracting fewer factors would have left variability unexplained that was potentially associated with an actual source. In addition, the scree plots and method described in Section 2.2. supported the choice of factors to extract. Other configurations primarily identified the same general factors described by these first two methods with only small variations, and thus are not discussed here or utilized further.

In analyzing the results of factor analysis, it is also worthy of note that of the species measured by the DRUM instrument, sulfur is the only element not strictly associated with primary emissions. While other elements are primarily transported from source to receptor with relatively little change in their chemical composition, sulfur is often emitted

as a gaseous reactive species subject to oxidation in both the gas and aqueous phases. As a result, the source/receptor relationship for this species may not be as clearly defined by a transport pathway as for other species.

### **3.3.1. THREE-MODE FACTOR ANALYSIS**

In the Three-Mode FA, following the techniques described in section 2.2., six factors explaining approximately 61% of the variance in the data were extracted (See Table B-2). Appendix B shows additional figures with information regarding the FA and factor extraction process.

Table B-4 shows the rotated component matrix produced by the PCA method, with key elements within each factor shown. Elements (at each size fraction) with scores greater than 0.3 were considered to have enough variance explained to be included as a characteristic component of that factor. The scores for each size-resolved elemental component indicate the correlation of the factor to that component. Each factor is then understood as a pattern of covariability wherein each component's importance to that factor is indicated by its component score. It is important to note that the absolute magnitude of each variable (i.e. each size-resolved elemental concentration) is not important to determining its importance to a factor. Rather, it is the degree to which an element covaries with a factor that determines its score. Therefore, an element with only very small absolute concentrations may still play an important part in a factor.

The absolute mass of each key element within a factor is considered once a factor is established, as it provides an indication of how much of the summed aerosol mass associated with each factor is associated with key elements. This can help establish which elemental components comprise the majority of aerosol mass within a factor and can help

establish the source type. For instance, in factor 1, a number of coarse and accumulation mode elements have high component scores, indicating that the majority of their respective variances are explained by the variability of factor 1. Of these key components with high scores, coarse mode aluminum, silicon, calcium, and iron, elements typically associated with dust, make up the majority of the mass, while other components such as coarse lead make up only a small fraction of the aerosol mass associated with this factor (see Figure B-1). As a result, dust is assigned as the primary source type associated with this factor based on the fingerprint created by the elemental component scores and the aerosol composition at time stamps dominated by the factor. Had coarse mode lead been a major mass fraction associated with this factor another source type might have been assigned. Rather, we expect that the lead is more likely a trace pollutant in the dust source region, or perhaps was mixed or associated with the dust when it was advected through a polluted or industrialized region. Table 3 shows the assigned source types based on the component scores and elemental mass fractions for each factor.

Table 3 Dongsha Three-Mode Factor Source Types and Associated Elemental Composition.

Factor / Assigned Particle Type	Description	Major Mass Fraction Components
1. Dust	Typical elements associated with dust, primarily in the coarse mode with some accumulation mode species. Small amounts of coarse mode lead are also seen.	Coarse Mode: Al, Si, Ca, Fe, K Accumulation Mode: Al, Si
2. Pearl River Delta Pollution	Accumulation mode heavy metals with coarse and accumulation mode zinc. Assigned to Pearl River region based on backtrajectories.	Coarse and accumulation mode zinc dominates, with peaks in accumulation mode lead.
3. Industrial Ultrafine Sulfur	Factor closely tracks the ultrafine sulfur signal, with some ultrafine magnesium contributing.	Ultrafine sulfur, some ultrafine magnesium.
4. Coastal China Pollution	Primarily ultrafine heavy metals associated with industry emissions, along with some coarse zinc. Assigned to coastal China region based on backtrajectories.	Ultrafine zinc, iron, silicon, and lead dominate the mass, with some coarse mode zinc also associated with this factor.
5. Sea Salt	This factor tracks the coarse and some accumulation mode chlorine and magnesium associated with sea salt closely. Several peaks also seem to covary with factor 1, the dust signal.	Coarse mode chlorine and magnesium comprise the majority of the mass associated with this factor, with some coarse sulfur.
6. Industrial Accumulation Mode Sulfur or Fly Ash	This factor very closely tracks the accumulation mode sulfur signal. This could be associated with oxidation or cloud processing of sulfur emissions (secondary sulfate production) from anthropogenic combustion sources.	Accumulation mode sulfur, with lesser amounts of magnesium, aluminum, and potassium.

Figure 26a shows the factor score time series for the Dongsha experiment. A smoothed version of the same plot, generated by using a 24-hour boxcar average of each factor, is shown in Figure 26b to assist with interpretation of factor variability. The source attributed to each factor is shown in these figures.

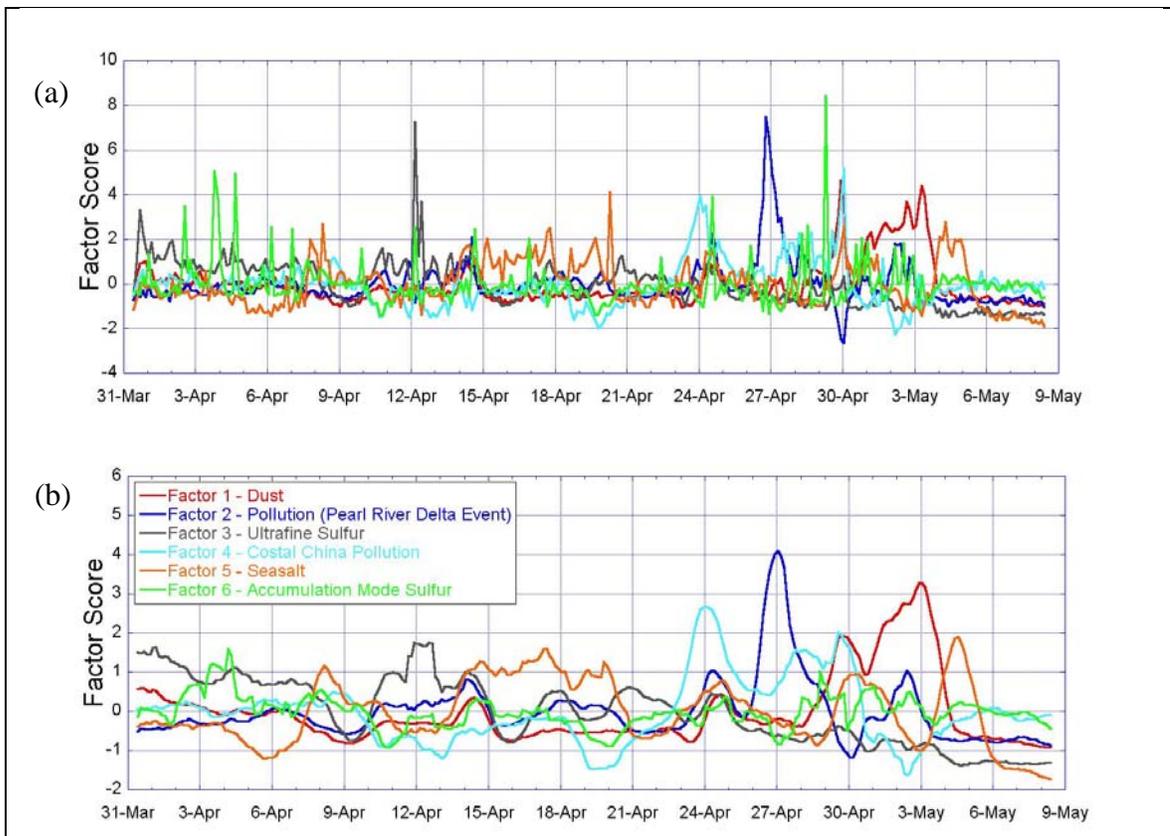


Figure 26 Timeline of Dongsha Three-Mode Factor Analysis scores for (a) the raw factor scores, and (b) factors smoothed by 24-hour boxcar average.

### 3.3.2. COARSE-MODE FACTOR ANALYSIS

In the Coarse-Mode FA, three primary factors were identified explaining approximately 64% of the variability in the coarse mode, whereas only two (mainly) coarse mode factors were identified in the initial factor analysis. The first factor in this coarse analysis, coarse factor 1, is a dust factor, similar to the first dust factor originally identified (factor 1), but with some differences before the large dust plume starting on 29 April. In fact, coarse factor 2 seems to account for much of the variability in factor 1 before this 29 April event, such as the 14 April event, and the dust between 24 and 27 April. In the three-mode analysis, factor 5 was identified as a sea salt signal, and here appears to closely track coarse factor 3, which was also identified as the sea salt signal.

The coarse-mode analysis can therefore be interpreted as a refinement of the coarse mode signals seen in the three-mode analysis. Dust which appeared primarily in factor 1 is now separated into coarse factors 1 and 2, with coarse factor 1 dominating the large plume towards the end of the study period, and factor 2 dominating the dust variability occurring earlier. Finally, the coarse factor 3 sea salt signal is more clearly independent of the dust sources than it was in the Three-Mode FA.

The results of the coarse-mode analysis indicate that dust and sea salt dominate the Dongsha coarse mode aerosol, a feature consistent with expectations. The description of each of these coarse mode factors is given in Table 4, while Figure 27 shows the time series of factor scores for the coarse-mode factor analysis. The average mass fraction of selected elements are shown in Figure 28 for each of the coarse factors and the whole study period. The results support the factor analysis, with chlorine and magnesium having the highest fractions at time stamps when coarse factor 3 dominates. Crustal elements likewise comprise the largest fraction of the dust factors (coarse factor 1 and 2), while sulfur associated with pollution comprises a greater fraction of the polluted dust (coarse factor 2).

Table 4 Dongsha Coarse-Mode Factor Source Types and Associated Elemental Composition.

Coarse Factor / Assigned Particle Type	Description	Major Mass Fraction Elements
1. Dust Type I	Typical dust elements. Factor is associated with the large dust plume starting after 28 April.	Al, Si, Ca, Fe, K comprise the majority of the mass associated with this factor.
2. Dust Type II	Typical dust elements, but with additional components of coarse sulfur and magnesium not associated with coarse factor 1. These additional elements may secondary components associated with pollution laden or cloud processed dust.	Similar crustal components noted in coarse factor 1 are major fractions of mass associated with this factor. Additionally, coarse sulfur and magnesium constitutes a greater fraction of the mass here than in coarse factor 1.
3. Sea Salt	Sea salt associated with chlorine and magnesium.	Ultrafine sulfur, some ultrafine magnesium.

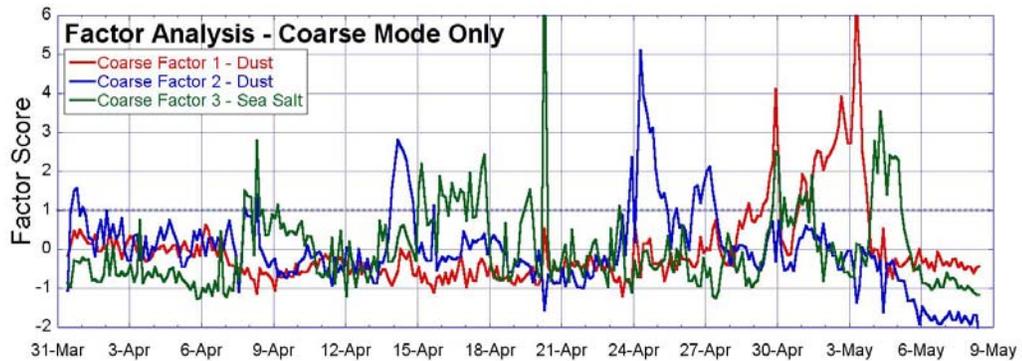


Figure 27 Timeline of Dongsha Coarse-Mode Factor Analysis scores.

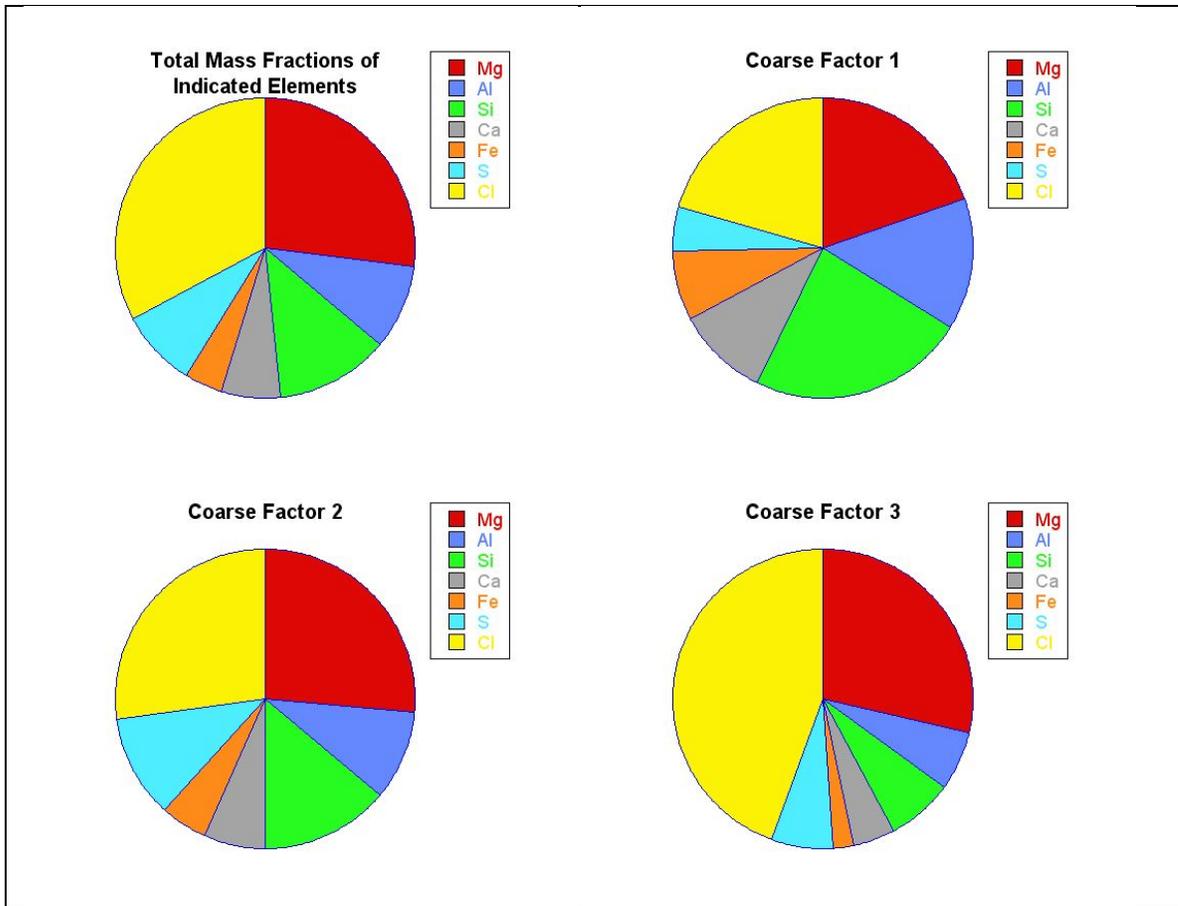


Figure 28 Dongsha coarse-mode mass fractions of indicated elements for all time stamps, and for time stamps when each factor score was greater than 1.

### 3.4. BACKTRAJECTORY ANALYSIS

Backtrajectory analysis was performed for the time period of the Dongsha experiment. Five day backtrajectories were run at receptor heights of 100 m, 500 m, 1000 m, and 1500 m to show the modeled estimates of transport both within, and just above, the boundary layer. Additional filtering was performed in some cases to analyze only at trajectory endpoints within 2000 m of the surface.

#### 3.4.1. DONGSHA NORMAL CASE

The backtrajectory Normal Case consists of all backtrajectories generated during the DRUM sampling period of the Dongsha experiment. A total of 320 five-day

backtrajectories were generated at each receptor height. Figure 29 shows the individual backtrajectories, colored by arrival time at the receptor, and the SCF contour plots for the experiment.

A clear height dependence exists in trajectory source region. The 100m and 500m trajectories indicate that likely source regions are more often to the north and east of Dongsha, while by 1500m, most trajectories originate south and west of the island. This shift highlights the effect of the low-level wind shear and temperature inversions noted earlier. Above the boundary layer, the 1500m trajectories have longer residence times over Southeast Asia in regions of significant springtime biomass burning. Smoke may be transported over Dongsha from these areas, but any temperature inversions will likely minimize transport through the boundary layer so that smoke may not be observed at the surface. The backtrajectory heights plotted in Figure 30 also show that backtrajectories arriving within the boundary layer tend to exhibit flow patterns which remain in the boundary layer over the open ocean. This observation is consistent with the notion of inversion layers developing over this area and maintaining a separation between the marine boundary layer and air aloft.

Furthermore, as transport to Dongsha within the boundary layer is expected to dominate surface observations, the low level backtrajectories are anticipated to explain more of the variance in measured aerosol composition than trajectories from higher levels. As such, based on Figures 29 and 30, source regions for this study are expected to include Coastal China along with the Gobi and Taklimakan deserts, Korea, Southern Japan, and marine areas.

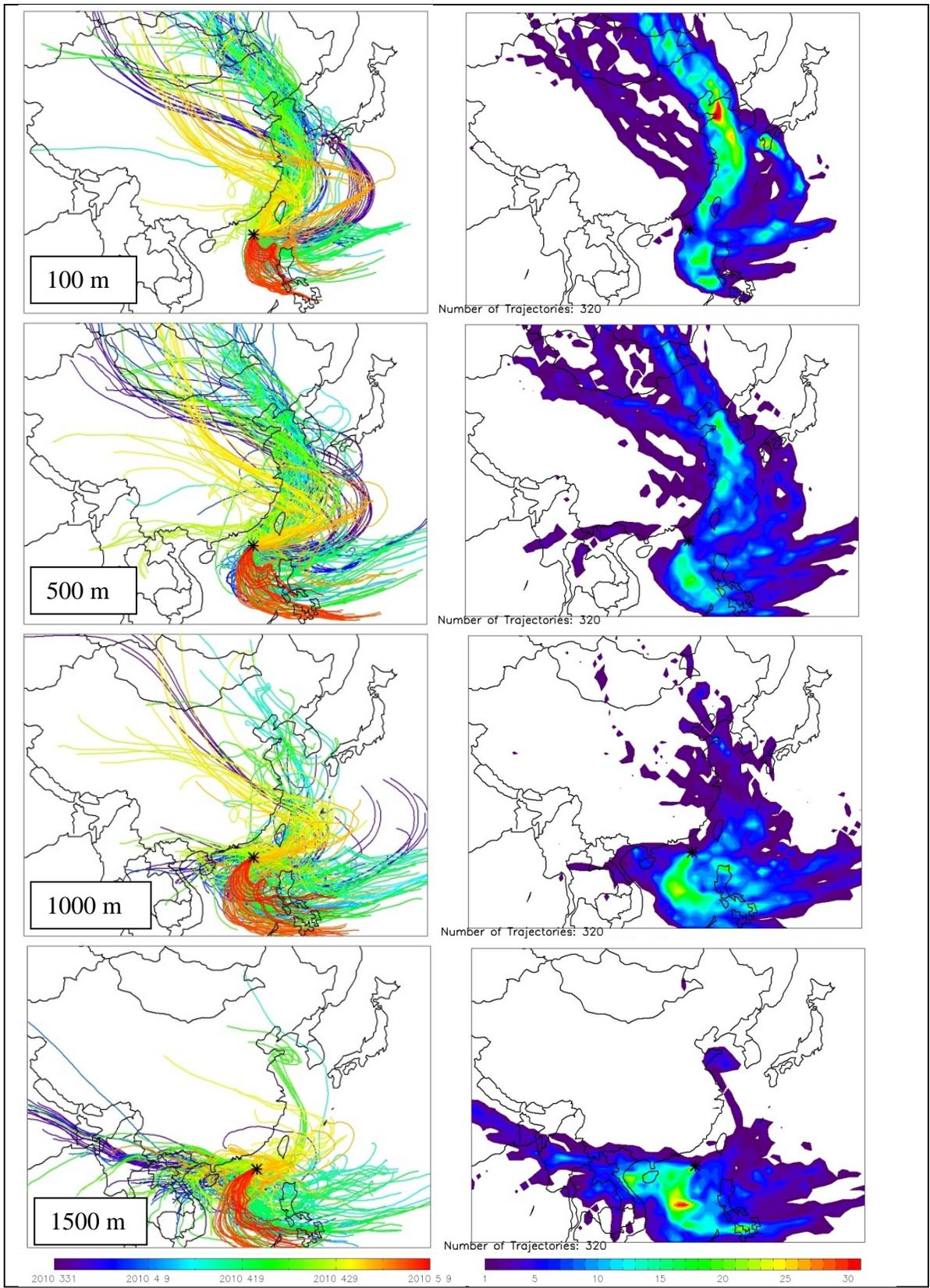


Figure 29 Dongsha Normal Case backtrajectories and SCF at 100m, 500m, 1000m, and 1500m

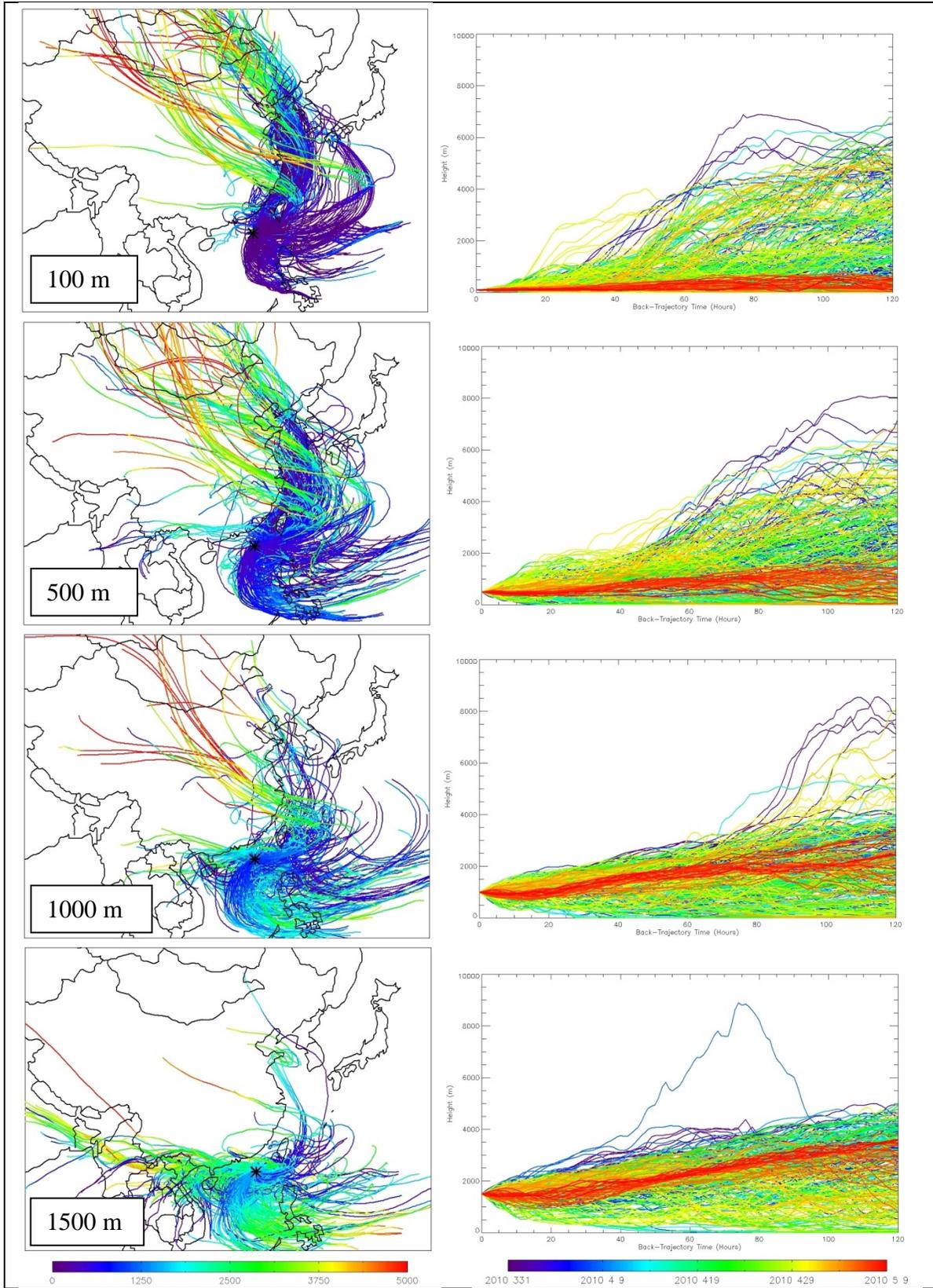


Figure 30 Dongsha Normal Case backtrajectory heights at 100m, 500m, 1000m, and 1500m

### 3.4.2. DONGSHA THREE-MODE FACTOR ANALYSIS SPECIAL CASES

Special cases for each of the three-mode and coarse-mode factors that were identified in the previous section are shown below based on a criterion of accepting time stamps with factor scores greater than 1 into the special case (i.e. a cutoff criterion of 1 standard deviation from the mean) [Ashbaugh *et al.*, 1985; Polissar *et al.*, 2001b]. This methodology effectively considers the special case to occur when a factor reaches greater than one standard deviation from the mean as opposed to a specific concentration. Backtrajectories colored by arrival time are displayed in Figure 31 and tend to indicate, as would be expected if this backtrajectory analysis has explanatory power, that distinct time periods (similarly colored groups of backtrajectories) associated with a factor tend to have similar transport patterns. Throughout this analysis, it is important to remember that the meteorology yields only approximates of source regions, and leading and trailing effects are possible; nevertheless, certain features do appear to differentiate between factors.

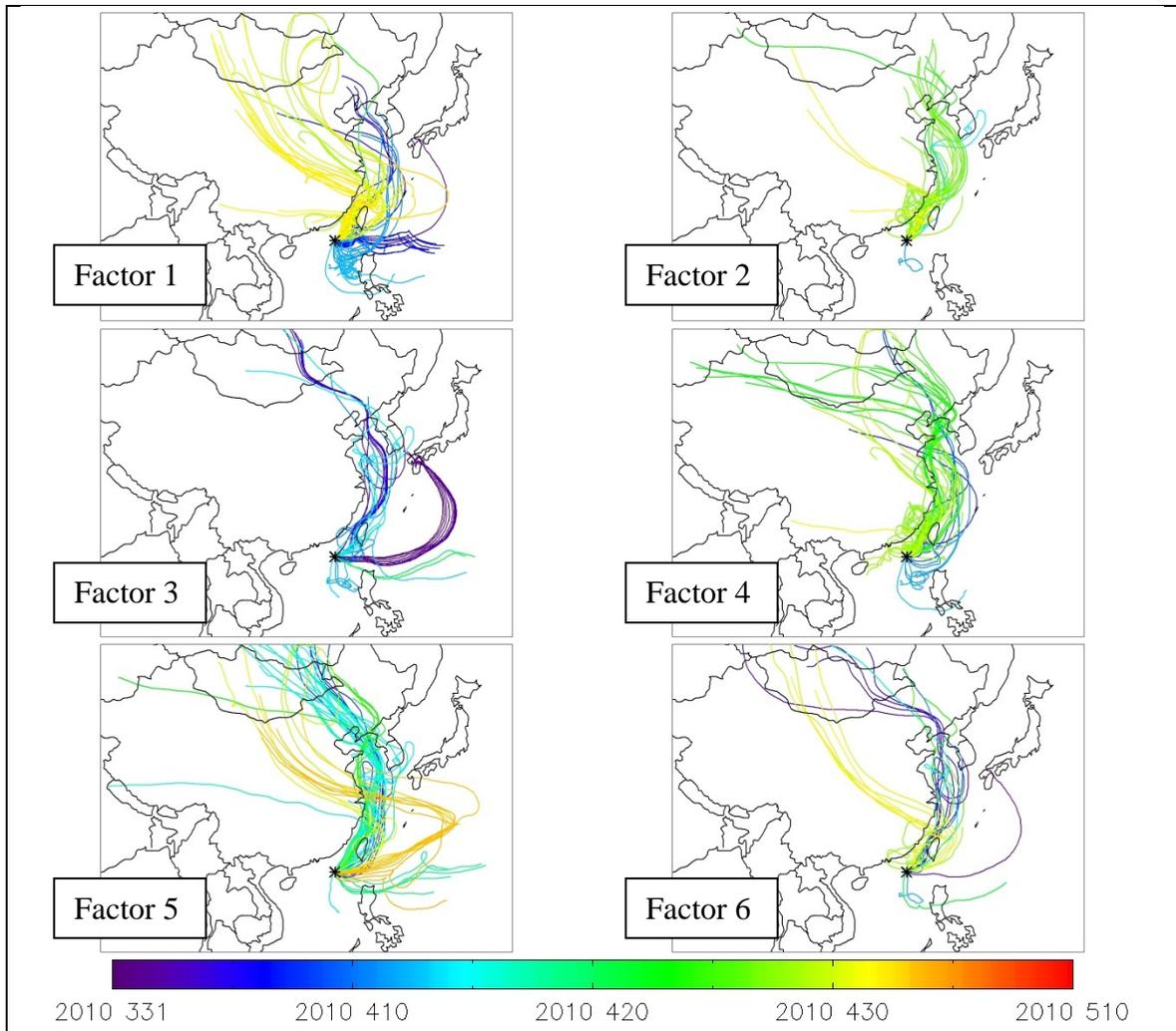


Figure 31 Dongsha Three Mode Special Case backtrajectory plots for 100m receptor heights.

Special Case RTA plots, shown in Figure 32, provide an indication of the residence time of each factor compared to the normal case residence time. Factor 1, initially characterized as having a significant dust component, indeed has larger residence times further inland over the Taklimakan and Gobi deserts, both known to be significant dust aerosol sources [Cohen *et al.*, 2010c; S W Kim *et al.*, 2010; X Wang *et al.*, 2008]. Factor 5 is also seen to have significant residence time over these desert areas, which, while not specifically identified as a dust factor in the FA, did seem to covary to some degree with elements of coarse factor 2 (one of the dust coarse factors). Factors 2 and 4 have some residence time over the edge of the Gobi desert, though perhaps not as much as factors 1

and 5, while both also have slightly more residence time over heavily populated coastal areas. Overall, the residence time distributions of all factors appear only slightly different, although the residence time distributions for dust factors do generally seem more consistent with desert regions.

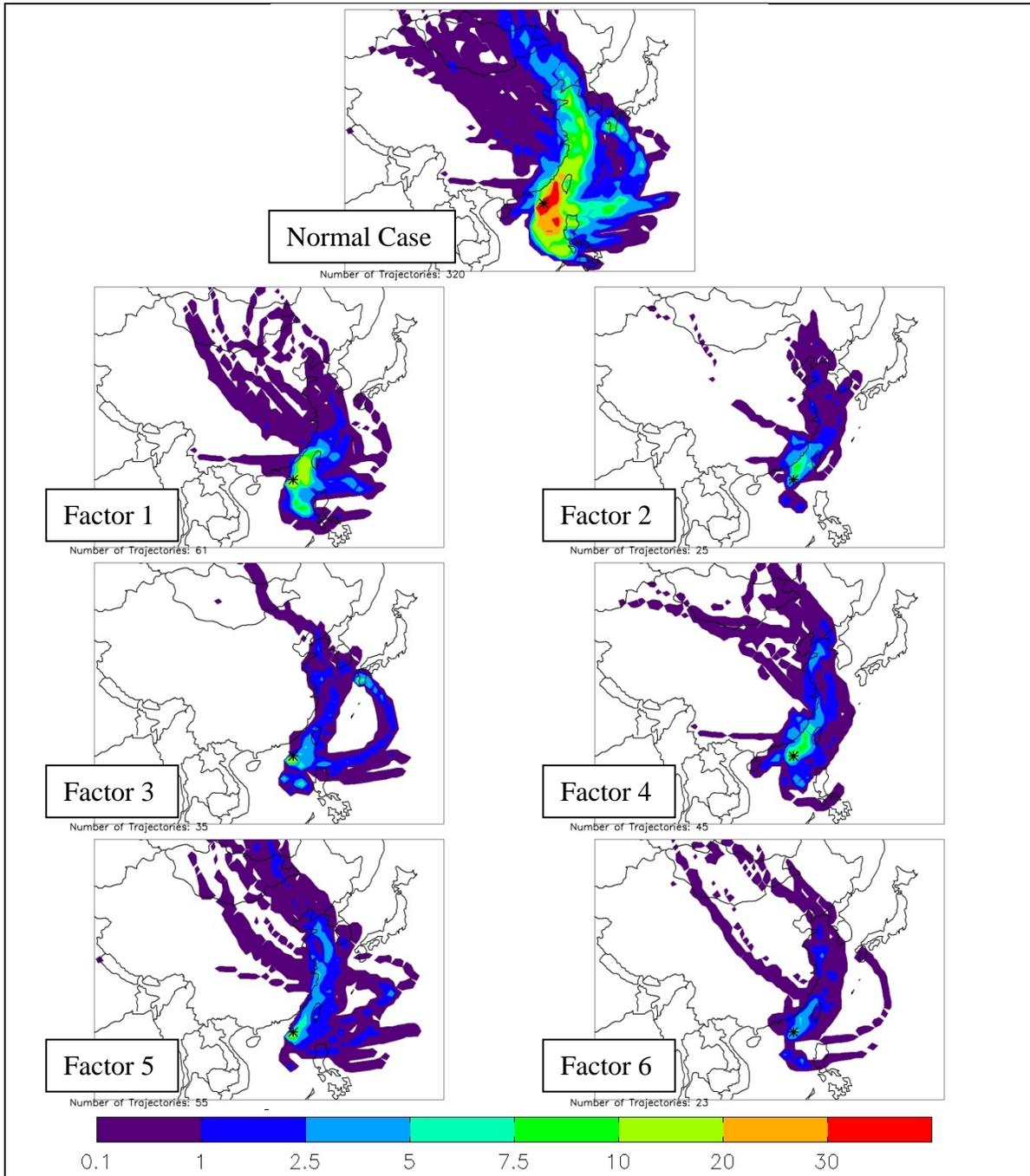


Figure 32 Dongsha Three Mode Special Case RTA plots for 100m receptor heights.

In the PSCF analysis (Figure 33), factor 1 strongly identifies the inland Taklimakan desert region, with evidence of some possible leading effects to the northwest. Nearly all of the trajectories in the normal case that extend to some parts of this region are associated with factor 1. Regions identified with factor 5 in the PSCF are considerably more spread out among the domain, compared with the other factors. As factor 5 is a sea salt factor, it is possible that the lack of a specific area identified by PSCF is indicative of non-synoptic-transport effects, such as wind speed, that are more important to sea salt variability. This possibility will be explored further in the coarse mode analysis.

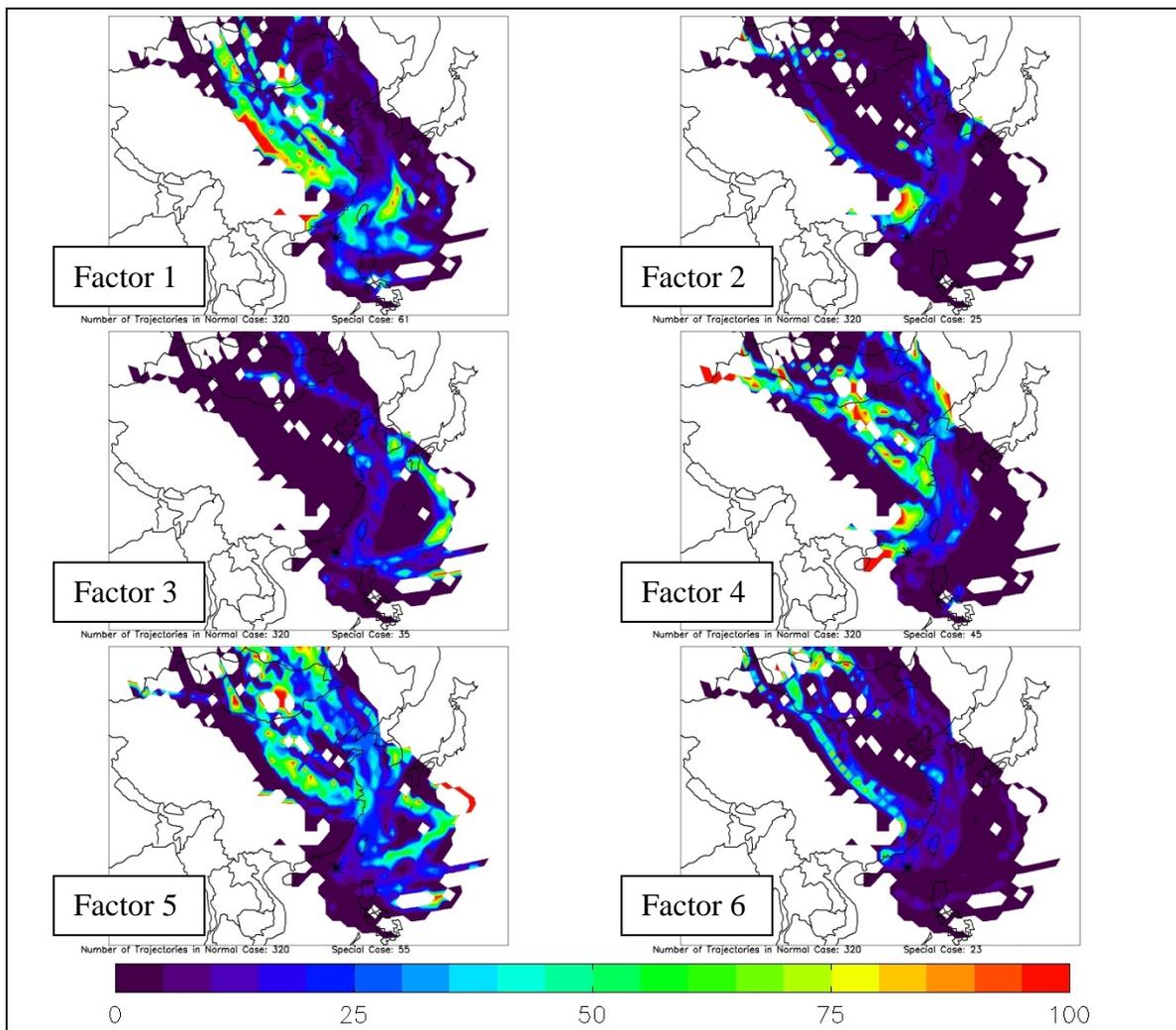


Figure 33 Dongsha Three Mode Special Case PSCF plots for 100m Receptor Heights.

The PSCF for Factor 4 does select several regions with potential desert sources that were also identified by factors 1 and 5, yet it is not considered a dust factor based on the chemical analysis. The unexpected areas identified by factors 4 and 5 typify the caution which must be exercised when using backtrajectory analysis; while a region may be highlighted in the PSCF or other functions, it does not necessarily mean that a source is definitively there. Even if the model is correct and air parcels did in fact originate over these desert regions during periods when the factor 4 scores are high, we note there are other reasons why dust may not be present in the DRUM samples. A trajectory may have had aerosol removed due to rainout; soil moisture or wind speeds over the area may have prevented significant generation of dust; or it is possible that dust was in fact present, but was not associated with the factor. For instance, we note that there were several time periods when both factor 4 and factor 1 were significant according to the criterion used. As we pointed out earlier, multiple sources, and therefore multiple factors may be impacting the receptor at the same time, which necessarily highlights the same regions since the backtrajectories are the same.

Note that factors 1, 4, and 5 all have peaks on 30 April in Figure 26a. Similarly, Figure 31 shows that many of the backtrajectories responsible for identifying the desert regions in the PSCF plot are from the 24 April peak in factor 4, during which there was also a lesser peak in factors 1 and 5. This could be interpreted as a smaller dust event that passed over a pollution laden air mass and mixed the two sources together with sea salt aerosol. During this period, despite the sources having separate locations, the same areas would be identified for all three factors.

Since the Dongsha study had limited data (only six weeks of DRUM samples were available), only a few strong peaks are identified in each factor timeline. Further, as was shown climatologically, the synoptic transport in the region during the study tends to be strongly dominated by only a few unique locations. If more data were collected over a broader range of meteorological conditions over a longer time period, the power of the backtrajectory analysis would be expected to increase. Given the limited data, the PSCFs calculated here are used primarily to determine if the indicated source region estimates clearly agree or disagree with potential source areas suggested by the FA results.

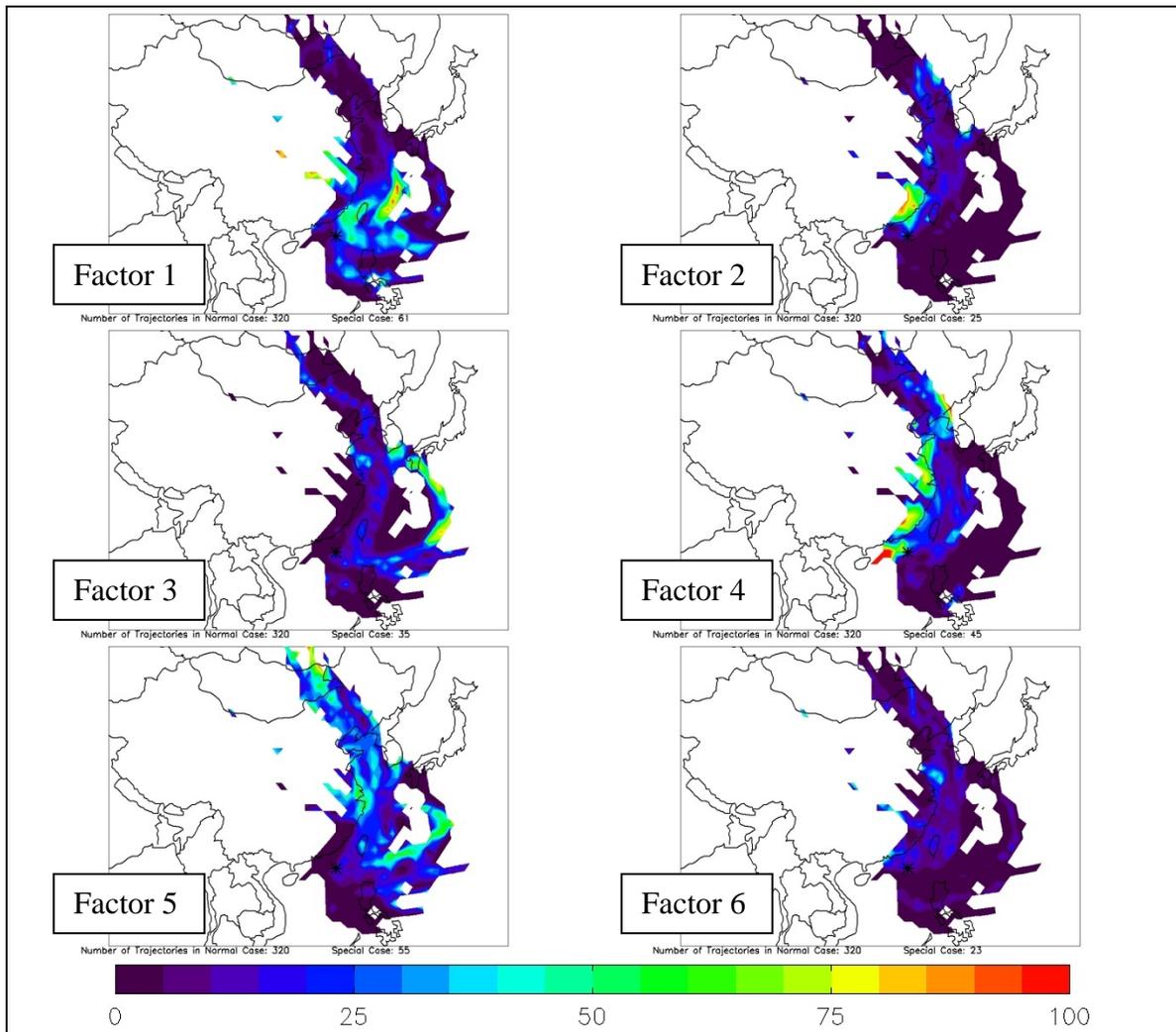


Figure 34 Dongsha Three Mode Special Case PSCF plots for 100m Receptor Heights with 10 endpoint cutoff.

Continuing the analysis of the applicability of PSCF source region estimates, Figure 34 shows the same PSCF, but with a normal case 10 endpoint per grid box minimum criterion. We can see that some of the features and source areas identified as having to do with factors are no longer included. However, the most significant regions (ones with larger normal case residence times) are still identified. Factor 2 is seen as having a large PSCF at, and to the immediate north of the heavily industrialized Pearl River Delta region. As this factor was identified as a pollution event, it is primarily associated with just one time period (26 to 27 April), and the backtrajectories point to one area, this factor was determined to likely result from one pollution plume originating over the nearby Pearl River Delta and the region to its north. In addition, during this period, a sounding from that time for the region identified by the factor 2 PSCF shows a low-level inversion and generally westerly winds (Figure 35; compiled by the University of Wyoming, available at <http://weather.uwyo.edu/upperair/sounding.html>), both of which are consistent with a plume from this region.

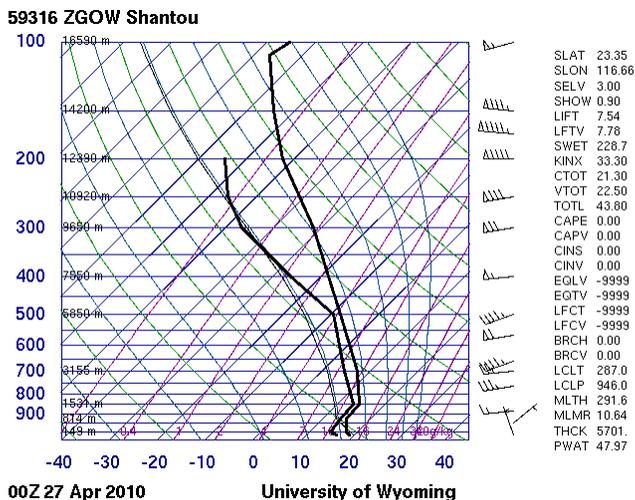


Figure 35 Atmospheric sounding from the China coast just to the Northwest of the Pearl River Delta showing a low-level temperature inversion.

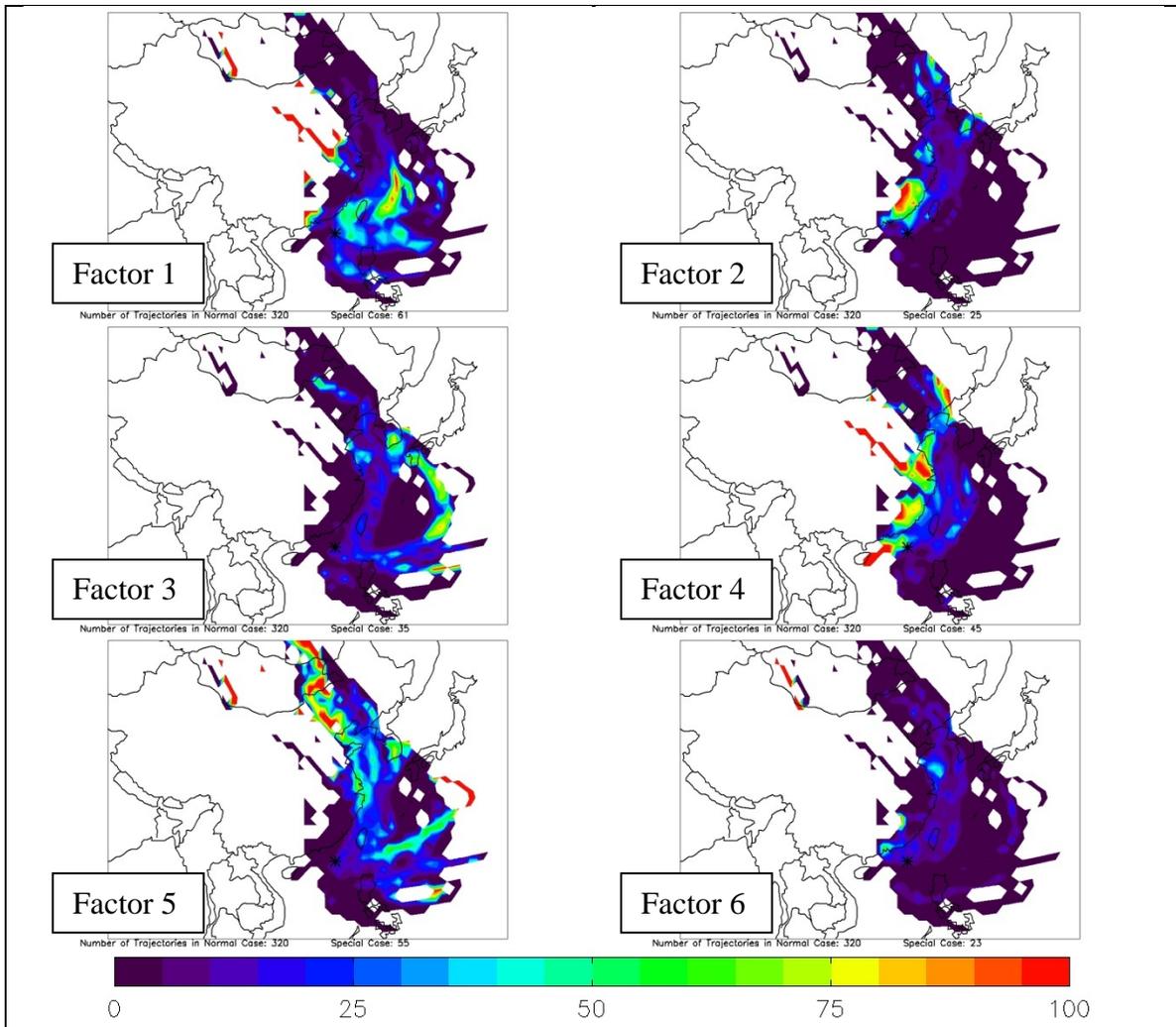


Figure 36 Dongsha Three Mode Special Case PSCF plots for 100m receptor heights, filtered for endpoints within 2000m above ground height.

Figure 36 shows the same PSCF analysis as Figures 33 and 34, but only for trajectory endpoints which are within 2000 m of the ground. The differences between factors 4 and 5 are immediately more apparent than in the earlier figures; factor 5 now identifies the Gobi desert and marine areas more clearly, while factor 4 selects coastal regions. Factor 1 still identifies more western desert regions, although data are far more sparse when looking only at low level endpoints. Since dust is generally lofted to higher altitudes when generated during storms, and since it is likely that one or two specific lofting events

dominate the source, the sparse, higher-altitude endpoints for factor 1 as shown in Figure 33 may be appropriate indicators of the source regions for factor 1.

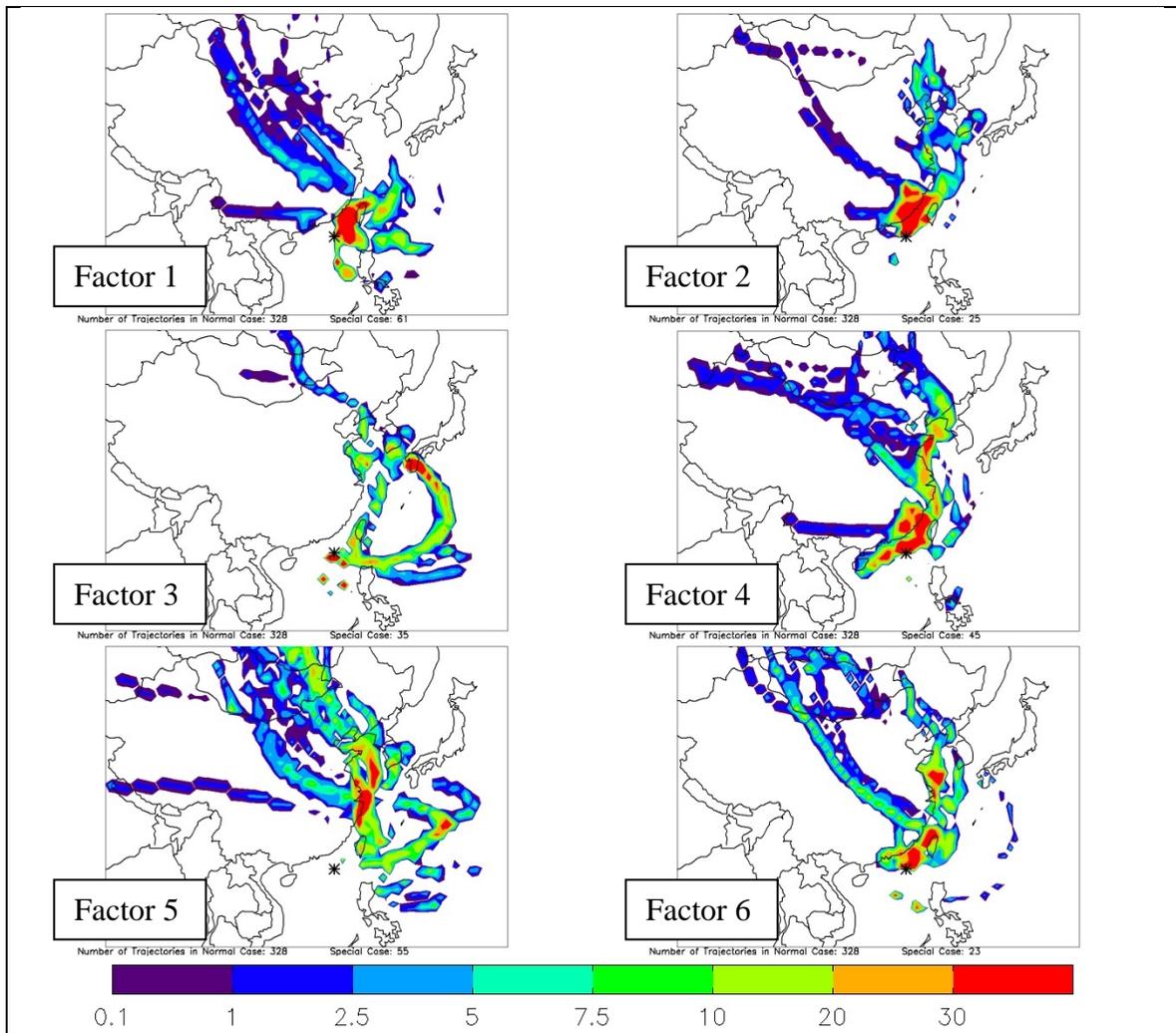


Figure 37 Dongsha Three Mode Special Case RTDF plots for 100m receptor heights.

The RTDF in Figure 37 can be considered as an additional measure of source areas with the greatest impact for each factor. The potential utility of the RTDF plot comes in trying to differentiate between regions with similar PSCF values for each factor. The factor 4 PSCF identified coastal regions along with interior desert regions. The factor's RTDF however, shows coastal regions to have higher absolute differences in special case endpoints than the interior desert regions. This implies that these coastal areas had higher

overall residence times during times when the factor 4 pattern was seen in the measured DRUM data. Despite the same PSCF fraction of special case to normal case trajectory passes over desert and coastal areas, the coastal areas are expected to be of greater importance in the factor 4 source regions.

A similar indication is shown for factor 5, with marine areas showing more importance than coastal or desert regions, with the notable exception of the western Mongolian Gobi region. While this may be an example of a trailing effect, it is also possibly an indication of the limitations of backtrajectory analysis in this circumstance. Further, the factor 1 RTDF shows a situation opposite to what is expected, wherein close marine regions are heavily emphasized over the expected interior desert region. Again, this may simply be a leading effect, or it may be indicative of method limitations.

A final discussion on factors 3 and 6, not discussed up to this point, must first consider their expected source. Factor 3 consists primarily of the ultrafine signal, which is dominated by sulfur in the DRUM data. As Dongsha is a fairly remote site, with most air masses arriving from distant sources and thus representing aged (larger-diameter) aerosol, ultrafine sulfur is expected to be largely local, from either the diesel generator or nearby shipping associated with the major shipping lanes in the SCS. Local wind and operating conditions would be expected to contribute to the variability of this signal, rather than long range transport. Backtrajectory analysis would be expected to yield no new information, and therefore the special figures should approximate the normal case. In reality though, with a small data set such as this one, the local wind conditions will likely have at least some correlation to the larger synoptic state of the atmosphere, yielding a special case subset that is not truly random. It would therefore be expected that the

backtrajectory analysis function would show some improperly identified source regions. That is indeed what is seen in the factor 3 figures, though they are generally lower than other factors, indicating that the synoptic state of the atmosphere is less important for factor 3.

Factor 6 is a case where little consistent understanding of source regions can be gleaned from the backtrajectory analysis figures. The RTA figures do not show figure 6 to be significantly different from other factors, however, the remainder of the figures comparing the special and normal cases yield little indication of consistent source regions. Some coastal regions are perhaps highlighted, but generally the factor seems to be not notably different from the normal case. Since factor six was the last of the PCA factors extracted from the DRUM data it also contains the least amount of variance explained. A possible explanation for its limited impact is that factor 6 is merely a residual component of a more important source. Alternatively, as the factor is characterized by large, narrow peaks of accumulation mode sulfur, it is possible that the HYSPLIT model cannot adequately resolve the variability in the synoptic state of the atmosphere at a fast enough time resolution. A final possibility is that the peaks are associated with secondary aerosol production, which does not have a defined source region.

Further discussion and results of the special case backtrajectory analysis at a variety of receptor heights are given in Appendix D. As the receptor height for the backtrajectory analysis is increased, source areas tend to move to the south and west with the change in normal case residence times. Little additional utility was gained by the additional heights in terms of interpreting the factors for their respective source regions, however, as low-

level transport was considered the dominant synoptic-scale transport mechanism into the SCS, this is not unexpected.

### **3.4.3. DONGSHA COARSE-MODE FACTOR SPECIAL CASES**

The special case backtrajectories are similarly plotted for each of the three coarse-mode factors in Figures 38 and 39. Based on the FA results, coarse factors 1 and 2 are both dust sources, with the second factor identified as potentially more pollution laden dust. Figure 38a clearly shows that the two dust factors identify different regions and different synoptic atmospheric states. Coarse factor 1 selects much the same region as factor 1 from the three-mode analysis, with the western Gobi and Taklimakan deserts being identified. Plots for Factor 2 however, show source regions more to the east and north, and at lower altitudes, compared to factor 1. This could be interpreted as dust from the eastern Gobi being transported through more heavily polluted coastal areas into the SCS, consistent with the FA results.

The sea salt signal, coarse factor 3, does not clearly identify any specific region, and instead appears to select many of the potential source areas from the normal case. An additional comment concerns the persistent presence of sea salt in each of the coarse mode factors. The FA discussion identified coarse factor 3 as the sea salt factor, however, in reality, it was merely an increase in measured sea salt, as all of the data points contained a large amount of coarse mode sea salt as determined by coarse mode elemental Cl and Mg. As Dongsha is a low-lying remote island, consistently large sea salt concentrations are expected. However, Figure 38 (b) and (c) show an interesting result that can potentially be explained by increased generation of sea salt aerosol at high wind speeds. Low altitudes and lower trajectory ages further out over marine areas in factor 3,

when compared to the first two coarse factors, indicate that the elevated sea salt factor occurred during times when the model predicts low altitude, high speed trajectories over the open ocean.

Overall, the refined coarse-mode backtrajectory analysis is largely consistent with the FA in its indication of potential sources for each factor. As a result, while none of the backtrajectory analyses would on their own be convincing of evidence for the source/receptor relationship of the special case factors, when interpreted alongside the FA results, they demonstrate that general agreement between the two methods can be achieved.

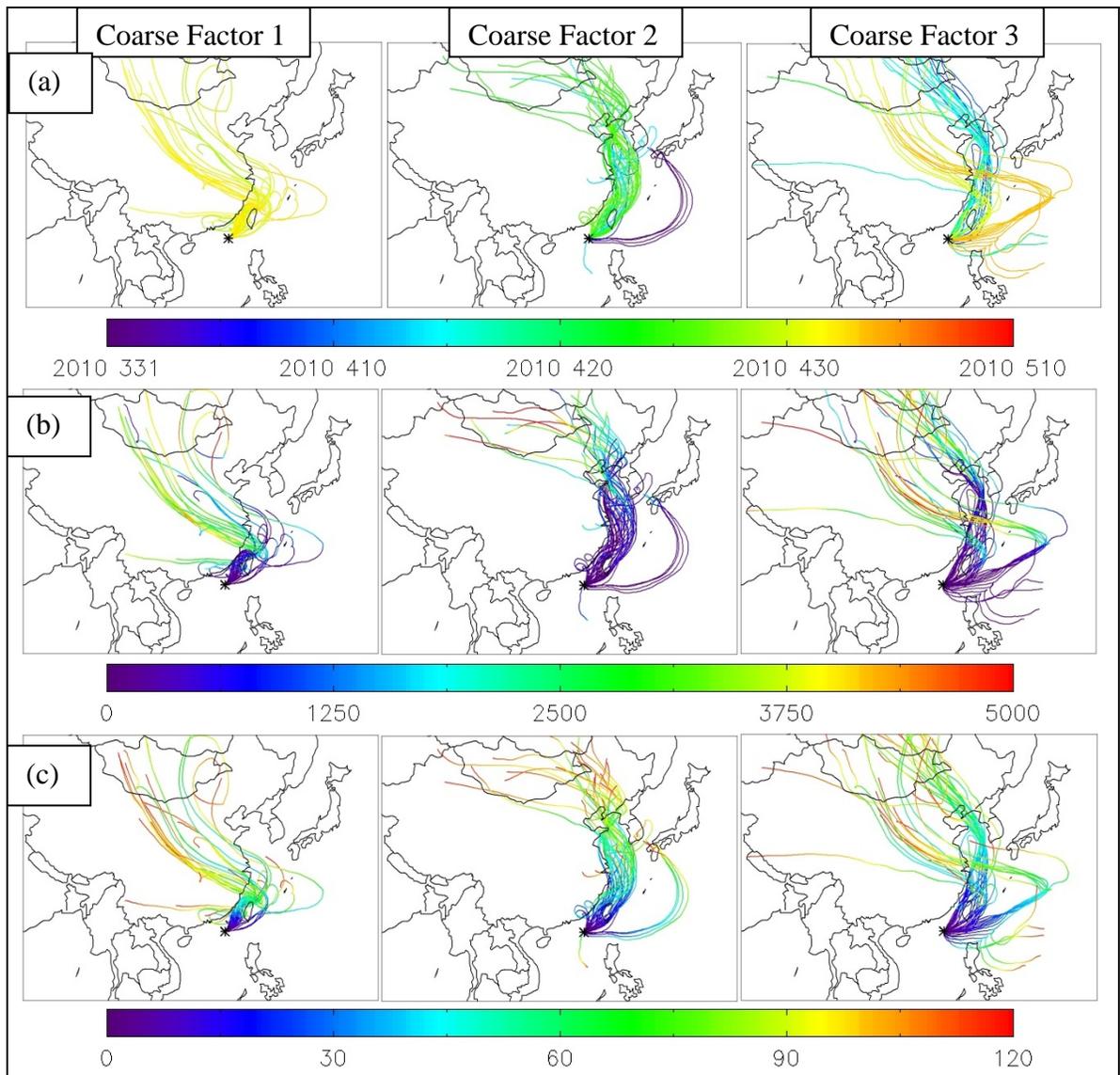


Figure 38 Dongsha Coarse-Mode Special Case backtrajectory plots for 100m Receptor Heights. Shown are trajectory (a) arrival times, (b) altitudes (m), and (c) ages (hours).

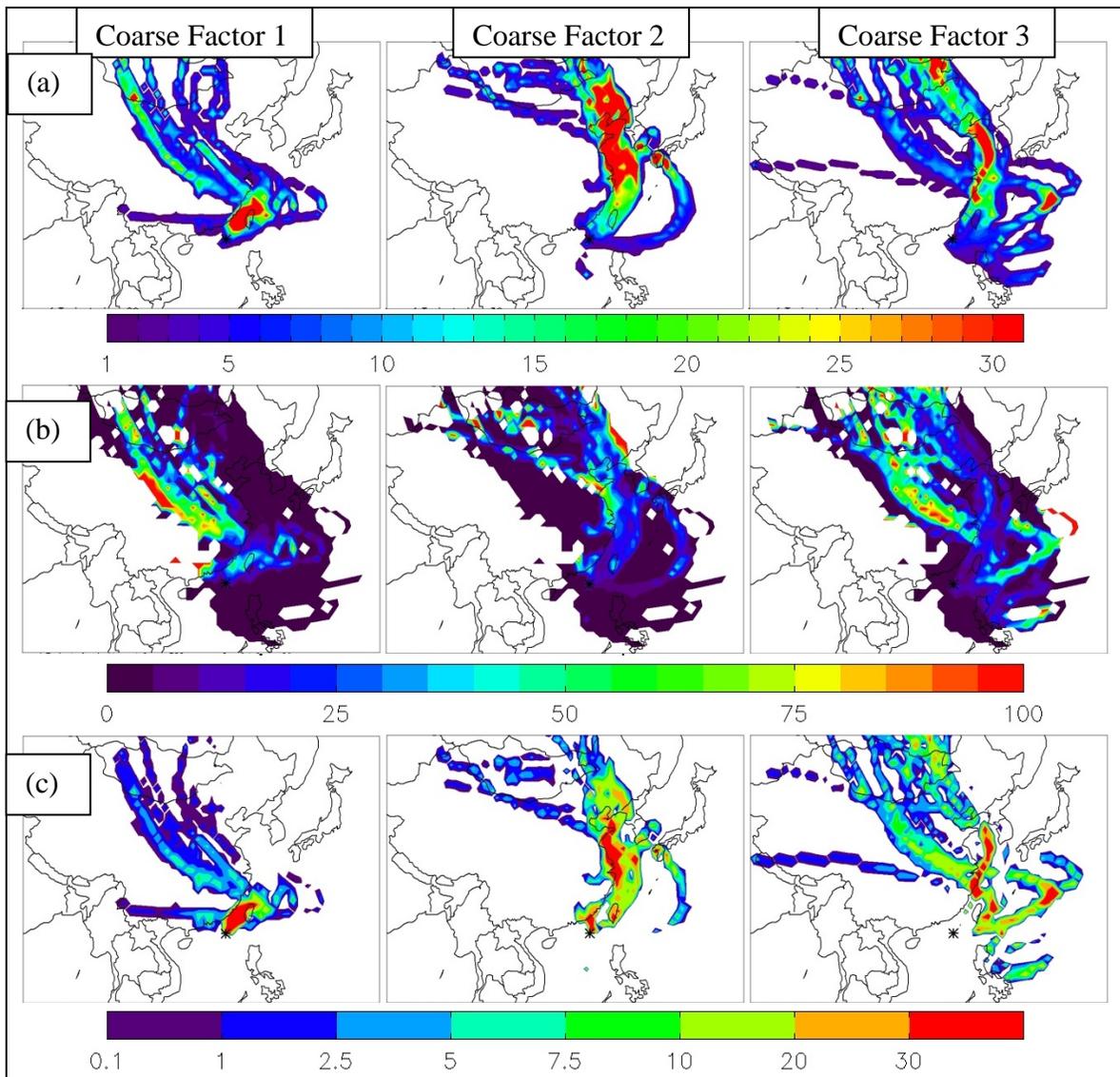


Figure 39 Dongsha Coarse-Mode Special Case backtrajectory plots for 100m Receptor Heights. Shown are (a) special case SCF, (b) PSCF, and (c) RTDF plots.

### 3.5. DISCUSSION AND ADDITIONAL ANALYSES

Thorough analysis of the nature of the aerosol environment at Dongsha requires examination of the results of all methods in relation to one another. Where the results agree, greater confidence can be established, whereas areas in which methods disagree may yield information about which aspects of the analyses have less certainty or utility in this circumstance. To that end, we examine the results of factor analysis and backtrajectory analysis alongside other modeled and observed data that are available for

the region, in order to better characterize the nature of the “whole column” of aerosol which exists at Dongsha in the springtime. The ultimate goal is to assess which aerosol sources and modes of transport we can reliably ascribe to the measurements at Dongsha Island in Spring 2010.

### **3.5.1. STUDY PERIOD – REGIONAL AVERAGE**

Two estimates of average Aerosol Optical Depths (AOD) during the Dongsha experiment period are shown in Figure 40 [*Giglio et al.*, 2006; *Zhang and Reid*, 2006]. Evidence of a significant biomass burning source of aerosol in Southeast Asia can be seen along with moderate AODs in East Asia associated with pollution sources.

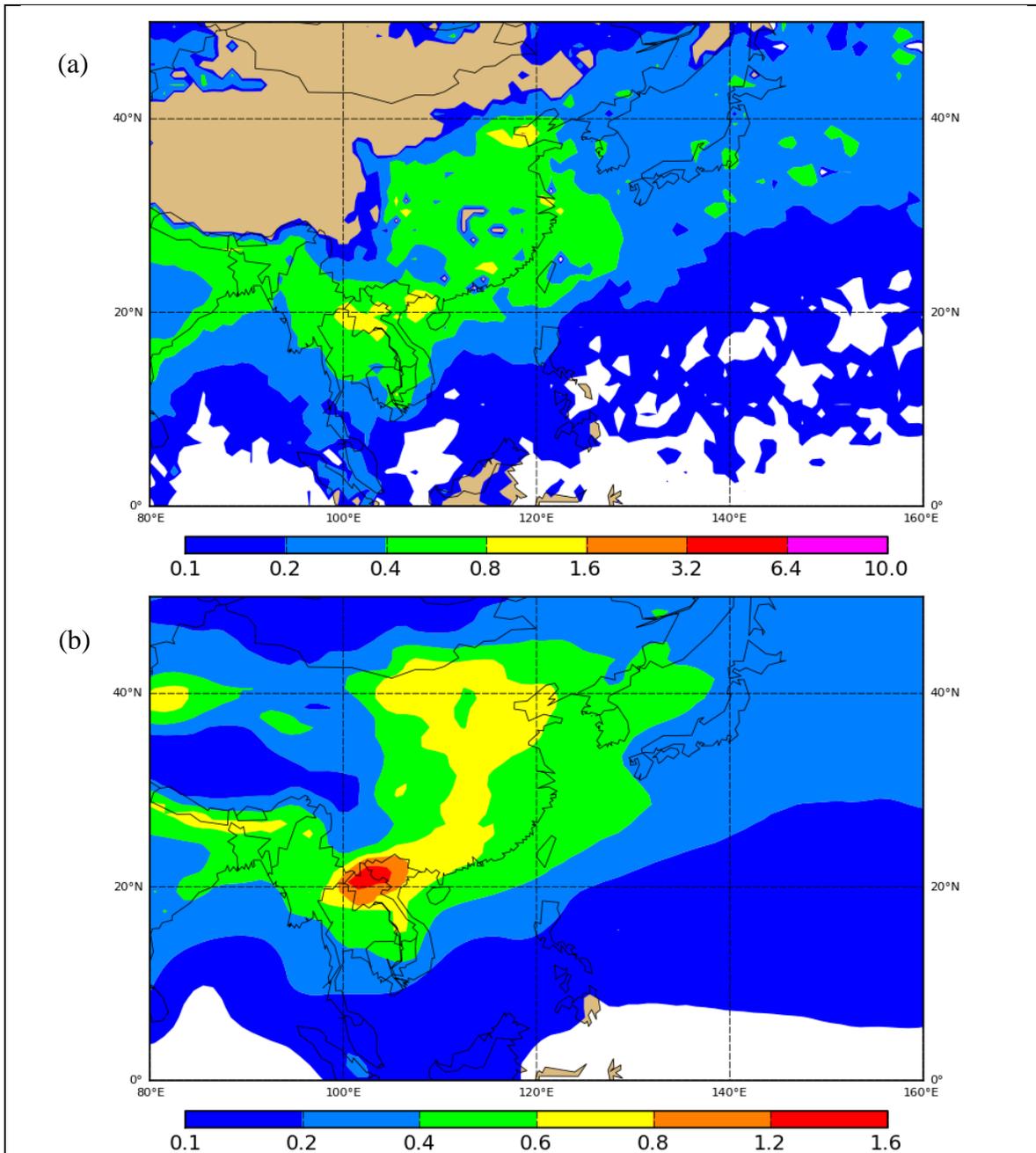


Figure 40 Regional average AODs during the Dongsha Experiment from (a) MODIS and (b) NAAPS.

### 3.5.2. STUDY PERIOD – DONGSHA TIME SERIES

In-situ measurements from AERONET and the Dongsha nephelometer are shown in Figure 41 (a) and (b). The AERONET radiometer measures column AOD and therefore is sensitive to the expected intermittent smoke plumes aloft along with other

aerosol distributed throughout the column [Holben *et al.*, 2001]. AODs, further split into expected contributions from fine and coarse mode aerosol, are incomplete in their coverage of the period, but nevertheless indicate a general picture of column aerosol conditions. Early in the study period, AODs primarily consisted of fine mode contributions consistent with smoke and pollution particles, with an AOD of 1.5 measured on 10 April indicating the possibility of a large smoke plume impacting the receptor at this time. By mid-April, fine mode contributions are below 0.5. Coarse mode extinction, more typical of dust influence, increased after 21 April, near the time when dust factors started to become more prominent in the DRUM data [O'Neill *et al.*, 2003].

The nephelometer measures surface aerosol and is therefore insensitive to smoke aloft that was detected by the radiometer unless that smoke is mixed down to the surface. The highest scattering coefficients detected during the study period were at two peaks centered around 24 and 27 April, the same time during which large scores for the two primary pollution factors (2 and 4) occurred. These factors consisted largely of accumulation mode aerosol that, depending on concentration, would be expected to scatter more significantly than coarse dust. Indeed, the nephelometer measured lower scattering coefficients during the period with higher dust factors starting around 28 April. Interestingly, earlier in the study period, between 31 March and roughly 23 April, light scattering seems to qualitatively align with the variability in DRUM factor 3, which was attributed to locally produced ultrafine sulfur. Ultrafine sulfur from a diesel generator or plumes from ships passing near the receptor would in fact be expected to produce significant scattering, which may be further evidence of a local source sporadically impacting the site.

The NAAPS model was run for the Dongsha receptor to produce predicted surface concentrations, along with column AODs separated by constituent contribution to extinction. The results of the model (Figure 41(c)) are encouraging, in that they are generally consistent with the DRUM, nephelometer, and AERONET radiometer, though the exact time stamps of events are not always aligned. Generally speaking, the NAAPS AODs tend to agree with the conceptual discussion of column aerosol from AERONET. A strong peak in AOD is predicted due to smoke aloft, though NAAPS has the peak arriving two days before the high AOD point from AERONET on 10 April. It should be noted that there were no valid retrievals from AERONET during the three days before this measurement; an earlier plume may have been present that persisted until the measured high AOD data point. Overall, the primary contributor to column aerosol in the NAAPS model is smoke, which is higher at the beginning of the study and tends to decrease towards the end of April and into May. Additionally, some elevated dust extinction is noted starting around 21 April, which is consistent with AERONET and DRUM dust arrival times.

NAAPS data for surface concentrations show relatively little impact from smoke, though the concentrations at the beginning of the study period do tend to be slightly higher than concentrations toward the end. The model would therefore seem to be mixing a small amount of smoke to the surface, though as a pure biomass burning signal is difficult to extract from either the DRUM or nephelometer data, actual detection of smoke at the surface is uncertain. The arrival of dust in the model coincides well with the 24 April peak in the DRUM coarse factor 2 dust signal, though NAAPS dust concentrations are decreasing by the time the coarse factor 1 dust plume peak arrives.

Nevertheless, the start and end times for the later Dongsha dust events in the DRUM data and NAAPS model agree well.

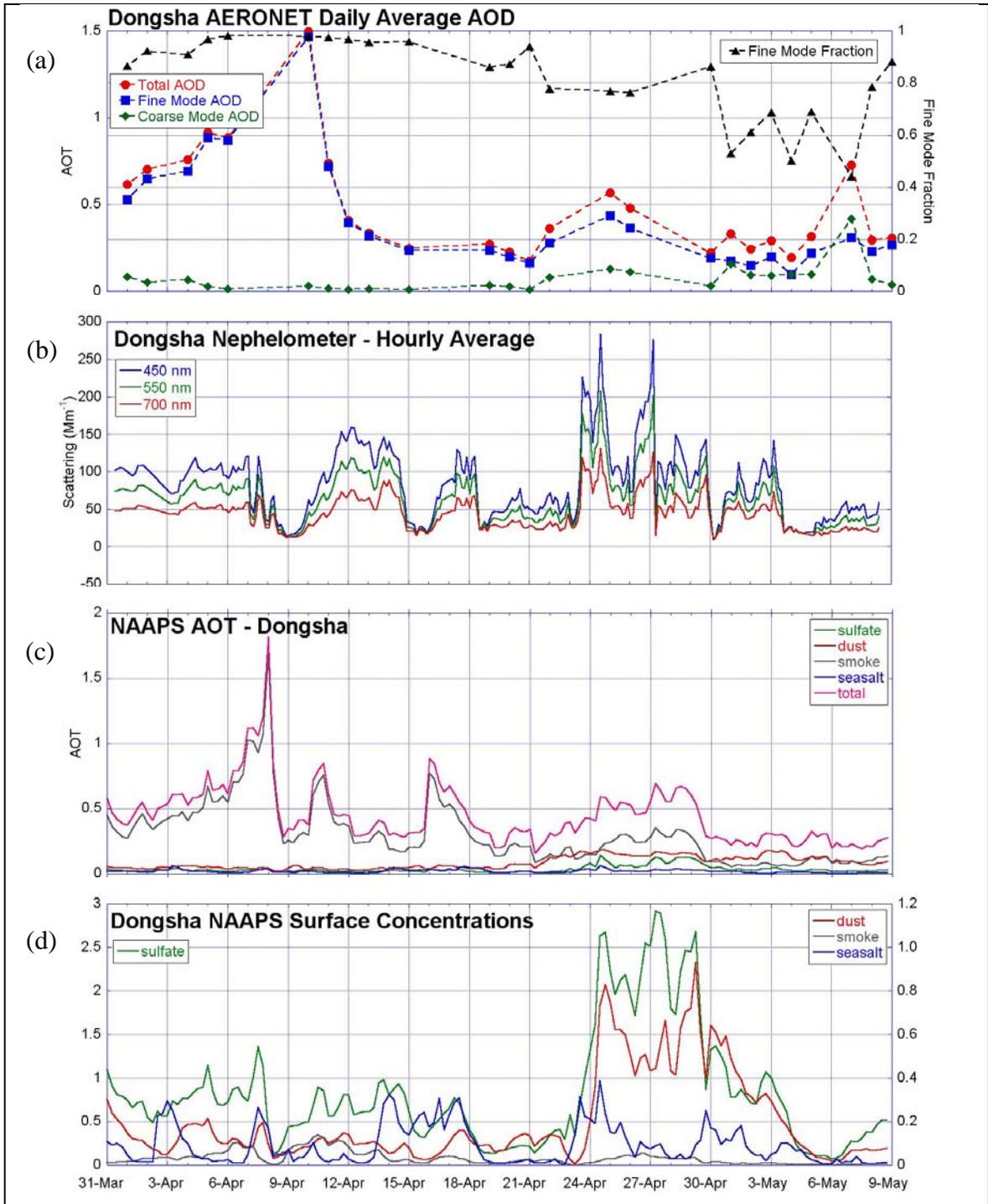


Figure 41 Dongsha receptor (a) daily averaged column AOD from AERONET, (b) backscattering from TSI three-wavelength nephelometer, (c) column AOD for various aerosol constituents from NAAPS, and (d) surface concentrations ( $\mu g/m^3$ ) for various aerosol constituents from NAAPS.

## 4. SUMMARY AND CONCLUSIONS

An initial study of atmospheric aerosol and transport into the South China Sea in springtime was conducted between 31 March and 9 May 2010 as part of the larger 7SEAS campaign in Southeast Asia. Measurements were taken at sea level on Dongsha Island, a remote atoll in the northern SCS, which served as one of two 7SEAS supersites set up for intensive measurement in the region over a multi-year timeframe. Size-resolved aerosol elemental composition data at three-hour time resolution were generated by X-ray fluorescence (XRF) analysis of aerosol collected on an eight-stage DRUM aerosol cascade impactor. The data were then subjected to factor analysis to extract common patterns of variability. Source types were assigned to each identified factor based on the correlation scores for the size-resolved elemental concentrations, and the elemental mass fractions associated with each factor. These factors were then used as a basis for interpreting the temporal variability of aerosol plumes impacting the Dongsha receptor.

Known aerosol sources in East and Southeast Asia include dust, anthropogenic pollution from industry and population centers, sea salt, and smoke from biomass burning. The elements analyzed by XRF were best suited to identification of aerosol from primary emissions including crustal elements associated with dust, heavy metals typical of industrial pollution, and chlorine and magnesium associated with sea salt. Sulfur is measured, but aerosol sulfate is largely formed from gas-phase precursors in secondary oxidation processes, and thus can be difficult to attribute to primary emissions.

Potassium, a commonly applied tracer of biomass burning, is quantified in the XRF analyses, but potassium is also associated with other types of sources such as dust that were frequently present in high concentration at the site. As a result, unequivocal attribution of aerosol to biomass burning sources was difficult using only DRUM data.

Additional analysis of the source/receptor relationship was conducted using the HYSPLIT model to generate backtrajectories every three hours throughout the study. During periods when a factor score was greater than 1, special case subsets of the backtrajectories were created for each identified factor. These special cases were then used to differentiate between factors and the normal background case based on the synoptic scale transport patterns that the model can resolve. A general indication of likely source regions for the general period of the study, and for each factor, was then established based on analysis of these backtrajectories. These analyses tended to confirm the assignments of sources using factor analysis, although the limited time period and generally consistent transport directions during the study resulted in only minimal differences between most of the special cases and the normal background case. Finally, additional data from a co-located three-wavelength nephelometer, a co-located AERONET radiometer, and aerosol forecasts from the U.S. Navy NAAPS model were used to qualitatively compare the timing, magnitude, and type of aerosol impacts at Dongsha.

The results of the Factor Analysis of the DRUM data indicated that dust, sea salt, and pollution had all impacted the site, and that the magnitude of these impacts varied throughout the study time period. A total of six factors associated with different aerosol chemical and size characteristics were extracted, showing that anthropogenic pollution

sources and sea salt tended to dominate the early part of the study, while a large dust plume impacted Dongsha towards the end of April and into early May. Using only the coarse-mode DRUM aerosol data in the factor analysis showed that two distinct dust signals were identified, along with a separate sea salt signal. While these results do not definitively indicate two separate dust sources, it does point to the likelihood that the dust measured at different times was chemically distinct, perhaps due to either different dust source regions, or different mixing or coagulation with pollution plumes during transport.

The large plume of dust and anthropogenic pollution at the end of the study dominated many of the factors as well as much of the total mass. As a result, many factors identified this time period strongly. Should such a plume not have occurred, it is possible that the same number of sources would not have appeared as clearly in the factor analysis. However, as some factors did identify earlier periods, it is likely that even without such a plume, the method would still have identified some potential source types.

Interestingly, considerable vertical wind shear in the SCS tended to yield significantly different backtrajectory source regions for air parcels arriving within or above the boundary layer over Dongsha. In addition, the AERONET radiometer, nephelometer, and NAAPS model were used to compare Aerosol Optical Depths indicative of average aerosol distributed throughout the entire atmospheric column to aerosol concentrations at the surface. The results showed that smoke was likely a significant contributor to column aerosol, but not to aerosol at the surface where DRUM and nephelometer measurements were taken. Dust and pollution also appeared to be more prevalent at the surface than throughout the rest of the column.

Based on these results, aerosol transport into the SCS in the springtime is understood to be generally dominated by low-level transport from regions throughout East Asia to the northwest, north, and northeast of Dongsha, while transport above the marine boundary layer originates in Southeast Asia to the island's west and southwest. Aerosol types that reached the surface at Dongsha during Spring 2010 were estimated to include dust from the Taklimakan and Gobi deserts of inland China and Mongolia; pollution from the major industrial and population centers of central and coastal China, Korea and Japan; and sea salt from marine areas surrounding the island. Based on basic measured and modeled column aerosol data, smoke from the significant biomass burning regions in Southeast Asia were expected to be advecting over Dongsha above the marine boundary layer with the general westerly flow patterns at those altitudes. The lack of a clear smoke-related potassium signal in the DRUM XRF data, taken together with the nephelometer and NAAPS model predictions, indicated that mixing of smoke to the surface was likely minimal, consistent with the strong inversions typical of the region.

Finally, as backtrajectory analysis and seasonal climatology show generally consistent springtime transport within the SCS from year to year, the general conclusions from this study regarding aerosol source areas and types may represent a good general description of potential aerosol impacts at Dongsha for air masses arriving in the springtime. Additionally, in agreement with previous studies in the region, the results of this work indicate that the vertical heterogeneity of the aerosol environment in the SCS is significant, and therefore the location and type of measurement (e.g. column vs. specific layer) should be considered during planning and analysis of measurements in the region.

Relying simply on AOD may not, in fact, provide appropriate indications of the nature of aerosol in the SCS.

As a final comment on the methodology described in this work, the two techniques of factor analysis and backtrajectory analysis are shown to be compatible methods when conducting source apportionment studies. Though backtrajectory analysis is not suitable for use on its own in this case study, it would be expected to achieve greater reliability and applicability with longer datasets. In such circumstances, the greater number of backtrajectories that can be associated with identified source types would be expected to reduce uncertainty and improve the special case transport signals against noise generated by attempting to associate the variability in aerosol chemistry with changes in modeled synoptic transport.

#### **4.1. FUTURE WORK**

As study of the aerosol environment in the SCS continues, several findings from this research may be relevant for future work. The vertical heterogeneity of the aerosol environment at Dongsha in the springtime implies that aerosol at the surface may be significantly different from that aloft. Additional study of the role of convection, turbulence, and wind shear in the vertical mixing of aerosol to the surface near its source region would aid in the interpretation of in-situ measurements downwind, especially as different climatic regimes occur in the region. Further, at the receptor site, understanding how aerosol is exchanged between the free troposphere and the boundary layer can also be helpful in identifying aerosol types and sources. In particular, the conditions under which smoke advecting above the marine boundary layer can be mixed to the surface need to be understood.

XRF analysis of the DRUM samples were promising for detection of pollution, dust and sea salt impacts. In particular, the high time resolution of this instrument allows for in depth analysis of transport patterns and plume impacts not always available with other aerosol collection methods. However, the DRUM is not as well suited for detection of smoke from biomass burning or secondary production of organic and inorganic aerosol from various pollution and combustion sources. The ability to better detect sulfate, nitrate, and organic aerosol in parallel would improve understanding of the aerosol types impacting the receptor. Furthermore, identifying the presence of smoke aerosol through detection of known biomass burning tracers such as levoglucosan [Sullivan *et al.*, 2008] would greatly improve the results of future studies in the SCS. Use of an instrument such as a Particle Into Liquid Sampler (PILS) to detect levoglucosan, sulfate, nitrate, and ammonium at time resolutions on the order of 15 minutes in conjunction with the DRUM instrument would greatly augment the results and improve the description of which sources may be impacting the receptor.

The principal component factor analysis method applied here has some shortcomings, most notably, the generation of negative factor scores that have no physical interpretation. It would be interesting to repeat the analysis using positive matrix factorization (PMF), a technique that is presently in common usage in the atmospheric chemistry community, and compare the results. Having a PMF tool available for future studies in the SCS region would be a valuable addition to the available methods for source apportionment.

Results can additionally be applied to the validation of aerosol models of the region such as NAAPS. Knowledge of the timing, relative size, and source types of plumes

moving across the SCS is of important to testing and validating these models, and can be provided by this type of analysis. Measurements at Dongsha Island can provide indications of which situations NAAPS may be better at predicting and help to constrain where future improvements to the model should be focused.

Lastly, the methods described in this work are easily adapted to these potential improvements in the chemical and meteorological understanding of the Dongsha aerosol environment. Factor analysis can be extended to include PILS data in addition to DRUM data to analyze aerosol variability and extract factors based on this additional information. Other statistical techniques can also be used to help differentiate between long-range transport and local aerosol impacts. Multiple regression methods can be used to analyze the variability of a aerosol signals with respect to various potential predictors such as HYSPLIT backtrajectories, measured local wind conditions, and atmospheric characteristics such as stability. Ultimately, as the Dongsha and 7SEAS experiments move forward, the use of a variety of methods and measurements in conjunction with each other can improve our understanding of the aerosol environment and its effects on climate, air quality, and aerosol-cloud-precipitation effects in the South China Sea.

## 5. REFERENCES

7SEAS-Whitepaper (March 4, 2009), A Program to Study Pollution-Meteorology Feedbacks in Southeast Asia: The Seven SouthEast Asian Studies (7SEAS) Mission, edited.

Anderson, T. L., and J. A. Ogren (1998), Determining aerosol radiative properties using the TSI 3563 integrating nephelometer, *Aerosol Science and Technology*, 29(1), 57-69.

Ashbaugh, L. L., W. C. Malm, and W. Z. Sadeh (1985), A RESIDENCE TIME PROBABILITY ANALYSIS OF SULFUR CONCENTRATIONS AT GRAND-CANYON-NATIONAL-PARK, *Atmospheric Environment*, 19(8), 1263-1270.

Cahill, T. A., C. Goodart, J. W. Nelson, R. A. Eldred, J. S. Nasstrom, and P. J. Feeny (1985), Design and evaluation of the DRUM impactor, in *Proceedings of the International Symposium on Particulate and Multiphase Processes*, edited by T. Ariman and T. Nejat, pp. 319-325, Washington, D. C.

Cliff, S. S. (2012), Personal Communication, edited.

Cohen, D. D., J. Crawford, E. Stelcer, and V. T. Bac (2010a), Characterisation and source apportionment of fine particulate sources at Hanoi from 2001 to 2008, *Atmospheric Environment*, 44(3), 320-328.

Cohen, D. D., J. Crawford, E. Stelcer, and T. B. Vuong (2010b), Long range transport of fine particle windblown soils and coal fired power station emissions into Hanoi between 2001 to 2008, *Atmospheric Environment*, 44(31), 3761-3769.

Cohen, D. D., J. Crawford, E. Stelcer, and V. T. Bac (2010c), Long range transport of fine particle windblown soils and coal fired power station emissions into Hanoi between 2001 to 2008, *Atmospheric Environment*, 44(31), 3761-3769.

Coutant, B. W., J. Engel-Cox, and K. E. Swinton (2003), Compilation of Existing Studies on Source Apportionment for PM<sub>2.5</sub>, in *Second Draft Technical Report (Revision 1)*

Contract No. 68-D-02-061 Work Assignment 1-05, edited, U.S. Environmental Protection Agency.

Draxler, R. R. (2004), HYSPLIT4 users's guide, edited, Air Resources Laboratory, Silver Spring, Md.

Draxler, R. R., and G. D. Hess (1997), *Description of the HYSPLIT4 modeling system*, U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, Air Resources Laboratory, Silver Spring, Md.

Draxler, R. R., and G. D. Hess (1998), An overview of the HYSPLIT\_4 modelling system for trajectories, dispersion and deposition, *Australian Meteorological Magazine*, 47(4), 295-308.

Engelbrecht, J. P., E. V. McDonald, J. A. Gillies, and A. W. Gertler (2008), Department of Defense Enhanced Particulate Matter Surveillance Program (EPMSP), edited, Desert Research Institute, Report Contract Number: W9124R-05-C-0135/SUBCLIN 000101-ACRNAB.

Giglio, L., I. Csizsar, and C. O. Justice (2006), Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors, *Journal of Geophysical Research-Biogeosciences*, 111(G2).

Haller, M., and A. Knochel (1996), X-ray fluorescence analysis using synchrotron radiation (SYXRF), *Journal of Trace and Microprobe Techniques*, 14(3), 461-488.

Han, J. S., K. J. Moon, S. Y. Ryu, Y. J. Kim, and K. D. Perry (2005), Source estimation of anthropogenic aerosols collected by a DRUM sampler during spring of 2002 at Gosan, Korea, *Atmospheric Environment*, 39(17), 3113-3125.

Hayasaka, T., S. Satake, A. Shimizu, N. Sugimoto, I. Matsui, K. Aoki, and Y. Muraji (2007), Vertical distribution and optical properties of aerosols observed over Japan during the atmospheric brown clouds-east Asia regional experiment 2005, *Journal of Geophysical Research-Atmospheres*, 112(D22), 12.

Henry, R. C. (1987), CURRENT FACTOR-ANALYSIS RECEPTOR MODELS ARE ILL-POSED, *Atmospheric Environment*, 21(8), 1815-1820.

Henry, R. C. (1991), Chapter 5 Multivariate Receptor Models, in *Data Handling in Science and Technology*, edited by K. H. Philip, pp. 117-147, Elsevier.

Hogan, T. F., and T. E. Rosmond (1991), THE DESCRIPTION OF THE NAVY OPERATIONAL GLOBAL ATMOSPHERIC PREDICTION SYSTEMS SPECTRAL FORECAST MODEL, *Monthly Weather Review*, 119(8), 1786-1815.

Holben, B. N., et al. (2001), An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET, *Journal of Geophysical Research-Atmospheres*, 106(D11), 12067-12097.

Hopke, P. K. A Guide to Positive Matrix Factorization, edited, U.S. Environmental Protection Agency.

Hopke, P. K. (1985), *Receptor modeling in environmental chemistry*, 319 pp., Wiley, New York.

Hopke, P. K. (2003), Recent developments in receptor modeling, *Journal of Chemometrics*, 17(5), 255-265.

Huang, S. L., K. A. Rahn, and R. Arimoto (1999), Testing and optimizing two factor-analysis techniques on aerosol at Narragansett, Rhode Island, *Atmospheric Environment*, 33(14), 2169-2185.

Juntto, S., and P. Paatero (1994), ANALYSIS OF DAILY PRECIPITATION DATA BY POSITIVE MATRIX FACTORIZATION, *Environmetrics*, 5(2), 127-144.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bulletin of the American Meteorological Society*, 77(3), 437-471.

Kim, E., P. K. Hopke, T. V. Larson, and D. S. Covert (2004), Analysis of ambient particle size distributions using unmix and positive matrix factorization, *Environmental Science & Technology*, 38(1), 202-209.

Kim, S. W., S. C. Yoon, J. Kim, J. Y. Kang, and N. Sugimoto (2010), Asian dust event observed in Seoul, Korea, during 29-31 May 2008: Analysis of transport and vertical distribution of dust particles from lidar and surface measurements, *Science of the Total Environment*, 408(7), 1707-1718.

O'Neill, N. T., T. F. Eck, A. Smirnov, B. N. Holben, and S. Thulasiraman (2003), Spectral discrimination of coarse and fine mode optical depth, *Journal of Geophysical Research-Atmospheres*, 108(D17), 15.

Paatero, P., and U. Tapper (1993), ANALYSIS OF DIFFERENT MODES OF FACTOR-ANALYSIS AS LEAST-SQUARES FIT PROBLEMS, *Chemometrics and Intelligent Laboratory Systems*, 18(2), 183-194.

Paatero, P., and U. Tapper (1994), POSITIVE MATRIX FACTORIZATION - A NONNEGATIVE FACTOR MODEL WITH OPTIMAL UTILIZATION OF ERROR-ESTIMATES OF DATA VALUES, *Environmetrics*, 5(2), 111-126.

Pallant, J. (2007), *SPSS survival manual : a step by step guide to data analysis using SPSS for Windows*, Open University Press, Maidenhead.

Perry, K. D., S. S. Cliff, and M. P. Jimenez-Cruz (2004), Evidence for hygroscopic mineral dust particles from the Intercontinental Transport and Chemical Transformation Experiment, *Journal of Geophysical Research-Atmospheres*, 109(D23), 9.

Poirot, R. L., P. R. Wishinski, P. K. Hopke, and A. V. Polissar (2001), Comparative application of multiple receptor methods to identify aerosol sources in northern Vermont, *Environmental Science & Technology*, 35(23), 4622-4636.

Polissar, A. V., P. K. Hopke, and J. M. Harris (2001a), Source regions for atmospheric aerosol measured at Barrow, Alaska, *Environmental Science & Technology*, 35(21), 4214-4226.

Polissar, A. V., P. K. Hopke, and R. L. Poirot (2001b), Atmospheric aerosol over Vermont: Chemical composition and sources, *Environmental Science & Technology*, 35(23), 4604-4621.

Polissar, A. V., P. K. Hopke, P. Paatero, Y. J. Kaufmann, D. K. Hall, B. A. Bodhaine, E. G. Dutton, and J. M. Harris (1999), The aerosol at Barrow, Alaska: long-term trends and source locations, *Atmospheric Environment*, 33(16), 2441-2458.

Raabe, O. G., D. A. Braaten, R. L. Axelbaum, S. V. Teague, and T. A. Cahill (1988), CALIBRATION STUDIES OF THE DRUM IMPACTOR, *Journal of Aerosol Science*, 19(2), 183-195.

Reid, E. A., J. S. Reid, M. M. Meier, M. R. Dunlap, S. S. Cliff, A. Broumas, K. Perry, and H. Maring (2003), Characterization of African dust transported to Puerto Rico by individual particle and size segregated bulk analysis, *Journal of Geophysical Research-Atmospheres*, 108(D19), 22.

Reid, J. S., E. A. Reid, A. Walker, S. Piketh, S. Cliff, A. Al Mandoos, S. C. Tsay, and T. F. Eck (2008), Dynamics of southwest Asian dust particle size characteristics with implications for global dust research, *Journal of Geophysical Research-Atmospheres*, 113(D14).

Reid, J. S., et al. (2003), Comparison of size and morphological measurements of coarse mode dust particles from Africa, *Journal of Geophysical Research-Atmospheres*, 108(D19), 28.

Reid, J. S., et al. (2009), Global Monitoring and Forecasting of Biomass-Burning Smoke: Description of and Lessons From the Fire Locating and Modeling of Burning Emissions (FLAMBE) Program, *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2(3), 144-162.

Satake, S., et al. (2004), Characteristics of Asian aerosol transport simulated with a regional-scale chemical transport model during the ACE-Asia observation, *Journal of Geophysical Research-Atmospheres*, 109(D19), 16.

Seinfeld, J. H., and S. N. Pandis (2006), *Atmospheric chemistry and physics from air pollution to climate change*, 2nd ed. ed., 1203 pp., Wiley, Hoboken, N.J.

Stohl, A. (1996), Trajectory statistics - A new method to establish source-receptor relationships of air pollutants and its application to the transport of particulate sulfate in Europe, *Atmospheric Environment*, 30(4), 579-587.

Stohl, A., C. Forster, A. Frank, P. Seibert, and G. Wotawa (2005), Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, *Atmospheric Chemistry and Physics*, 5, 2461-2474.

Sullivan, A. P., A. S. Holden, L. A. Patterson, G. R. McMeeking, S. M. Kreidenweis, W. C. Malm, W. M. Hao, C. E. Wold, and J. L. Collett (2008), A method for smoke marker measurements and its potential application for determining the contribution of biomass burning from wildfires and prescribed fires to ambient PM(2.5) organic carbon, *Journal of Geophysical Research-Atmospheres*, 113, 14.

Tabachnick, B. G., and L. S. Fidell (2008), *Using multivariate statistics*, Pearson/Allyn & Bacon, Boston.

Takemi, T., M. Yasui, J. X. Zhou, and L. C. Liu (2006), Role of boundary layer and cumulus convection on dust emission and transport over a midlatitude desert area, *Journal of Geophysical Research-Atmospheres*, 111(D11), 17.

VanCuren, R. A., S. S. Cliff, K. D. Perry, and M. Jimenez-Cruz (2005), Asian continental aerosol persistence above the marine boundary layer over the eastern North Pacific: Continuous aerosol measurements from Intercontinental Transport and Chemical Transformation 2002 (ITCT 2K2), *Journal of Geophysical Research-Atmospheres*, 110(D9).

Wang, S. H., et al. (2011), First detailed observations of long-range transported dust over the northern South China Sea, *Atmospheric Environment*, 45(27), 4804-4808.

Wang, X., J. P. Huang, M. X. Ji, and K. Higuchi (2008), Variability of East Asia dust events and their long-term trend, *Atmospheric Environment*, 42(13), 3156-3165.

Watson, J. G., L. W. A. Chen, J. C. Chow, P. Doraiswamy, and D. H. Lowenthal (2008), Source apportionment: Findings from the US Supersites program, *Journal of the Air & Waste Management Association*, 58(2), 265-288.

Zeng, Y., and P. K. Hopke (1989), A STUDY OF THE SOURCES OF ACID PRECIPITATION IN ONTARIO, CANADA, *Atmospheric Environment*, 23(7), 1499-1509.

Zhang, J. L., and J. S. Reid (2006), MODIS aerosol product analysis for data assimilation: Assessment of over-ocean level 2 aerosol optical thickness retrievals, *Journal of Geophysical Research-Atmospheres*, 111(D22), 17.

## Appendix A

### **Matrix Effects Correction**

Self absorption of photons emitted during XRF analysis is known to occur for elements lighter than calcium, particularly for larger particles, and requires examination to determine if a correction factor should be applied [E A Reid *et al.*, 2003]. The ratio of light crustal elements to Fe is expected to remain roughly constant with increasing concentration of Fe. If matrix effects are present, the ratio will decrease and a linear correction factor can be applied if effects are significant. Shown in the following figures are these plots for Mg, Al, and Si. Some possible matrix effects are seen for Mg in size fraction 2, but as Mg is associated with sea salt in addition to dust, large sea salt concentrations can skew the ratios higher. Despite this possibility, any matrix effects in these elements were judged to be small and only potentially impacting a small number of data points; as a result, no correction was applied.

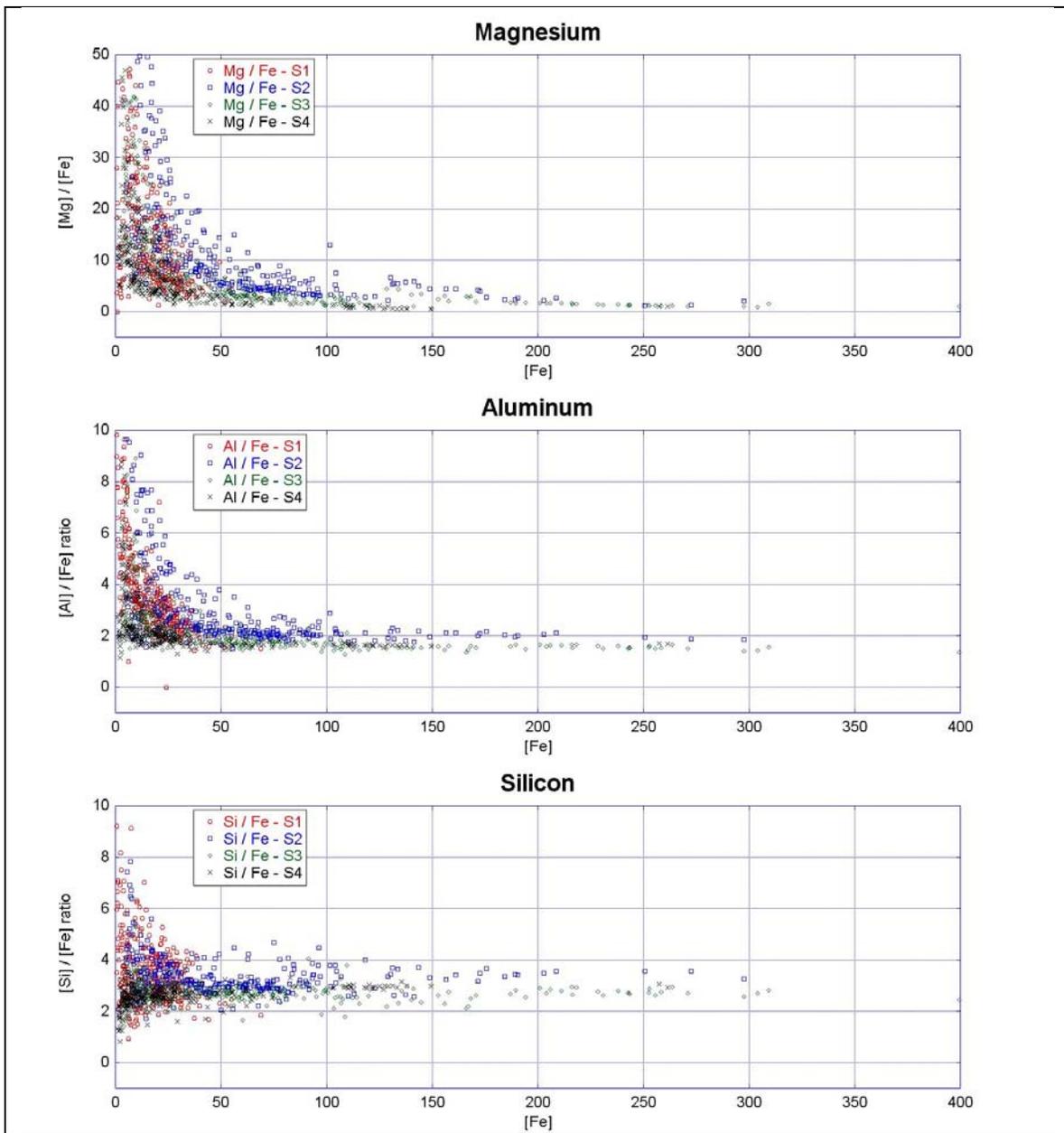


Figure 42 Concentration ratios ( $\text{ng} / \text{m}^3$ ) against Fe for light elements in Dongsha DRUM data.

## Appendix B

### Three-Mode Factor Analysis

Table 5 KMO and Bartlett's Test Statistics for the Dongsha Three Mode Factors

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.891
Bartlett's Test of Sphericity Approx. Chi-Square	38843.021
df	3240
Sig.	.000

Table 6 Total Variance Explained for Dongsha Three Mode Factors

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	22.995	28.389	28.389	22.995	28.389	28.389	17.963	22.177	22.177
2	7.969	9.838	38.227	7.969	9.838	38.227	9.078	11.207	33.384
3	6.549	8.085	46.312	6.549	8.085	46.312	7.546	9.316	42.700
4	4.807	5.935	52.247	4.807	5.935	52.247	5.924	7.313	50.013
5	3.909	4.826	57.072	3.909	4.826	57.072	4.724	5.832	55.845
6	3.349	4.135	61.208	3.349	4.135	61.208	4.344	5.362	61.208

Table 7 Scree Plot for Dongsha Three Mode Factors

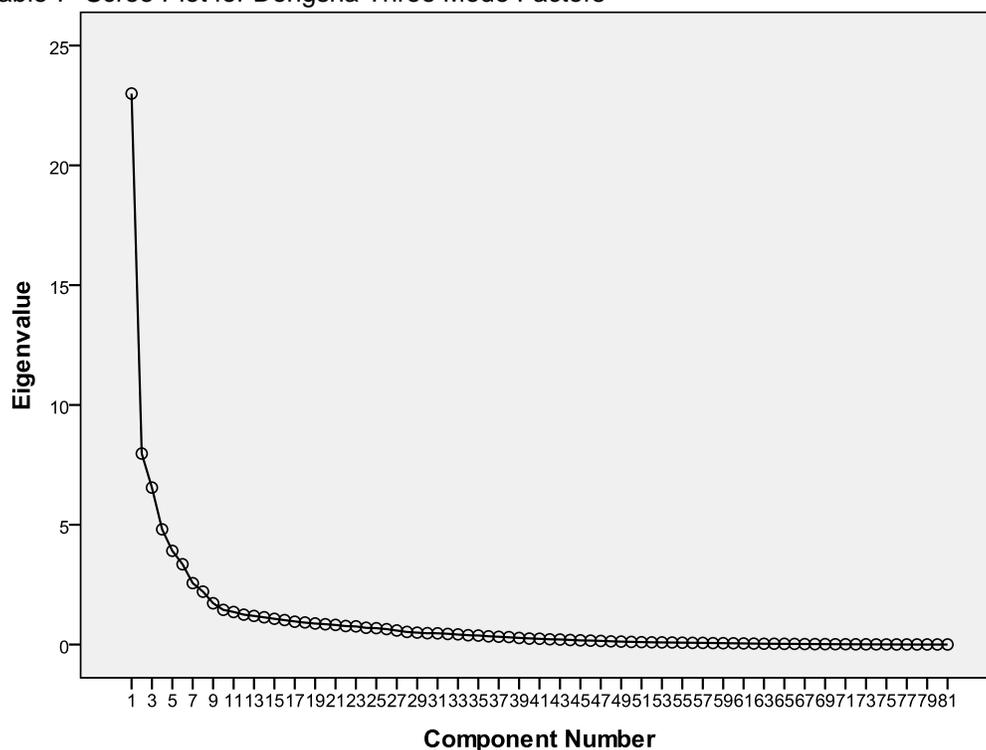


Table 8 Rotated Component Matrix for Dongsha Three Mode Factors

	Factor							Factor						
	1	2	3	4	5	6		1	2	3	4	5	6	
Mg - Coarse					.799		Ni - Accum		.622					
Al - Coarse	.928						Cu - Accum		.662					
Si - Coarse	.927						Zn - Accum		.895					
P - Coarse					.747		Ga - Accum				.482	-.319		
S - Coarse		.306	.456		.487		As - Accum		.908					
Cl - Coarse					.882		Se - Accum		.570					.448
K - Coarse	.804						Br - Accum		.806					
Ca - Coarse	.915						Rb - Accum							.677
Ti - Coarse	.896						Sr - Accum	.617						
V - Coarse					.686		Y - Accum		.656					
Cr - Coarse		.490		.331			Zr - Accum		-.311					
Mn - Coarse	.900						Mo - Accum		.734					
Fe - Coarse	.923						Pb - Accum		.836					
Co - Coarse	.815						Mg - Ultrafine			.947				
Ni - Coarse							Al - Ultrafine			.913				
Cu - Coarse				-.418	.318		Si - Ultrafine	.495		.516	.403			
Zn - Coarse		.644		.496			P - Ultrafine			.945				
Ga - Coarse	.553						S - Ultrafine			.957				
As - Coarse					.327		Cl - Ultrafine							
Se - Coarse	-.562						K - Ultrafine			.836				
Br - Coarse	.333			.419	.598		Ca - Ultrafine	.678			.473			
Rb - Coarse	.790						Ti - Ultrafine	.657			.433			
Sr - Coarse	.518		.349		.308		V - Ultrafine			.670				
Y - Coarse	-.517						Cr - Ultrafine				.636			
Zr - Coarse	.349						Mn - Ultrafine	.361			.787			
Mo - Coarse							Fe - Ultrafine	.653			.626			
Pb - Coarse	.926						Co - Ultrafine							
Mg - Accum					.357	.800	Ni - Ultrafine			.753				
Al - Accum	.784					.522	Cu - Ultrafine						.312	
Si - Accum	.834	.341					Zn - Ultrafine	.347			.715			
P - Accum						.883	Ga - Ultrafine							
S - Accum						.873	As - Ultrafine	.363			.668			
Cl - Accum					.566		Se - Ultrafine			.414				
K - Accum	.558					.709	Br - Ultrafine			.689				
Ca - Accum	.937						Rb - Ultrafine							
Ti - Accum	.936						Sr - Ultrafine							
V - Accum	.360	.733					Y - Ultrafine							
Cr - Accum		.473		.496			Zr - Ultrafine							
Mn - Accum	.442	.789		.306			Mo - Ultrafine							
Fe - Accum	.901						Pb - Ultrafine				.646			
Co - Accum	.857													

Extraction Method: Principal Component Analysis.  
 Rotation Method: Varimax with Kaiser Normalization.  
 a. Rotation converged in 7 iterations.

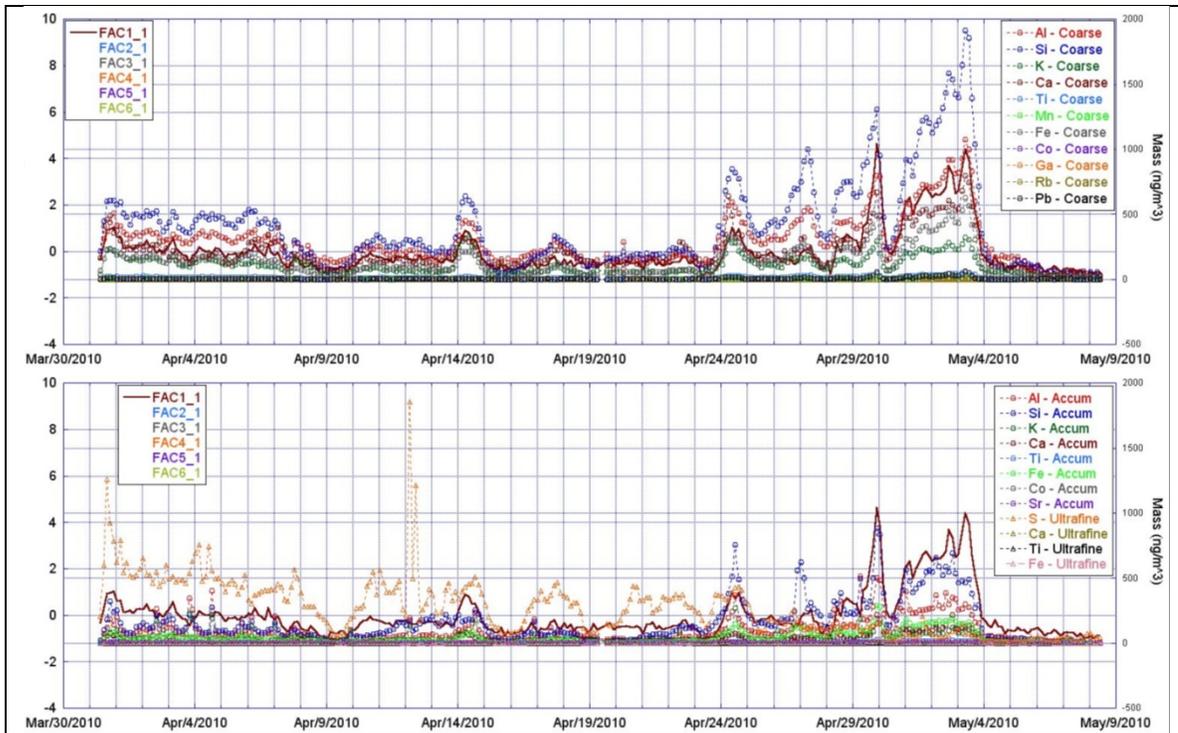


Figure 43 Three-Mode Factor 1 time series with coarse mode (top), and accumulation and ultrafine mode (bottom) concentrations ( $\text{ng} / \text{m}^3$ ) for elements with high factor 1 component scores.

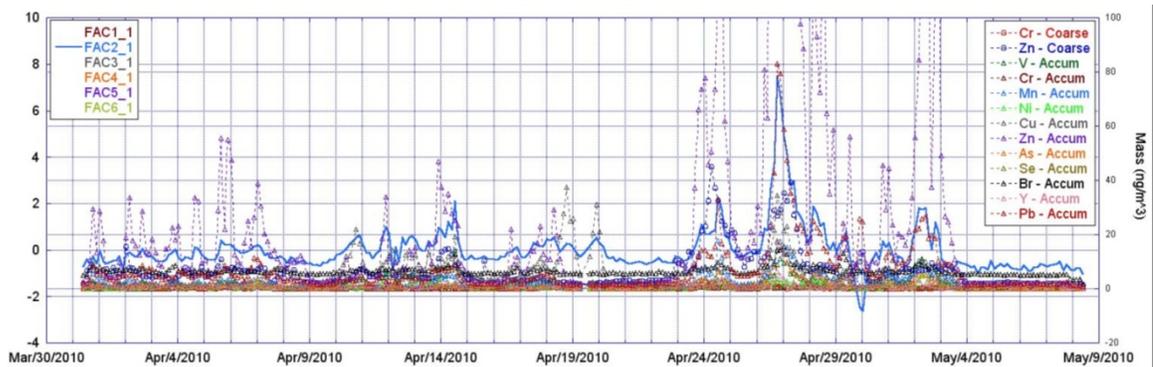


Figure 44 Three-Mode Factor 2 time series (similar to Figure B-1) for elements with high factor 2 component scores. Note that accumulation mode Zn concentration is often greater than scale shown, with peaks reaching as high as  $500 \text{ ng} / \text{m}^3$ .

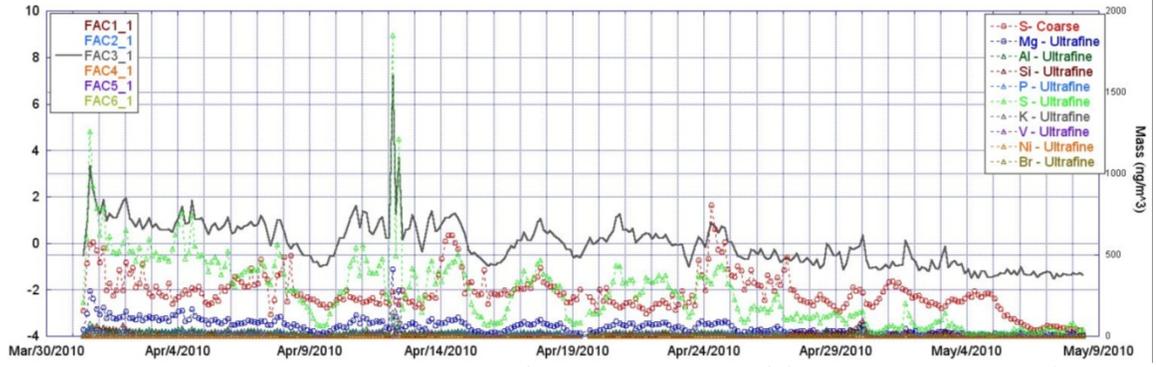


Figure 45 Three-Mode Factor 3 time series (similar to Figure B-1) for elements with high factor 3 component scores.

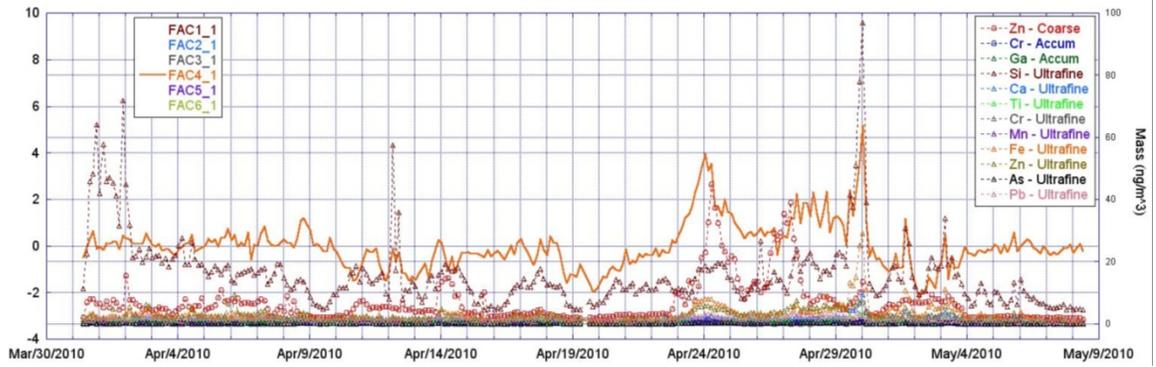


Figure 46 Three-Mode Factor 4 time series (similar to Figure B-1) for elements with high factor 4 component scores.

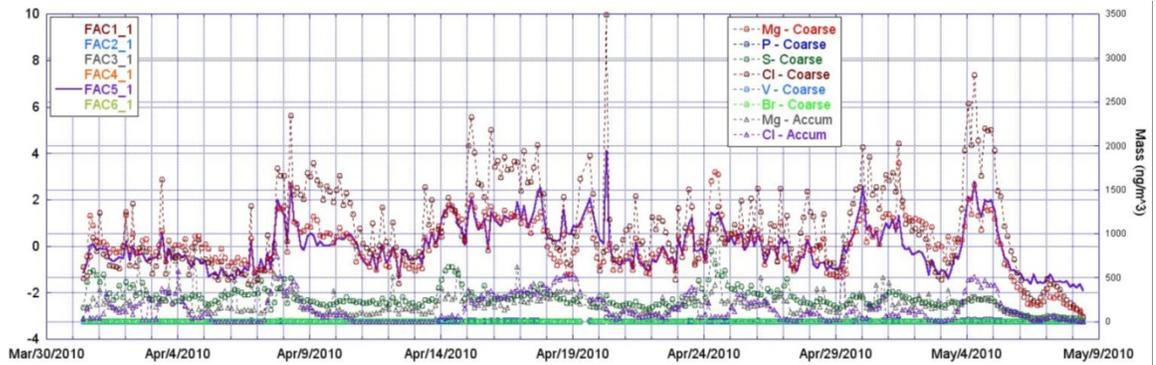


Figure 47 Three-Mode Factor 5 time series (similar to Figure B-1) for elements with high factor 5 component scores.

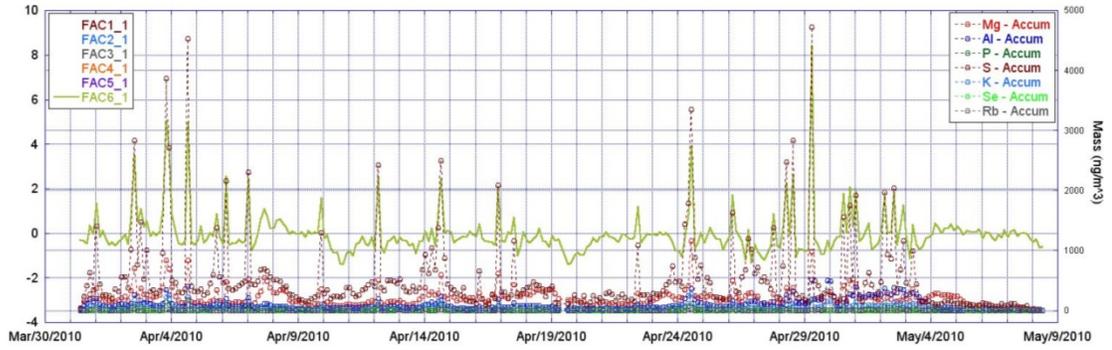


Figure 48 Three-Mode Factor 6 time series (similar to Figure B-1) for elements with high factor 6 component scores.

## Appendix C

### Coarse-Mode Factor Analysis

Table 9 KMO and Bartlett's Test Statistics for the Dongsha Coarse Mode Factors

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.894
Bartlett's Test of Sphericity Approx. Chi-Square	11053.331
df	351
Sig.	.000

Table 10 Total Variance Explained for Dongsha Coarse Mode Factors

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.157	41.321	41.321	11.157	41.321	41.321	8.798	32.586	32.586
2	3.852	14.266	55.587	3.852	14.266	55.587	4.899	18.145	50.730
3	2.179	8.070	63.656	2.179	8.070	63.656	3.490	12.926	63.656

Table 11 Scree Plot for Dongsha Coarse Mode Factors

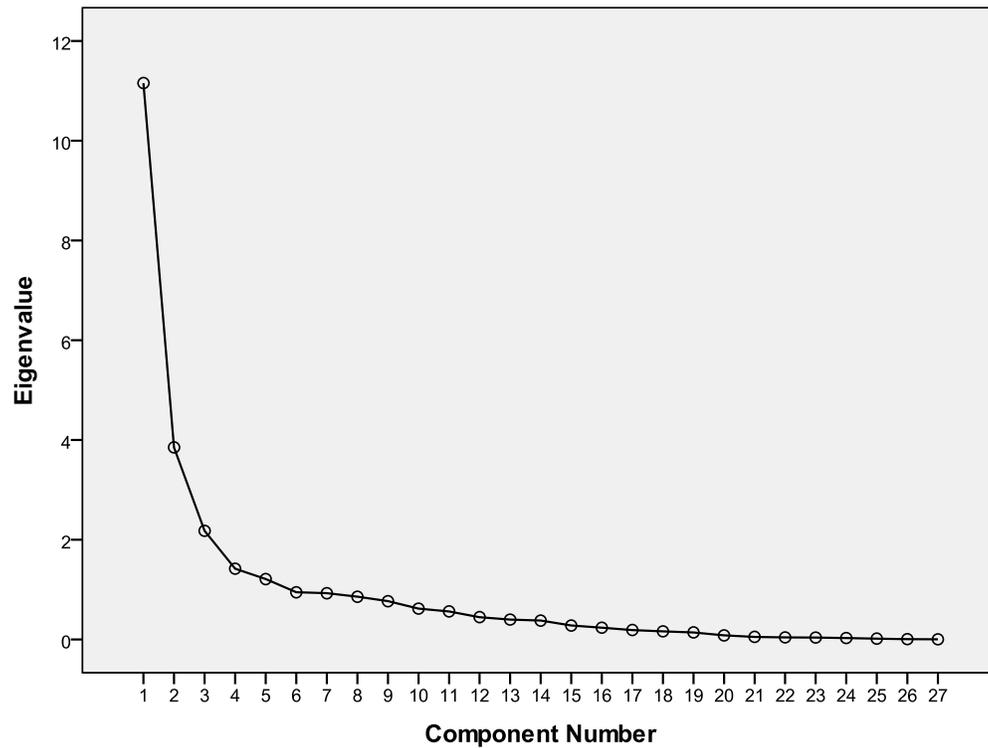


Table 12 Rotated Component Matrix for Dongsha Coarse Mode Factors

	Factor		
	1	2	3
Mg - Coarse		.577	.676
Al - Coarse	.876	.446	
Si - Coarse	.902	.372	
P - Coarse			.897
S - Coarse		.898	
Cl - Coarse			.933
K - Coarse	.673	.672	
Ca - Coarse	.833	.434	
Ti - Coarse	.866	.428	
V - Coarse			.823
Cr - Coarse		.379	-.334
Mn - Coarse	.860	.412	
Fe - Coarse	.881	.426	
Co - Coarse	.747	.466	
Ni - Coarse			
Cu - Coarse			
Zn - Coarse		.751	
Ga - Coarse	.626		
As - Coarse			.409
Se - Coarse	-.543	-.349	
Br - Coarse		.686	.406
Rb - Coarse	.853		
Sr - Coarse	.321	.797	
Y - Coarse	-.567		
Zr - Coarse	.462		
Mo - Coarse	.326		
Pb - Coarse	.936		

Extraction Method: Principal Component Analysis.  
 Rotation Method: Varimax with Kaiser Normalization.  
 a. Rotation converged in 5 iterations.

## Appendix D

### Additional Special Case Backtrajectory Analysis

Additional figures for the Dongsha special case backtrajectory analysis are shown below. Generally, the source regions can be seen shifting from north to south with increasing receptor height along with the shift in the normal case source regions. Consistent with the understanding of predominantly lower level boundary layer transport to the Dongsha surface, the expected source region plots tend to have less utility with increasing receptor height.

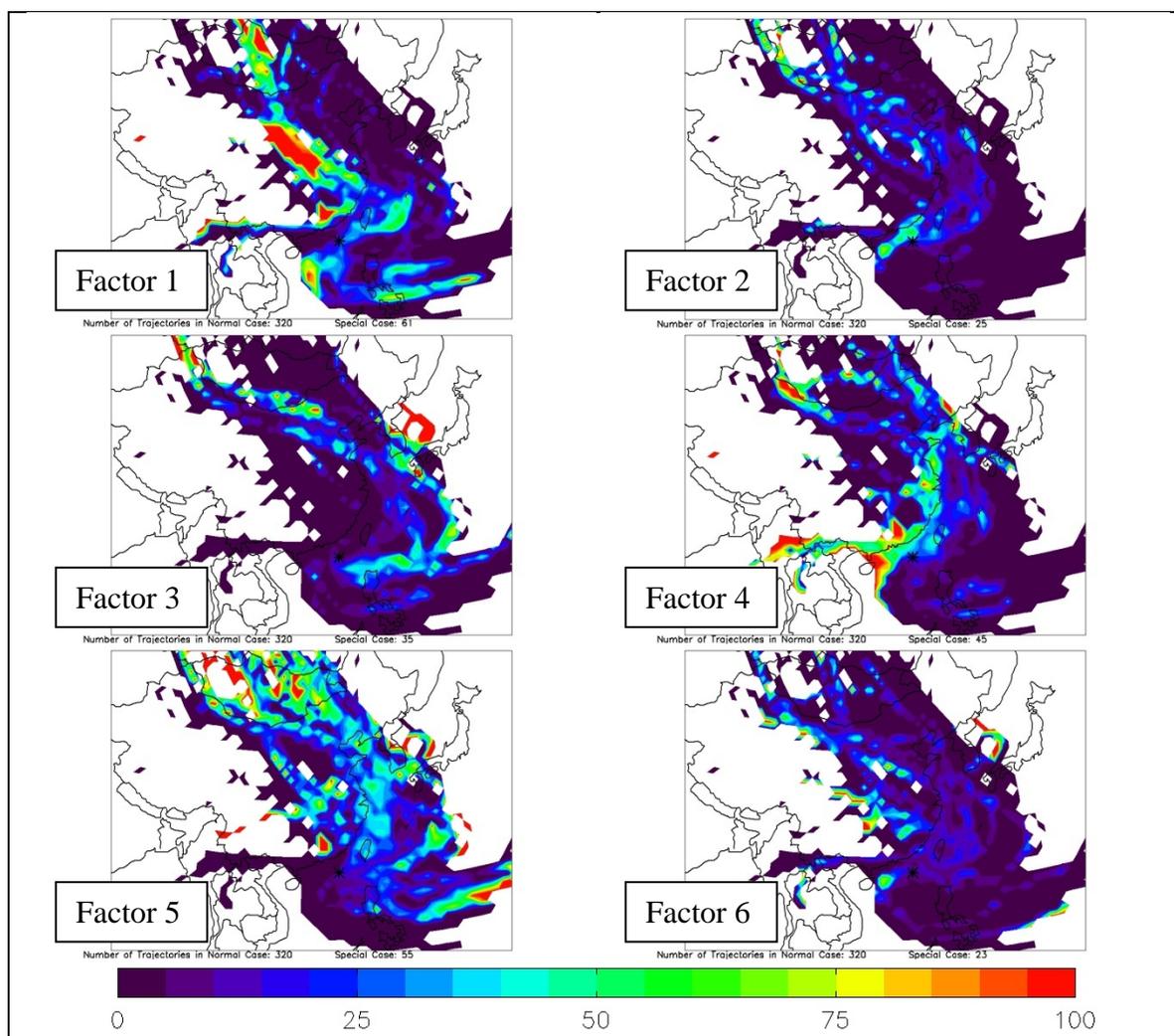


Figure 49 Dongsha Three Mode Special Case PSCF plots for 500m Receptor Heights.

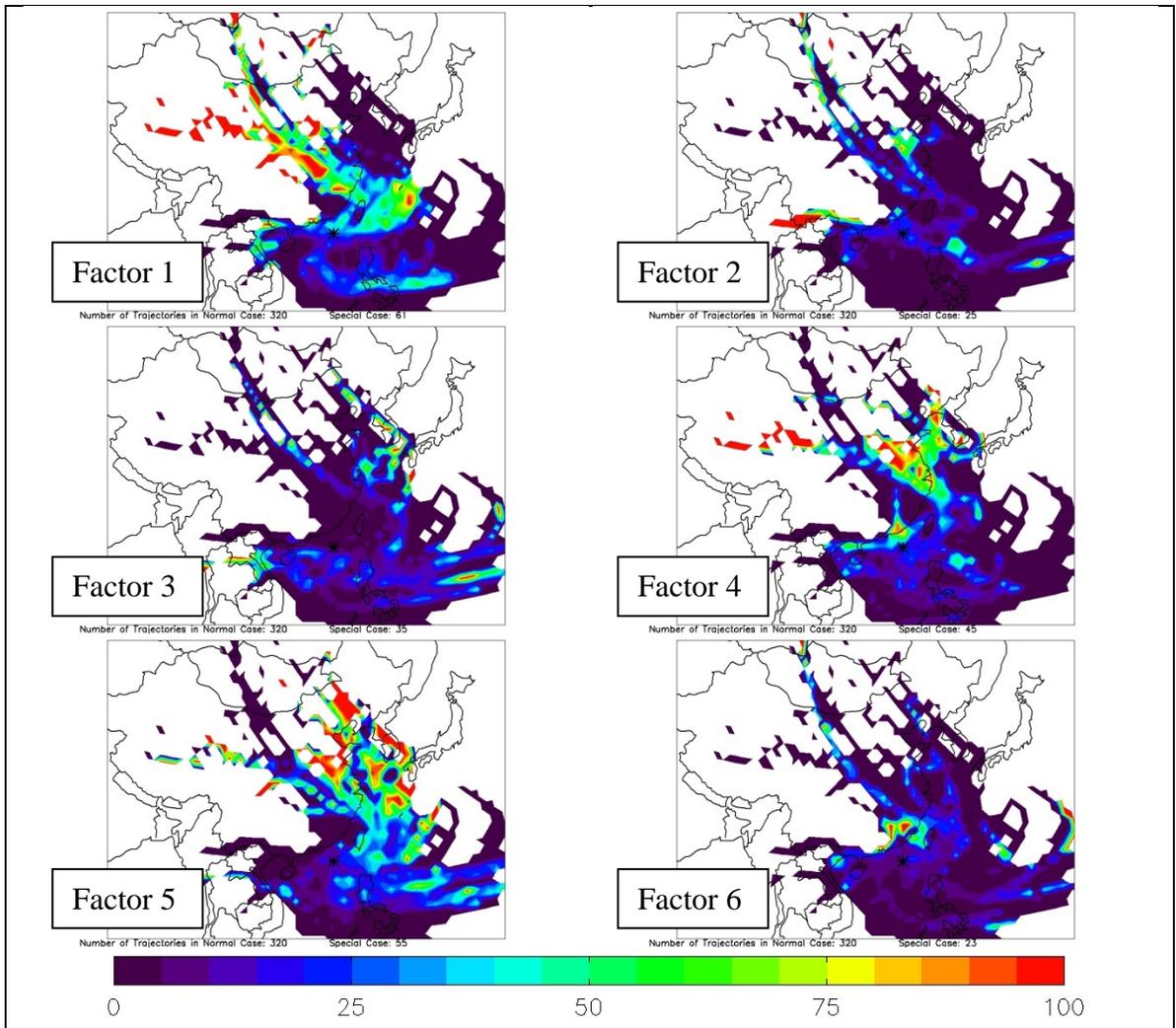


Figure 50 Dongsha Three Mode Special Case PSCF plots for 1000m Receptor Heights.

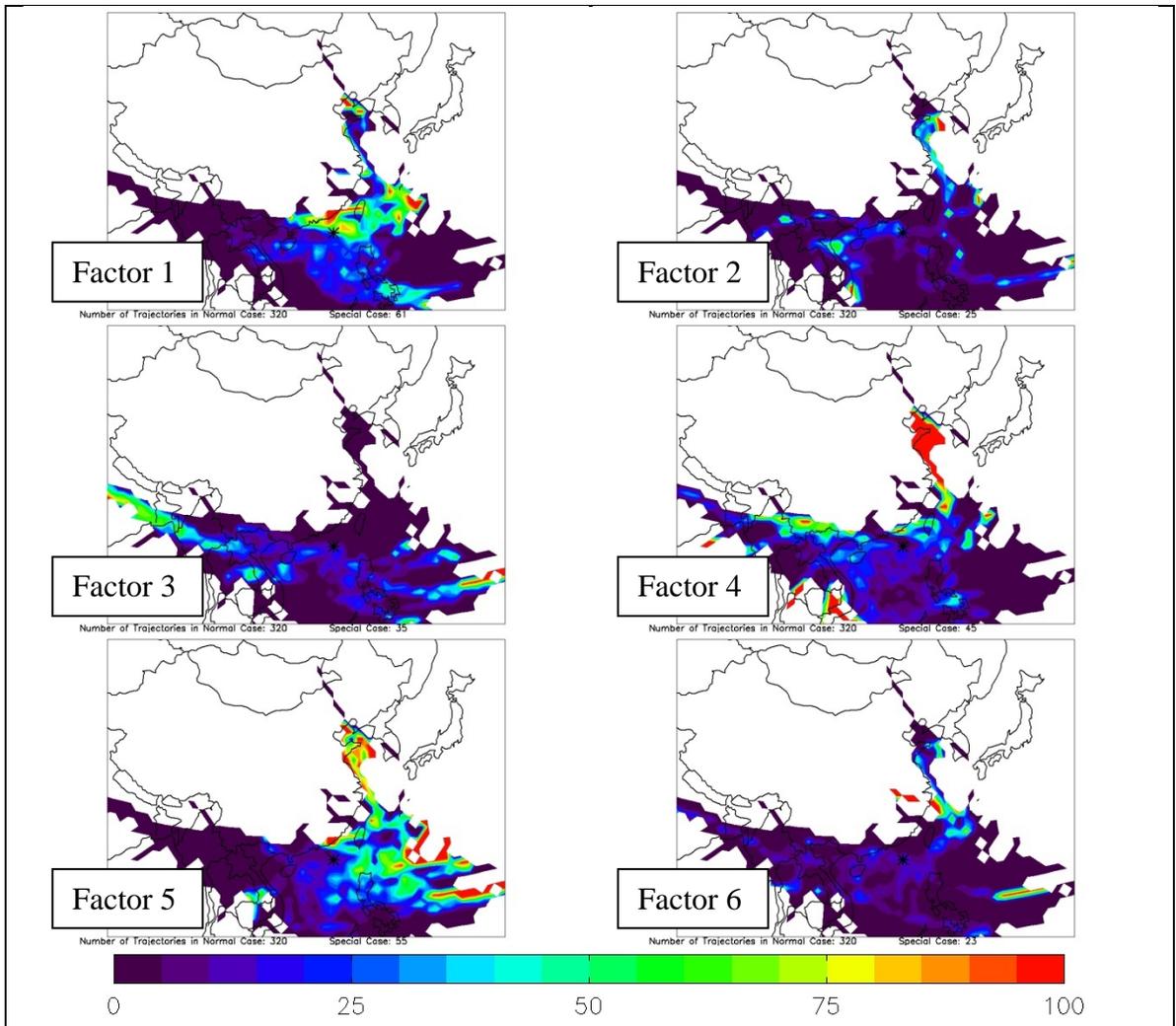


Figure 51 Dongsha Three Mode Special Case PSCF plots for 1500m Receptor Heights.