ESTIMATING THE DECISION-MAKING VALUE OF INFORMATION OBTAINED FROM SURFACE OBSERVATIONS

A thesis submitted to
the Faculty of Graduate and Postdoctoral Affairs
in Partial Fulfillment of the requirements for the degree

Master of Applied Science

by

Sahar Saeednooran

B.Sc.

Department of Civil and Environmental Engineering
Carleton University

Ottawa-Carleton Institute of Civil and Environmental Engineering

September 2014

Abstract

Regulatory air quality management, as well as scientific understanding of air quality processes and models, relies heavily on observation systems. While spaceborne observations have resulted in a new wealth of information, in-situ, surface observations continue to provide invaluable information to environmental managers and scientists. Surface observations are mainly used as a representation of the true state of the atmosphere, i.e., to provide concentrations of pollutants of concern. One of the ultimate uses for air quality observations is through collective decision metrics that go beyond single point measurements. Examples of these decision metrics include regional attainment or population health metrics.

In this study, spatial interpolation (i.e., Kriging method) is conducted to account for spatial patterns of ozone concentrations. Kriging Standard Deviations (SD) are calculated to quantify the spatially varying predicted ozone uncertainties. Furthermore, the adjoint of gas-phase CMAQ is employed to relate collective ozone decision metrics (e.g., short-term ozone mortality) to ozone concentrations. Finally, the estimated uncertainties in ozone concentrations are propagated through the adjoint model to assess the contribution of different sources of uncertainty to the overall uncertainty in mortality risk reduction benefits associated with improved air quality. New monitoring stations are recommended to add at the locations with large contributions so as to reduce the estimated health benefits uncertainty the most (i.e., maximize the uncertainty-based information of the health benefits). In addition, Shannon entropy is used to quantify the information value added by the observations at different locations across the U.S. and potential sites are prioritized based on the information gained.

Acknowledgements

There are a number of people I wish to thank for their help and contribution to this thesis. First of all, I would like to deeply thank my supervisor, Dr. Amir Hakami, for giving me the opportunity to pursue this research, and for the guidance and support he has given me in my master's program. His insight and contagious enthusiasm has been a source of inspiration through this work.

I would also like to acknowledge the kind support of the faculty, staff and students in the Civil and Environmental Engineering department at Carleton University. In particular, the members of the atmospheric research group (Dr. Morteza Mesbah, Matthew Russell, and Amanda Pappin) have been tremendously helpful in different aspects of my thesis.

I also owe much gratitude to my parents and my little sister. Mom, Dad, Samira you have provided support, encouragement and interest in my thesis work. Even though we are thousands of miles away, you were always there whenever I needed you. I would not be who am I today without you.

Finally, the most special thanks go to my best friend and colleague, my husband. Pedram without your unconditional support and love throughout this process, this thesis would have never been completed.

Table of Contents

List of Tables

List of Figures

List of Abbreviation

ADE	Atmospheric diffusion equation
AQS	Air Quality System
BCON	Boundary Condition Preprocessor
BF	Brute-Force
CMAQ	Community Multi-scale Air Quality model
COI	Cost of Illness
CTM	Chemical Transport Model
CWS	Canada Wide Standards
DALY	Disability-Adjusted Life Years
DDM	Decoupled Direct Method
DFS	Degrees of Freedom for Signal
DOF	Degrees of Freedom
FIM	Fisher Information Matrix
FORM	First- order Reliability Method
HDDM	Higher-order Decoupled Direct Method
ICON	Initial Condition Preprocessor
IDW	Inverse Distance Weighting
KPP	Kinetic pre-processor
LM	Linear Regression Model
MCS	Monte Carlo Simulation
NAAQS	National Ambient Air Quality Standards
NN	Nearest Neighborhood

OK	Ordinary Kriging
PDF	Probability Density Function
PM	Particulate Matter
QALY	Quality –Adjusted Life Years
SA	Sensitivity Analysis
SD	Standard Deviation
SK	Simple Kriging
SOI	Sphere of Influence
SORM	Second- Order Reliability Method
TLM	Tangent Linear Model
UK	Universal Kriging
US EPA	United States Environmental Protection Agency
VOC	Volatile Organic Compound
VSL	Value of a Statistical Life
WHO	World Health Organization
WTP	Willingness to Pay

List of Symbols

$\mathcal{P}(x)$	probability density function of the state x
$\mathcal{P}^{A}(x)$	posterior probability density
$\mathcal{P}^{B}(x)$	prior probability density
$\mathcal{H}$	Shannon entropy
$\mathcal{I}^{Shanon}$	Shannon information content
$\boldsymbol{A_0}$	posterior covariance matrix
$\boldsymbol{B}_0$	prior covariance matrix
n	number of error components in the state vector
$\mathcal{P}(y\|x)$	conditional probability of the observations
$\boldsymbol{Q}$	an orthogonal matrix
$\boldsymbol{\Sigma}$	a diagonal matrix
$\mathcal{I}^{\mathrm{DFS}}$	degrees of freedom for signal information
$\mathcal{F}$	Fisher information matrix
$\mathcal{I}^{\mathrm{FIM}}$	Fisher information content
$\hat{Z}(x)$	estimated value at the point x
$Z(x)$	observed value at the point x
λ_i	interpolation weight for each sample point
μ	mean value
ε'	random spatially correlated component
ε"	spatially uncorrelated random noise
h	distance between the points

γ	semi-variance
C_i	concentration of species *i*
$\boldsymbol{u}$	three-dimensional wind field
$\boldsymbol{K}$	turbulent diffusivity tensor
R_i	chemical reaction rate of species *i*
E_i	emission rate of species *i*
p_j	model parameter or input
S_{ij}	semi-normalized sensitivity of species i with respect to the emissions of species j
ϵ_j	scaling factor
$\boldsymbol{F}$	Jacobian matrix
δ_{ij}	Kronecker delta function
J	adjoint cost function
λ	adjoint variable
g	local cost function
φ	adjoint forcing term
$\boldsymbol{s}$	location of the points
σ_e^2	mean-squared error
$\boldsymbol{\Gamma}$	modeled (theoretical) semivariogram between all pairs of sample locations
$\boldsymbol{g}$	theoretical semivariogram values between each measured location and the prediction location

C_0	nugget
$C_0 + C_1$	sill
a	range
S	standard error of estimate
M_0	baseline non-accidental mortality rate
P	population
β	epidemiological concentration response factor
ΔM	change in mortality
V_{SL}	value of statistical life
Var	variance
Cov	covariance
U_{hourly}	hourly uncertainty
U_{local}	local uncertainty
$U_{overall}$	overall uncertainty

1 Chapter: Introduction

1.1 Overview

Ground-level ozone (O_3) is among the air pollutants of most concern. Numerous epidemiological studies have reported the association between O_3 and a number of serious short and long-term health effects (Bonn, 2003; Bell et al., 2004; Hubbel et al., 2005). In order to account for the adverse effects of ozone, a number of metrics are commonly used. Health-adjusted life year measures transform any type of mortality and morbidity into an equivalent number and quality of life years. The Air Quality Health Index (AQHI) also has been developed for estimating short-term health risk caused by degraded air quality. This metric is based on the combined impact of three pollutants (NO_2, O_3, and $PM_{2.5}$). Moreover, monetary measures express health state in a monetary value such as Willingness to Pay (WTP). WTP considers how much individuals would be willing to pay for a reduction in their risk of death or disease (Robinson, 2008). Health benefits of mortality risk reduction are of great significance and can be monetized by using Value per Statistical Life (VSL) as a metric of WTP.

Air quality data is required for identifying and tackling air quality problems. Thereby, ambient air pollutants such as ozone are continuously measured at monitoring sites. One of the main objectives of ozone monitors is to insure compliance with the existing ozone standards (e.g., National Ambient Air Quality Standards (NAAQS) in the U.S. and Canada-Wide Standards (CWS) in Canada). Since it is not possible to monitor O_3 levels at all locations and all times, selection of sites to give a reliable and realistic picture of ozone concentrations becomes an important issue and at the same time a difficult task.

Determining the location (configuration) and the number of monitoring stations is referred to as the network design problem. An increased number of monitoring sites will lead to a decrease in sampling error but will increase network cost. Thus, the design technique must be able to evaluate the trade-off between the economic risk due to inadequate information and the cost of a network capable of providing the required information. This evaluation can be viable in light of air quality monitoring objectives.

Some of the specific monitoring objectives can be: to quantify ambient air quality and its variation in space and time; to provide data for air pollution control regulations; to determine representative concentrations in areas of high population density; and to discover highest concentrations expected to occur in the area covered by the network. Based on these objectives, most of the available ozone monitors are situated within urban centers with higher ozone concentrations or areas of generally higher population density. In 2007, there was a total of 1198 state and local O_3 monitors providing data in the United States (U.S. EPA, 2014). Figure 1-1 shows that the total number of monitoring locations is a reasonably high number and most of the U.S. domain exhibits some monitoring. However, the spatial distribution of available data for analysis is far from evenly distributed across the domain and the emphasis of the network is on human population.

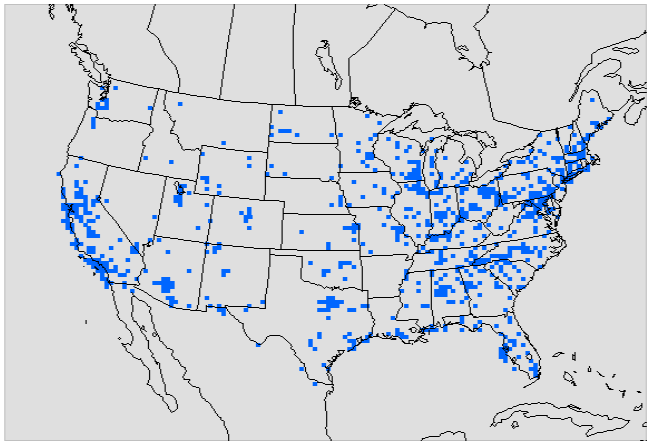


Figure 1-1: Spatial distribution of ozone monitoring sites within U.S. in 2007

It usually happens that an existing monitoring network needs to be modified to maximize its capacity to meet new demands. In other words, it may be desirable that the new network maximizes the amount of information it will provide concerning a certain policy objective. The information can be measured by the reduction of uncertainty known as uncertainty-based information metrics, including Shannon information, Degrees of freedom for signal, and Fisher information.

To reduce uncertainty maximally, variance reduction analysis can be adopted. The variance may be due to spatial interpolation estimation at non-monitored locations.

Obviously even by using the best available interpolation technique, there would be an estimation error. As a result, some studies proposed adding monitoring stations at the points of maximum interpolation errors to improve the regional variance of estimation (Rouhani, 1985). However, these studies ignored the overall effect of a new measurement on the accuracy of the whole field. Considering such an influence helps the selection of prospective locations which provide maximum information content.

1.2 Scope of the work

The primary goal of this dissertation is to determine the locations of new ozone monitors so as to achieve the best predictive capability and obtain the most informative set of observations. This goal is met through evaluating the reduction in uncertainty of our objective function (i.e., mortality risk reduction benefits associated with ozone reductions) realized by adding new monitors at different locations. In order to quantify the uncertainty of health benefits, uncertainty analysis is performed.

The uncertainty analysis to quantify the health benefits comprise three elements:

- *Quantification of uncertainty associated with the estimated ozone concentrations*. As ozone measurements are only available at discrete points and the decision-maker may require the concentrations at all grid cells, spatial interpolation is required which introduces errors into the input concentrations. Kriging standard deviations (SD) are used in this study to quantify the spatially varying interpolation errors as a measure of input uncertainties.

- *Computation of the effect of changes in ozone concentrations on mortality risk, i.e., sensitivity analysis*. The backward sensitivity analysis (adjoint of CMAQ) is employed to estimate the sensitivity of health benefits from reduced short-term mortality with respect to ozone reduction all over the model domain.

- *Evaluation of the uncertainty in mortality risk reduction benefits induced by the uncertainties in ozone concentrations, i.e., uncertainty propagation*. The perturbation method of uncertainty propagation is applied to assess how the health

metric variance will decline for a given reduction in the concentration variance of one or more grid cells.

The last goal of this research is to examine the existing ozone monitoring network to determine the extent of information gain resulting from the addition of one or more monitoring stations. To achieve this aim, Shannon Information theory, which measures the decrease in the (co-)variance of the error after adding new monitors at different locations (nodes), is applied.

1.3 Organization

The thesis is divided into the following five chapters:

Chapter 1 Focuses on the research background and the major objectives of this research.

Chapter 2 Overviews a number of different topics required to understand the scientific issues discussed later in the thesis. Ozone negative impacts and the need for ozone monitoring network design are explained briefly. An overview of uncertainty-based information metrics is followed with their applications in air monitoring network design. Spatial interpolation methods (deterministic and geostatistical interpolators) are also introduced. Finally, sensitivity analysis methods are reviewed and particular emphasis is given to describing adjoint (backward) sensitivity analysis.

Chapter 3 Describes the methods and tools used in this work. Kriging interpolation method is introduced to quantify the uncertainty of estimated ozone concentrations. Adjoint sensitivity analysis is also discussed as it is used for estimating the sensitivity of health benefits from reduced short-term mortality with respect to ozone reduction. An overview of an uncertainty propagation method (perturbation method) is given to discover the main contributors to uncertainty in

the health metric. Finally, Shannon information content is represented to measure the information content of monitoring networks.

Chapter 4 Presents the results of uncertainty analysis and Shannon entropy (i.e., suggested locations for new monitoring stations) and their discussion.

Chapter 5 Summarizes the findings and the open issues discussed in the previous chapters. A discussion of possible future areas of research concludes the chapter.

2 Chapter: Background

"Everything is related to everything else, but near things are more related than distant things."- Tobler

2.1 Ozone and Health metrics

This section provides a general overview of ground-level ozone (i.e., tropospheric ozone) which is considered as a serious threat for public health and the environment. Detrimental effects of elevated ozone concentrations on human health and quantification of these effects (i.e., health metrics) are discussed as well.

2.1.1 Overview of ground-level-ozone

Ground-level ozone (O_3) is among the air pollutants of most concern. It is a secondary pollutant mostly produced by photochemical reactions of Nitrogen Oxides (NO_x) and Volatile Organic Compounds (VOCs) in lower parts of the atmosphere. Hence, NO_x and VOCs are usually referred to as ozone precursors. Anthropogenic sources such as transportation (e.g., motor vehicles), combustion processes in energy production and industry, chemical solvents, biomass burning and agricultural practices are the major origins of ozone precursors (WHO, 2008).

The formation of ozone depends upon several environmental variables such as temperature, wind speed, wind direction, and mixing height (Cardelino and Chameides, 1995). Generally, high ozone levels occur in warm, sunny and stagnant atmospheric conditions. Thus, concentrations reach their peak in the heat of the afternoon and early evening. In the above-mentioned conditions, both emission of ozone precursors and their photochemical reactions increase which usually leads to higher ozone concentrations (Diem and Comrie, 2002).

Although NO_x can contribute to ozone formation through photochemical reactions, fresh NO_x plumes can also remove ozone by titration. Freshly emitted NO reacts with O_3 to form NO_2 on a time-scale of about 2 minutes. Since more than 90% of the emissions of oxides of nitrogen occur as NO, an amount of ozone equivalent to the flux of the emitted

NO will be converted to NO2. Lower NO concentrations are the primary causes for the lesser ozone destruction and higher concentrations at the rural locations, as in urban areas NO concentration is higher due to peak vehicular activity in the evening hours. Hence higher ozone in the rural locations may be the combined effect of slower titration by NO in the evening hours and the greater ozone formation due to precursor transport from the nearby polluted locations.

Studies have shown that ozone-pollution causes harm to humans, animals and plants. A brief summary of detrimental ozone effects on human health and vegetation is presented in the following section.

2.1.2 Impacts of tropospheric ozone

There are two ways in which air pollutants, such as ozone, may affect human health: acute (short-term) effects that are characterized by sudden and severe exposure; and chronic (long-term) effects that are characterized by prolonged or repeated exposures over a period that may ultimately lead to death.

Exposure studies have reported the association between ambient ozone and a number of serious health effects (Hubbel, 2005). Epidemiological studies have also linked several short and long-term negative impacts to ground-level ozone exposure (Bonn, 2003).

Numerous experimental studies have been carried out on the acute effects of ozone exposure on humans. Recent meta-analysis and time-series studies were consistent in showing clear association between short-term O_3 exposure and premature mortality (Anderson et al., 2004; Ito et al., 2005; Levy et al., 2005; Bell et al., 2004). Bell et al. (2004) found 0.52% increase in non-accidental mortality and 0.64% increase in cardiovascular and respiratory mortality for a 10 ppb increase in the daily average ozone concentration. Ito et al. (2005) conducted a meta-analysis of single-city studies and estimated 0.39% per 10-ppb increase in 1-hour daily maximum ozone level. Levy et al. (2005) also performed an empirical Bayes meta-regression to assess the health effects of ozone on premature mortality. The results of meta-regression suggested a considerable

0.41% increase in daily mortality per 10-ppb increase in 1-h maximum ozone concentrations.

In addition to short-term effects, ozone has been suspected of playing a role in long-term health effects. Chronic effects are considered to be the result of cumulative effects caused by repeated exposure. Few epidemiological studies have been carried out on reduced lung function, increased frequency of respiratory symptoms, development of asthma and total mortality (e.g., Jerrett et al., 2009). The epidemiologic evidence is inconclusive with regard to whether long-term exposure is related to total mortality or other chronic health effects. However, the repeated cycles of damage, inflammation, and repair in humans and the morphological findings from the animal toxicological studies suggest that repeated short-term exposures should be avoided until more is known about the effects of long-term ozone exposure (McConnell et al., 2002; U.S. EPA 2013).

Epidemiological studies of air pollution in Canada have also presented several acute and chronic detrimental effects of ozone (e.g., Burnett et al., 1997; Judek et al., (2005); Farhat et al., 2013). Time-series study by Farhat et al. (2013) in 12 Canadian cities investigated the association between ozone and mortality. They reported that 0.56% to 2.47% increase in mortality can be observed for a 10 ppb increase in 1-hr daily maximum ozone level. A significant association between tropospheric ozone and respiratory diseases was found in 16 cities across Canada by Burnett et al. (1997).

Elevated ozone exposure can also have adverse effects on plants and ecosystems. It can affect crop yield and cause reductions in growth and biomass of forest tree species in laboratory and glasshouse studies (Heck et al., 1982; Chappelka and Samuelson 1998)

2.1.3 Health metrics

Common health threats related with ground-level ozone were discussed in the previous section. In order to account for different health effects (morbidity and mortality) due to environmental impacts, a number of metrics have been proposed by economists and

social scientists. These metrics can help policy development and monitor progress over time in reaching goals.

Health-adjusted life year measures are indications of health metrics that transform any type of mortality and morbidity into an equivalent number and quality of life years. The most frequently used measures are Quality-Adjusted Life Years (QALYs) and Disability-Adjusted Life Years (DALYs), which are often applied in policy making.

The Air Quality Health Index (AQHI) also has been developed in Canada for estimating short-term health risk caused by degraded air quality. This evaluation metric was derived based on the combined impact of three pollutants (NO_2, O_3, and $PM_{2.5}$). It differs from the older Air Quality Index (AQI), which determined the health risk based on the highest value of only one pollutant. The AQHI is currently used by several provinces across Canada and presented on the Weather Network, Environment Canada weather website and Health Canada website. This index is scaled to range from 1 to 10, and increases as health risk increases (Environment Canada, 2012).

Moreover, there are metrics which involve economic valuation by placing a value on the expected health outcomes. These metrics are used to translate changes in health risks into monetary values such as Cost of Illness (COI) or Willingness to Pay (WTP) (Robinson, 2008). The COI approach focuses only on expenditures and on human capital costs, while WTP is a more comprehensive measure, considering how much individuals would be willing to pay to secure beneficial change in risk. In cost-benefit analyses, where the social costs of alternative air pollution control strategies are compared to the value of the associated health risk reductions, WTP estimates are more applicable. Decision-makers can use the results of these analyses to determine the most suitable policy (Zhang, 2002; Robinson, 2008).

A particularly important consideration is the value assigned to increased mortality risks. This value typically accounts for a substantial fraction of the economic costs associated with air pollution impacts on human health (Hubbel et al., 2005). Mortality risk valuation studies usually identify the amount an individual would pay for a small reduction in the

risk of death (i.e., willingness to pay values). In this study, health benefits of mortality risk reduction are monetized by using the Value of a Statistical Life (VSL) as a metric of WTP.

2.2 Ozone monitoring

Ambient air pollutants are continuously measured at monitoring sites for determining compliance with air quality standards. United States Environmental Protection Agency EPA (US EPA) has set National Ambient Air Quality Standards (NAAQS) for six principal pollutants, which are called "criteria" pollutants. Of the six pollutants, ground-level ozone and Particulate Matter (PM) are the most widespread health threats.

Several guidelines and standards exist for ozone concentration. According to the latest revision, the standard of ozone concentration was set to 75 ppb based on the annual fourth-highest daily maximum 8-hr average concentration, averaged over 3 years, by US EPA. The regulations are periodically reviewed to ensure that they are sufficiently stringent to prevent short and long-term health effects (CRS, 2010). Similarly, Environment Canada and Health Canada defined air quality standards for fine particulate matter and ground-level ozone. Canada-Wide Standards (CWS) have established the maximum allowable ozone level as 65 ppb (Annual fourth-highest daily maximum 8-hr concentration, averaged over 3 years). Canadian Ambient Air Quality Standards also provide more stringent long-term target in the range of 62-63 ppb for ground-level ozone concentrations in Canada (Environment Canada, 2000).

The main objective of ozone monitors is to insure compliance with standards. Monitoring data is applied for developing strategies of controlling ozone levels and as an input in risk assessment models (U.S. EPA. 1999). Although monitoring sites provide valuable information at their location, there is always a gap in the understanding of air quality in non-monitored areas due to widely variation of emission rates, meteorology, topography, and other local conditions. Thus, monitoring stations have their spatial representativeness.

For reliable prediction of ozone concentrations at non-monitored locations, number of monitoring sites and the appropriateness of their locations should be evaluated. Determining the optimum number as well as optimum configuration of ozone monitoring sites is referred to as the ozone network design problem. High expenses of continuous ozone measurements accentuate the role of optimal network design. There should be a balance between the economic risk due to inadequate information and the cost of a network capable of providing the required information (Singh, 2013). It has been reported that an urban network of 10 well-located O_3 monitoring stations can provide estimates of design values that differ little from networks having 25-32 stations (Ludwig et al., 1983).

2.2.1 Monitoring network design methods

In the present, most of the ozone (O_3) data available in North America has been collected at monitoring stations in and around heavily populated or industrialized areas. Several methodologies have been suggested for designing monitoring networks in the literature. Seinfeld (1972) ascertained the locations of stations based on the sensitivity of concentration measurements to changes in source emissions. Noll and Mitsutomi (1983) presented a network design procedure that assigns stations to high-dosage locations, without redundancy. They also identified the representative area of the assigned stations. Langstaff et al. (1986) defined the locations of the monitoring stations based on the human exposure patterns. To eliminate redundant monitoring sites, they proposed Sphere of Influence (SOI) based on the correlation between air quality at the stations and nearby points. Baldauf et al. (2002) established a monitoring network at locations according to maximum anticipated concentrations and highest population densities. In the recent work by Kanaroglou et al. (2004), a location-allocation approach was used to place monitors in areas with high spatial variability in traffic-related pollution and high population densities.

Some methods adopt variance reduction analysis for optimizing a monitoring network. The variance may be due to spatial interpolation estimation at non-monitored locations. A number of studies added monitoring stations at the points of maximum interpolation errors, while the others established the locations of monitors by minimizing a function of

the interpolation errors based on the objective of the monitoring network (e.g., Buell, 1975). Kriging variances can also be used in the design and optimization of monitoring networks (Cressie, 1993). The application of spatial interpolation methods in monitoring network design is discussed in details in section 2.4.1.2.2.

2.3 Uncertainty and Information

2.3.1 Background

Any model is an approximation of a complex real system. Engineers and scientists are always interested in understanding the performance of the models to make appropriate decisions. The behavior of the system can only be assessed according to available information, thereby involving some extent of uncertainty. There is a strong relationship between the concepts of uncertainty and information. Information deficiency may be related to missing, incomplete, imprecise, or vague knowledge which leads to uncertainty. Various information deficiencies determine the associated types of uncertainty.

There is often a need to quantify the amount of information provided by a set of observations. Information will increase as a result of adding new monitors and the gained information can be measured by the amount of uncertainty reduction. This indicates the difference between a priori (before the measurement) and a posteriori (after the measurement) uncertainty. This concept of information is usually referred to as uncertainty-based information as illustrated in Figure 2-1 (Klir, 2006).

If $\mathcal{P}^B(x)$ is the prior probability density and $\mathcal{P}^A(x)$ is the posterior probability density, the information content of the observations is measured by the decrease in uncertainty from $\mathcal{P}^B$ to $\mathcal{P}^A$. To quantify the uncertainty reduction, several information metrics are proposed. The metrics are distinguished from one another by the mathematical representation employed and the type of uncertainty involved. Three main types of information theoretic metrics are reviewed below (Klir, 2006; Sandu et al., 2012).

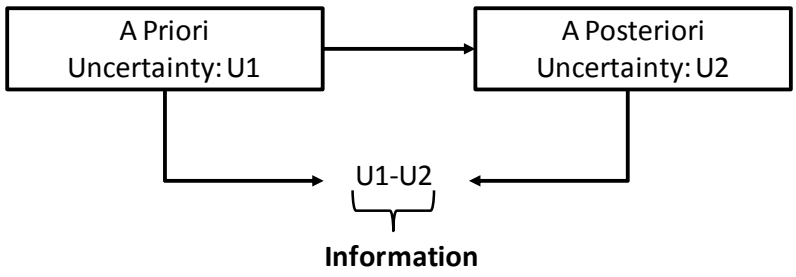


Figure 2-1: Uncertainty-based information (Klir, 2006)

2.3.2 Shannon information content

The question of how to measure the amount of uncertainty (and the associated information) in classical probability theory was first addressed by Shannon (1948). He introduced entropy as a meaningful way to measure the amount of uncertainty. Shannon entropy is the most widely used concept in information theory, which derives its name by analogy to entropy in thermodynamics and statistical mechanics. The entropy associated with a continuous probability density function (pdf) is defined as follows:

$$\mathcal{H}(\mathcal{P}) = -\int_{R^n} \mathcal{P}(x)\, log_b(\mathcal{P}(x))\, dx \qquad (2\text{-}1)$$

where b is the base of the logarithm. Common values of b are 2, Euler's number e, and 10 and the unit of entropy is bit for b=2, nat for b=e, and dit (digit) for b=10.

$\mathcal{P}(x)$ describes our knowledge of the system, i.e., it is the pdf of the state x. The entropy represents the number of internal states of the system being measured, consistent with the measurement. Shannon information content is given by the change in the entropy of the state before and after making the measurement. It can, then, be determined as:

$$\mathcal{I}^{Shannon} = \mathcal{H}(\mathcal{P}^B) - \mathcal{H}(\mathcal{P}^A) \qquad (2\text{-}2)$$

With the Gaussian assumption, the entropy of the prior pdf with n elements is given by relation (Rodgers, 2000):

$$\mathcal{H}(\mathcal{P}^B) = n \ln\left(\sqrt{2\pi e}\right) + \frac{1}{2} \ln \det(\boldsymbol{B}_0) \qquad (2\text{-}3)$$

where $\boldsymbol{B}_0$ is the prior covariance matrix and n is the number of error components in the state vector. By using the Bayes rule, the entropy of the posterior pdf can be written as:

$$\mathcal{H}(\mathcal{P}^A) = \int [\ln \mathcal{P}^B(x) + \ln \mathcal{P}(y|x) - \ln \mathcal{P}(y)]\, \mathcal{P}^A(x)\, dx \tag{2-4}$$

Here $\mathcal{P}(y)$ is the pdf of the observations and $\mathcal{P}(y|x)$ is the conditional probability of the observations.

Under the assumption that both the prior and the posterior pdfs are Gaussian, the Shannon information content of the observations would be:

$$\mathcal{J}^{Shanon} = \frac{1}{2}\ln\det(\boldsymbol{B_0}) - \frac{1}{2}\ln\det(\boldsymbol{A_0}) \tag{2-5}$$

$$= \frac{1}{2}\ln\det(\boldsymbol{B_0 A_0^{-1}}) = \frac{1}{2}\ln\det\left(\boldsymbol{A_0^{-1/2} B_0 A_0^{-1/2}}\right)$$

Matrices $\boldsymbol{A_0}$ and $\boldsymbol{B_0}$ are the posterior and prior covariance matrices, respectively. The greater the reduction in uncertainty is, the larger the $\mathcal{J}^{Shanon}$ will be.

2.3.3 Degrees of freedom for signal

The Degrees of Freedom for Signal (DFS) indicates the number of independent pieces of information available from measurements. Obviously, observation groupings with a high number of degrees of freedom are of interest (Sandu et al., 2012). To evaluate the number of DFS, a linear algebra approach can be used.

Consider an orthogonal matrix $\boldsymbol{Q}$ whose columns are eigenvectors of the symmetric matrix $\boldsymbol{B_0^{-1/2} A_0 B_0^{-1/2}}$:

$$\boldsymbol{Q^T\left(B_0^{-1/2} A_0 B_0^{-1/2}\right)Q = \Sigma} \tag{2-6}$$

with $\boldsymbol{\Sigma}$ a diagonal matrix. The matrix $\boldsymbol{L = B_0^{-1/2} Q}$ has the property that it transforms simultaneously the prior and the posterior covariance matrices ($\boldsymbol{B_0, A_0}$) to diagonal forms (Fisher, 2003) when it is symmetrically applied:

$$\boldsymbol{L^T B_0 L = I_{n\times n}}, \quad \boldsymbol{L^T A_0 L = \Sigma} \tag{2-7}$$

The diagonal elements of the transformed prior covariance matrix are equal to unity and each corresponds to an individual degree of freedom. The eigenvalues of the transformed matrix $\boldsymbol{\Sigma}$, on the other hand, can be interpreted as the relative reduction in variance in each of the *n* independent directions corresponding to the *n* components of error in the state vector. The degrees of freedom for signal measures the total reduction in variance and is defined as:

$$\mathcal{J}^{\mathrm{DFS}} = \mathrm{trace}(\mathbf{I}_{\mathrm{n\times n}} - \boldsymbol{\Sigma}) = n - trace\left(\boldsymbol{B_0^{-1/2} A_0 B_0^{-1/2}}\right) = \mathrm{n} - \mathrm{trace}(\boldsymbol{B_0^{-1} A_0}) \tag{2-8}$$

2.3.4 Fisher information matrix

The Fisher Information Matrix (FIM) (Fisher, 1922) associated with the probability density function $\mathcal{P}(x)$ is defined as:

$$\mathcal{F}(\mathcal{P}) = \int_{\mathbb{R}^n} \left[\frac{\partial(-\ln \mathcal{P}(x))}{\partial x}\right] \left[\frac{\partial(-\ln \mathcal{P}(x))}{\partial x}\right]^T \mathcal{P}(x)\, dx \quad \in \mathbb{R}^{n\times n} \tag{2-9}$$

The trace of the FIM offers a measure of the total level of uncertainty associated with the distribution. Under the assumption that the prior errors are normally distributed, the Fisher information matrix of $\mathcal{P}^B(x)$is just the inverse of the prior covariance matrix:

$$\mathcal{F}(\mathcal{P}^B) = \int_{\mathbb{R}^n} [\nabla \mathcal{J}^{\mathrm{B}}(x_0)]\, [\nabla \mathcal{J}^{\mathrm{B}}(x_0)]^T \mathcal{P}^B(x_0)\, dx_0 = \boldsymbol{B_0^{-1}} \tag{2-10}$$

Similarly, assuming that the posterior pdf is Gaussian, the posterior Fisher information matrix is:

$$\mathcal{F}(\mathcal{P}^A) = \int_{\mathbb{R}^n} [\nabla\, \mathcal{J}(x_0)]\, [\nabla\, \mathcal{J}(x_0)]^T \mathcal{P}^A(x_0)\, dx_0 = \boldsymbol{A_0^{-1}} \tag{2-11}$$

The information content of the observations can be measured as the trace of the posterior FIM minus the trace of the prior FIM. In the Gaussian case this reduces to the trace of difference between the posterior and prior covariance matrices:

$$\mathcal{J}^{\mathrm{FIM}} = \mathrm{trace}\big(\mathcal{F}(\mathcal{P}^A)\big) - trace(\mathcal{F}(\mathcal{P}^B)) = \mathrm{trace}(\boldsymbol{A_0^{-1}} - \boldsymbol{B_0^{-1}}) \quad (2\text{-}12)$$

2.3.5 Applications of information theory

Information theory started in electrical engineering and has been applied to diverse areas as network analysis, financial mathematics and mathematical statistics (Manning and Petit, 2003). Information theory has been used in atmospheric sciences for uncertainty studies, instrument development, and data selection. In the development of remote-sounding instruments, some information metrics like the entropy reduction and the degrees of freedom for signal are commonly employed (Rodgers, 2000). Applications of information theory in air quality have been relatively limited in number, although promising (Le and Zidek, 2006). Fleming (2007) applied basic principles from communication theory to assess seasonal and spatial patterns in the variability or uncertainty in ground-level ozone concentrations.

Three basic principles of uncertainty were developed to guide the use of uncertainty measures in different situations. Due to the connection between uncertainty and uncertainty-based information, these principles can also be interpreted as principles of information. The principles are: minimum uncertainty, maximum uncertainty, and uncertainty invariance (Klir, 2006). The principle of maximum uncertainty suggests how to utilize all the variable information and acknowledge our ignorance at the same time. The principle of minimum uncertainty, on the other hand, facilitates the selection of most valuable alternatives from available solutions of a problem with the least information loss (Harmanec, 1999). The last principle, the principle of uncertainty invariance, is recently used to guide meaningful transformations between various theories of uncertainty. This principle postulates that the amount of uncertainty should be preserved in each transformation of uncertainty from one mathematical framework to another. In this study, the location of monitoring stations is selected based on the principle of minimum uncertainty, also called maximum information.

Selected approaches for monitoring network design were noted in section 2.2.1. In the following section, the applications of information theory in monitoring network design will be discussed.

2.3.5.1 Applications in air monitoring network design

Prior applications of information theory have focused largely on the mathematically formalized design of optimal monitoring network configurations (Le and Zidek, 2006). Information theory can be applied to locate a given number of monitors in a network so that the information content in the collected data is maximized or the uncertainty is minimized (Caselton and Zidek, 1984; Silva and Quiroz, 2003). In this application, information theory can also be used to evaluate the effectiveness of the existing monitoring network of stations or optimize the information gained by locating an additional number of new monitors. Pickett and Whiting (1981) used a statistical measure of information content to assess the effectiveness of a particular monitoring network configuration in Canada.

Information metrics can be used as an objective function in different network design studies. Husain and Khan (1983) proposed a Statistical technique based on Fisher's information metric to design an optimal air monitoring network. Shannon's entropy was first applied by Caselton and Zidek (1984) for designing networks as a measure of the information content that reflects the reduction in uncertainty when a random variable is observed. Caselton and Zidek (1992) used it to rank monitoring sites for possible elimination, an idea extended by Wu & Zidek (1992). Recently, Ainslie et al. (2009) used the entropy-based approach of Le & Zidek (2006) to redesign a monitoring network in Vancouver (Canada) using hourly ozone concentration.

2.4 Spatial interpolation

When data are collected at discrete locations, spatial interpolation is commonly applied to create a continuous observation surface. Spatially continuous data plays a significant role in planning, risk assessment and decision making in environmental management. The goal of interpolation is to predict concentrations at unmeasured locations of interest based on available observations at surrounding area.

2.4.1 Spatial interpolation methods

Numerous methods have been developed for spatial interpolation in various disciplines. They fall into three main categories: deterministic methods, geostatistical methods, and combined methods. All spatial interpolation methods have the same basic mathematical formulation. The value (e.g., air pollutant concentration) at un-sampled location of interest can be represented as weighted averages of sampled data. The interpolation relationship is:

$$\hat{z}(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i) \quad \text{and} \quad \sum_{i=1}^{n} \lambda_i = 1 \tag{2-13}$$

where $\hat{z}$ is the estimated value at the point x_0, z is the observed value at the sample point x_i, λ_i is the interpolation weigh for each sample point, and n represents the number of sample points used for the estimation (Webster and Oliver, 2001). The difference among interpolation methods is mainly due to the means of allocating weights to sample data.

2.4.1.1 Deterministic interpolators

Deterministic spatial models estimate the dependent variable by empirical parameters. These models use mathematical formulas or other relationships to interpolate values and no estimation of error is available. The following categories provide a brief description of deterministic interpolation methods:

- *Nearest Neighborhood (NN)* utilizes no information about the system other than measured data points. It assigns an average value of the cell and its surrounding points for each grid cell. It means that all points in an area are set equal to this average value. NN is also known as Thiessen or Voronoï polygons. All locations within each polygon are assigned the same value (Webster and Oliver, 2001).

- *Inverse Distance Weighting (IDW)* estimates the value of a certain grid cell by a linear combination of the surrounding sampled points weighted by an inverse function of the distance, i.e., the closer data points to the location of interest, the more weight they receive in the averaging formula. IDW has some limitations

owning to the fact that spatial relationship between two locations is not simply a function of distance. However, it is one of the most widely used interpolation schemes (Lu and Wong, 2008).

- *Linear regression Model (LM)* seeks a functional relationship between the predicted variable and one or more independent variables (e.g., geographical coordinates, traffic, population, point source, and others) which are easy to measure (Burrough and McDonnell, 1998). Linear regression models are deterministic, but by considering some statistical assumptions about the probability distribution of the predicted variable, the method turns to stochastic. In that case, the standard error can be calculated as well (Sluiter, 2008).

- *Polynomial functions (splines)* are methods that fit mathematical functions to the observed values. The order of polynomials can change from first-order polynomial (linear) to higher-order polynomials. More details on the application of polynomial interpolation in air quality data are provided in the study by Junninen et al. (2004).

2.4.1.2 Geostatistical interpolators

Deterministic interpolation methods were discussed in the previous section. Those methods are directly based on the surrounding measured values or on specified mathematical formulas. A second group of interpolation methods, geostatistical methods, are based on statistical methods that include autocorrelation (statistical relationships among the measured points). Not only do these techniques have the capability of producing a prediction surface, but they can also provide some measure of the uncertainty and accuracy of the predictions.

Kriging, the most well-known method in geostatistics, was first introduced by D.G. Krige, a mining engineer, in 1951 for small amount of data. That method was formalized and developed later by Matheron (1962) and now is commonly known as kriging, with great power in spatial statistics.

Similar to IDW method, kriging calculates weights for measured points in deriving predicted values for unmeasured locations. However, the weights are based not only on distance between points, but also the correlation between sample points as a function of distance. Kriging has the advantage of providing unbiased estimates of values at unmeasured locations with minimum estimated variance (Diem and Comrie, 2002).

Burrough and McDonnell (1998) have reported that kriging is the best interpolation technique when data is sparse. It is a relatively fast interpolator that can be exact or smoothed depending on the applied method. The following section reviews three main types of kriging including simple, ordinary, and universal kriging.

2.4.1.2.1 Major kriging methods

Kriging recognizes that the spatial variation of any continuous attribute is too complex to be modeled by a simple, smooth mathematical function. Therefore, it is modeled as a stochastic surface or random field. Regionalized variable theory assumes that the spatial variation of any variable can be expressed as the sum of three major components:

$$Z(x) = \mu(x) + \varepsilon'(x) + \varepsilon'' \quad (2\text{-}14)$$

$\mu(x)$ is a structural component; $\varepsilon'(x)$ is a random, but spatially auto-correlated component, and ε'' is a spatially uncorrelated random noise which is related to measurement errors (Burrough and McDonnell, 1998).

The first step is to decide on a suitable function for $\mu(x)$. In the simplest case this can be thought of as a flat surface with no trend. The mean value of $\mu(x)$ is the mean value within the sample area. This means that the expected difference in the values for two points x and $x + h$ (where h is the difference between the points) is zero. i.e.,:

$$\mathbb{E}[Z(x) - Z(x + h)] = 0 \quad (2\text{-}15)$$

It is also assumed that the variance of the differences is a function of the distance between the points. In other words, they are spatially auto-correlated and the near points are more likely to have similar values:

$$\mathbb{E}[\{Z(x) - Z(x+h)\}^2] \; = \; \mathbb{E}[\{\varepsilon'(x) - \varepsilon'(x+h)\}^2 \,] \; = 2\gamma(h) \tag{2-16}$$

where $\gamma(h)$ is known as the semivariance. Under these two assumptions (i.e., stationarity of differences and stationarity in the variance of differences), the model can be expressed as:

$$Z(x) = \mu(x) + \gamma(h) + \varepsilon'' \tag{2-17}$$

Depending on the assumption about the mean value or systematic trend, different types of kriging can be applied. Classical methods of kriging fall into three categories:

1) Simple Kriging (SK), which is often referred to as "kriging with known mean". SK assumes second-order stationarity that is constant mean, variance and covariance over the domain of interest. In Figure 2-2, the observed data is given by the solid circles and the known constant mean -the solid line- is μ. For simplicity the measurement error is not represented in the following figures. The assumption of simple kriging is unrealistic for most problems and hence this method is rarely used (Wackernagel, 2003)

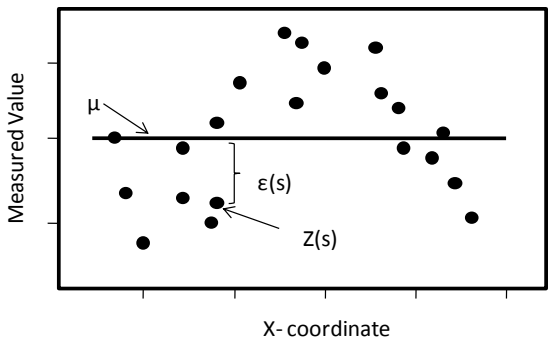


Figure 2-2: Simple Kriging concept

2) Ordinary Kriging (OK) has more relaxed assumptions than Simple Kriging. OK assumes second order stationarity, with an unknown mean which is determined during

the interpolation (De Smith et al., 2007). The constant unknown mean is shown in dotted line, μ in Figure 2-3. $\varepsilon(s)$ is the spatially correlated component which is the main focus of ordinary kriging.

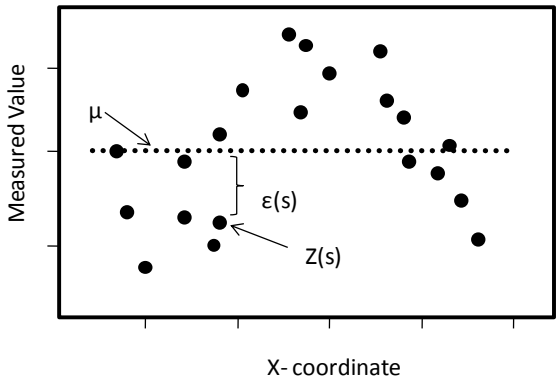


Figure 2-3: Ordinary Kriging concept

3) Universal Kriging (UK) is an extension of OK. In practice, environmental fields often exhibit non-constant mean values. This fact has led to the development of the universal kriging method (Matheron, 1969) which assumes that the mean follows a particular polynomial trend model. Second order polynomial is shown in Figure 2-4 as a trend - long dashed line - which is $\mu(s)$. By subtracting the second-order polynomial from the data, the errors, $\varepsilon(s)$, are obtained which are assumed to be random. The mean of all $\varepsilon(s)$ is 0.

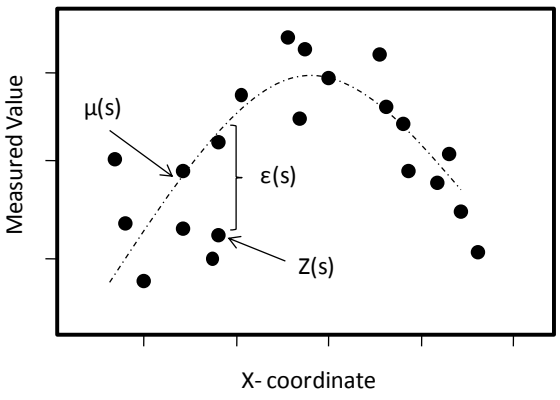


Figure 2-4: Universal Kriging concept

2.4.1.2.2 Kriging applications

The spatial interpolation methods like kriging have been applied widely in environmental sciences, including soil, groundwater, and acid precipitation modeling and air quality data (Cameron and Hunter 2002; Hengl et al., 2004; Tan et al., 2005). Studies of environmental health utilize the estimated concentrations over spatial fields to assess the health impacts of environmental pollution (Le and Zidek, 2006). For instance, in the study by Mulholland et al. in 1998, a kriging approach was used to interpolate ozone levels in the Atlanta region to investigate the relationship between ozone concentrations and increases in pediatric asthma rates.

Kriging has also been employed in the design of air pollution monitoring networks. The main goal is to locate the monitors of a network so that the kriging variance is minimized. Trujillo-Ventura and Ellis (1990) presented a multi-objective air pollution monitoring network design which uses multiple metrics to assess the spatial coverage of the network. Minimizing the spatial covariance of the interpolation method (i.e., simple kriging) is among their objective functions. There are other objective functions used in the literature based on the variance-covariance of the kriging estimation (e.g., minimizing the determinant of the matrix of estimation variance-covariances (Sacks and Schiller, 1988) and evaluating the trace of the matrix of kriging estimation variance-covariances (Fedorov and Mueller, 1989). Haas (1992) also selected the design criteria based on minimizing mean relative error and maximizing the information content. He optimized the continental-scale monitoring network by minimizing the kriging variance. Similarly, a study by Holland et al. (1998) applied kriging and other methods to evaluate the CASTNet network in the eastern U.S.

2.5 Sensitivity analysis

Atmospheric Chemical Transport Models (CTMs) are being extensively used for the study of air pollution and environmental policy making (Sandu et al., 2008). The response of these model predictions to system parameters or emission controls provides valuable information for the strategy design to improve air quality. Such information can

be obtained by sensitivity analysis (Nowak et al., 1998). Sensitivity Analysis (SA) is used to calculate the changes in an output variable at the receptor (e.g., ambient quality) with respect to small perturbations in an input variable (e.g., emission). Sensitivity information obtained from atmospheric models can be used in numerous applications such as design of air pollution control strategies, inverse modeling, air quality forecasting, and data assimilation (Hakami et al., 2006). SA is also applied to assess uncertainties in air quality models, i.e., how given model depends on the information fed into it (Saltelli et al., 2000).

The governing equation for atmospheric CTMs is the Advection-Diffusion Equation (ADE) (Seinfeld and Pandis, 2006):

$$\frac{\partial C_i}{\partial t} = -\nabla.(\boldsymbol{u}C_i) + \nabla.(\boldsymbol{K}\nabla C_i) + R_i + E_i \qquad (2\text{-}18)$$

where C_i is the mixing ratio of species i and varies spatially and temporally, $\boldsymbol{u}$ is the vector of wind field, $\boldsymbol{K}$ is the diffusivity tensor, R_i is the chemical reaction rate for species$i$, E_i is the emission rate of species i. Equation 2-18 is solved subject to specified initial and boundary conditions explained elsewhere (Seinfeld & Pandis, 2006).

Two kinds of sensitivity techniques exist in the literature, namely global and local sensitivity analysis. Global sensitivity methods take into account all variation ranges of the input parameters. Nonetheless, local methods concentrate on the local impact of input parameters on the model and do not attempt to fully explore the input space. It is based on computing the gradient of outputs with respect to the inputs around a given value. Thus, local SA is only valid within small neighborhoods of the nominal values of the model parameters (Zlatev, 2011). Subject of our study in this work is the local sensitivity analysis.

The first-order local sensitivity coefficients are defined as (Yang et al.,1997):

$$S_{ij} = \frac{\partial C_i(t)}{\partial p_j} \qquad (2\text{-}19)$$

where p_j is a model parameter or input, and C_i is the concentration of i^{th} species. The term local refers to the fact that these sensitivities describe the system around a given set of values for the parameters p. Due to the spatial and temporal variability in the magnitudes and units of different parameters of concern, semi-normalized sensitivity coefficients are defined by normalizing the absolute coefficients:

$$S_{ij} = P_j \frac{\partial C_i(t)}{\partial p_j} = P_j \frac{\partial C_i(t)}{\partial(\epsilon_j P_j)} = \frac{\partial C_i(t)}{\partial \epsilon_j} \quad (2\text{-}20)$$

where P_j represents an unperturbed field, and ϵ_j is a scaling variable with a nominal value of 1. Therefore, the semi-normalized sensitivity coefficient S_{ij} can be calculated by the partial derivative of a species concentration, to the scaling variable. Numerous local sensitivity techniques have been developed to compute sensitivity coefficients efficiently, including finite-difference (i.e., brute force) schemes, direct differentiation or adjoint differentiation methods (Cacuci et al., 2005; Saltelli et al., 2004).

2.5.1 The Brute-Force (BF) Method

Brute Force (BF) is a traditional technique in sensitivity analysis which is applied for calculating first order sensitivity coefficients. In this technique, each parameter is perturbed by a small amount and the model response is computed. Based on the finite difference approximation (central scheme), the sensitivity coefficients of BF method are simplified as:

$$S_{ij} \cong \frac{C_i|\,\epsilon_j+\Delta\epsilon_j - C_i|\,\epsilon_j-\Delta\epsilon_j}{2\Delta\epsilon_j} \quad (2\text{-}21)$$

Here, S_{ij} measures the ratio between the effect (absolute variation of output ΔC_i) and the cause (absolute variation of input $\Delta\epsilon_j$).

Simulating one-at-a-time perturbation of each sensitivity parameter becomes less viable and inefficient when a large number of sensitivity coefficients are desired. Furthermore, BF method faces numerical noise for small perturbations and this method is unrealistic

for large perturbations. Despite all of these shortcomings, BF simple implementation has been used in several air quality studies especially for quantification of source-receptor relationships (Fiore et al., 2009). This is mainly due to its straightforward application in any photochemical model, without any substantial programming required.

2.5.2 Decoupled Direct Method

Various approaches have been introduced to avoid the need for differencing numerous brute-force simulations by computing pollutant sensitivities directly within the photochemical model. Some of the approaches such as the coupled direct method and Green's function method have been reported to be unstable or impractical for use in comprehensive photochemical models (Yang et al., 1997). However, the Decoupled Direct Method (DDM) (Dunker, 1984) and adjoint sensitivity analysis (Elbern et al., 1997) have been found to be stable and efficiently compute sensitivity relationships and thus have been more widely implemented in zero-dimensional (Milford et al., 1992; Gao et al., 1995), and three-dimensional models (Yang et al., 1997; Hakami et al., 2003).

Decoupled direct method, also referred as Tangent Linear Model (TLM), is used to calculate local sensitivity coefficients of the form given in Equation 2-19. Both coupled and decoupled direct methods derive sensitivity equations by differentiating the model equations (ADE) with respect to the parameters p_j (Yang et al., 1997):

$$\frac{\partial S_{ij}}{\partial t} = -\nabla.\left(\boldsymbol{u}S_{ij}\right) + \nabla.\left(\boldsymbol{K}\nabla S_{ij}\right) + \boldsymbol{F}S_{ij} + E_j\delta_{ij} \qquad (2\text{-}22)$$

where S_{ij} is the semi-normalized sensitivity of species i to the emissions of species j. $\boldsymbol{F}$ is the Jacobian matrix ($F_{ik} = \frac{\partial R_i}{\partial C_k}$) of the reaction rates, and δ_{ij} is the Kronecker delta function. Sensitivity equations are similar to concentration equations (ADE) and the only difference between them stems from the non-linearity in the model processes (mainly the chemistry).

In coupled direct method, the sensitivity equations are solved simultaneously with the model equations. Decoupled direct method, on the other hand, solves the model and

sensitivity equations separately and as a result, instability issues are resolved (Dunker, 1984). More examples of first order sensitivity analysis using the DDM technique and applied to photochemical grid models were collected by Hanna et al. (2001), Milford et al. (1992) and Seefeld and Stockwell (1999).

A distinction between DDM and brute-force is that DDM predicts local sensitivity coefficients $\frac{\partial C}{\partial \epsilon_j}$, representing responsiveness to infinitesimal changes in a parameter, whereas brute-force predicts responses to finite changes, $\frac{\Delta C}{\Delta \epsilon}$.

High-order DDM (HDDM) is developed to calculate higher-order local sensitivity coefficients (Hakami et al., 2003) to predict nonlinearity in atmospheric response. HDDM has been applied in several studies over the last decade to evaluate the sensitivities of surface level ozone to precursor emissions (Cohan et al., 2005; Jin et al., 2008; Kim et al., 2009).

2.5.3 Adjoint (Backward) Method

Brute-force and DDM compute the sensitivities of all model outputs to specified model inputs and parameters. Thus, they are suitable for characterizing how concentrations are impacted by a limited number of changes in emissions or other parameters of interest. However, for some applications, it may be desired to investigate how specific model outputs are influenced by numerous model parameters. As the number of input parameters of interest grows, it would quickly become cumbersome to compute forward sensitivities to each parameter. For these cases, an adjoint model is an efficient method for calculating sensitivity of a few model outputs to numerous model parameters.

Mathematical description of an adjoint is simplified by introducing the forward tangent linear model, which is analogous to the formulation of DDM (Equation 2-22). Both DDM and TLM refer to the forward method for calculating sensitivity because the perturbation is carried forward through the various model processes and in time. The ADE in operator form would be:

$$\boldsymbol{C} = \boldsymbol{M}(\boldsymbol{p}) \tag{2-23}$$

where the model output vector $\boldsymbol{C}$ is a function of the model input parameter vector $\boldsymbol{p}$, and $\boldsymbol{M}$ is the operator matrix representing the combination of all processes in the model.

A perturbation in inputs such as emissions causes a perturbation in output such as the ozone concentration. The corresponding TLM (Equation 2-22) in operator form can be written as:

$$\delta\boldsymbol{C} = \boldsymbol{L}(\delta\boldsymbol{p}) \tag{2-24}$$

where $\boldsymbol{L}$ is the Jacobian of the non-linear model operator $\boldsymbol{M}$. To evaluate model output or use outputs for policy making purposes, a metric (cost function) is defined:

$$J = J(\boldsymbol{C}) \tag{2-25}$$

which is a function of concentrations (e.g., could be the average ozone concentration over the domain or the square of the difference between the predicted and observed concentrations). The goal of the adjoint is to efficiently determine $\frac{\partial J}{\partial \boldsymbol{p}}$, or $\frac{\partial C_i}{\partial \boldsymbol{p}}$ in the simplest case.

Similar to DDM sensitivities (Equation 2-20), it is possible to define the adjoint variable λ_i as:

$$\lambda_i = \frac{\partial J}{\partial p_i} \tag{2-26}$$

The adjoint can then be defined by applying Lagrange multipliers to Equation 2-22 and integrating by parts by taking into account its initial and boundary conditions (Sandu et al., 2005) as:

$$-\frac{\partial \lambda_i}{\partial t} = \boldsymbol{u}.\nabla\lambda_i + \nabla.(\boldsymbol{K}\nabla\lambda_i) + \boldsymbol{F}_i^T\lambda + \varphi_i \tag{2-27}$$

where $\boldsymbol{F}_i^T$ is the i^{th} row of the transpose Jacobian (or the i^{th} column of the original Jacobian) for the chemical reaction rates, φ_i ,the forcing term, is the derivative of the

local cost function with respect to the concentration ($\frac{\partial g}{\partial C_i}$), and λ_i is the adjoint variable for a specific time and location ($\frac{\partial J}{\partial C_i}$). J is the overall cost function which is calculated by integration of the local cost function (g) over different times (t) and locations (w):

$$J = \int_t \int_w g(\boldsymbol{C}, t, w)\, dt\, dw \tag{2-28}$$

The structure of adjoint equation is similar to the tangent linear equation and can be integrated (for the most part) by the same numerical integration routines as the ADE. The main difference is the negative sign in the adjoint equation which results in backward integration in time and reversal of wind field (Hakami et al., 2007).

There are two ways of integrating adjoint equation, known as continuous adjoint and discrete adjoint. In the continuous approach, the adjoint equation is discretized and solved numerically. However, in the discrete approach, the adjoint of descretized TLM is developed and integrated. These two approaches provide different adjoint models. Adjoint of CMAQ which is used in this study considers hybrid approach for different processes (e.g., advection, diffusion, emission, deposition and gas-phase processes). The continuous adjoint approach is used for advection process and the discrete adjoint approach is applied for all the other processes (Hakami et al., 2007). More information about the mathematical foundation of adjoint sensitivity method applied to atmospheric transport and chemistry models can be found elsewhere (Elbern et al., 2000; Sandu et al., 2005; Hakami et al., 2007).

2.5.4 Adjoint method applications

Adjoint models have various applications in environmental studies, including meteorology, oceanography and atmospheric chemistry. It has been used in meteorological studies since the 1970's (e.g., Marchuk, 1974; Lamb et al., 1975). In air quality applications, an adjoint model was first developed for Lagrangian tropospheric and stratospheric models (Fisher and Lary, 1995; Elbern et al., 1997). Elbern and Schmidt (1999) conducted the first adjoint of a 3D Eulerian CTM to include chemistry.

Vukicevic and Hess (2000) applied the adjoint method to study the sensitivity of an inert gas-phase tracer over the Pacific in the tracer model HANK. Menut (2003) utilized the adjoint of a chemistry transport model in order to identify the parameters to which the ozone and nitrogen oxides peaks are more sensitive. Martien et al. (2006) developed an adjoint sensitivity analysis procedure for a three-dimensional photochemical model to predict changes in an ozone response due to changes in NO_x and VOC emissions or other model data. More recently, Henze et al. (2007) developed the adjoint of global CTM model (GOES-Chem) that included aerosol dynamics and thermodynamics.

The adjoint model of several CTMs have been developed; STEM (Sandu et al., 2005), CHIERE (Vautard et al., 2000; Menut et al., 2000; Schmidt and Martin, 2003), TM4 (Meirinik et al., 2006), and DRAIS (Nester and Panitz, 2006). An adjoint version of STEM was applied for inverse modeling of the black carbon to explore the spatial distribution of black carbon in Asia for April of 2001 (Hakami et al., 2005). Hakami et al. (2006) also adopted an adjoint of STEM for sensitivity analysis of NAAQS nonattainment ozone metric in the continental US. They used this method to quantify the interstate transport of ozone and its impacts on air quality impairment in some states in the continental US. The adjoint method was found to be a powerful framework for formal analysis of interstate, trans-boundary, and intercontinental transport of pollution in their work.

The adjoint of CMAQ model for gas-phase processes, which is used in this study, was developed by Hakami et al. (2007). They also introduced some applications for adjoint sensitivity analysis at regional scales using the adjoint of CMAQ. This was done by defining different cost functions such as population exposure metric and environmental exposure metric. Finally, they calculated the sensitivities of these metrics with respect to local temperatures to show applicability of CMAQ adjoint to regional climate studies.

3 Chapter: Methodology

"Knowledge is an unending adventure at the edge of uncertainty."- Jacob Bronowski

3.1 Overview

Models are developed to catch the complexities of the real world. They attempt to relate output variables (dependent variables) with a set of input variables (predictor or independent variables) based on some simplifying assumptions. For example, changes in mortality rate (ΔM) can be estimated as an output of a model with respect to an input, such as changes in pollutant concentration (ΔC). Epidemiological methods would observe the changes in concentration of pollutants, and the corresponding changes in mortality rate are valuated, therefore developing a relationship such as:

$$\Delta \boldsymbol{M} = \boldsymbol{f}(\Delta \boldsymbol{C}) \tag{3-1}$$

where $\boldsymbol{C}$ is the concentration vector, $\boldsymbol{M}$ is the vector of mortality rate at different locations, and **f** is the function that relates input to output. Estimation of input parameters may be difficult and inaccurate, or in other words, uncertain. The uncertainties in input parameters propagate through the model and lead to uncertain predicted values. Gaining insights about the uncertainties can help us decide whether it is worthwhile to gather more information, carry out more attentive uncertainty assessments, or improve the model. Given this information, one could then, opt for the best choice which would reduce the uncertainty in the results by the largest degree (Morgan and Henrion, 1990). The objective of this thesis is to introduce uncertainty analysis for desired health metrics that are generalized by Equation 3-1.

3.2 Uncertainty analysis

Uncertainty analysis is defined by Morgan and Henrion (1990) as the computation of the total uncertainty induced in the output by quantified uncertainties in the inputs and model, and the attributes of the relative importance of the input uncertainties in terms of their contributions. Figure 3-1 summarizes the main steps of uncertainty analysis which is

followed in this study. Total model uncertainty is defined by the sum of the model uncertainty, variability, and uncertainty of input data. In this work, there is a sequence of two models, CMAQ (Community Multi-scale Air Quality) and Epidemiological model. CMAQ is a Chemical Transport Model (CTM) which is the most widely used air quality model in the U.S. and worldwide.

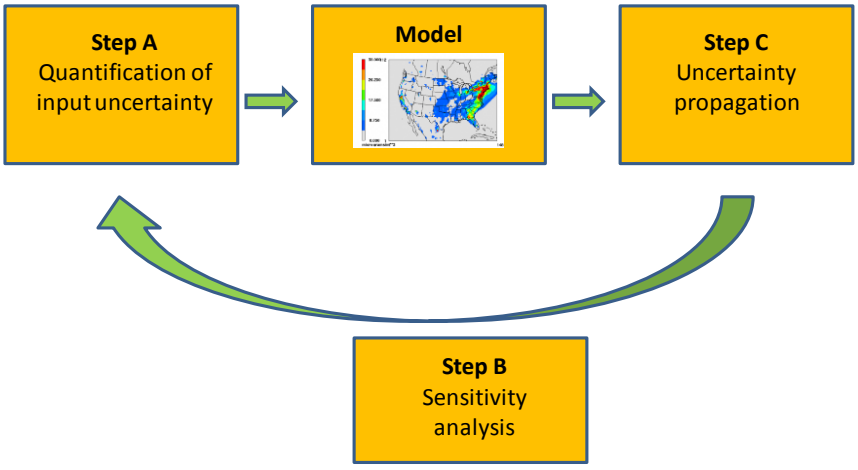


Figure 3-1: General sketch for uncertainty analysis

Figure 3-2 provides a representation of the input and output of each sub-model.

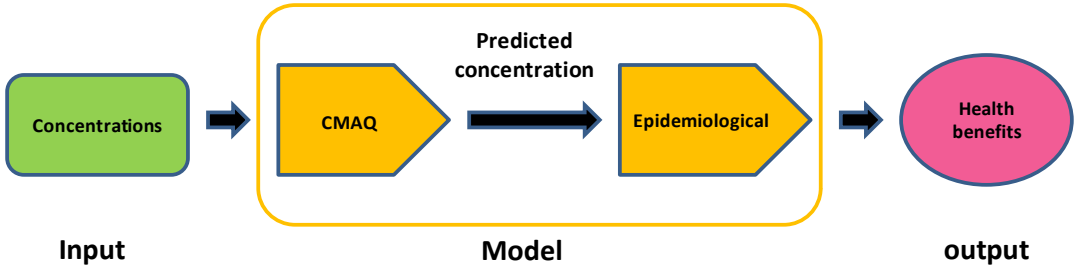


Figure 3-2: A schematic view of model components

Uncertainties associated with model formulation may be due to erroneous or incomplete representation of the dynamic and chemistry of the atmosphere, numerical solution techniques, and choice of modeling domain and grid structure. Our objective is estimating the value of information contained in surface observations; therefore, in this study uncertainty of input data is restricted to that of initial condition (i.e., observational

data). Consequently, other sources of uncertainty such as model formulation uncertainty are not considered.

CMAQ takes the observed concentrations as the model input and estimates future concentrations. Afterwards, an epidemiological model predicts a change in mortality or its monetized valuation (ΔM) associated with changes in predicted pollutant concentrations (ΔC) resulting from CMAQ (Figure 3-2). As a result, the information obtained from observations can be measured by the amount of uncertainty reduction in the estimated output (ΔM). In the scope of this work, the uncertainty in the output values (i.e., Health benefits associated with ozone reductions) can be reduced by improving the accuracy of model inputs (i.e., ozone concentrations) through adding new monitoring stations.

In Figure 3-1, three steps are identified:

- Step A: Quantification of input uncertainty

Air quality measurements are only available at certain locations and are usually accurate and precise at those locations. The model requires extensive inputs; therefore, when the model uses observations as the starting point (e.g., in data assimilation applications), for the grid cells without observation, interpolated values are necessary for the input variables. Spatial interpolation introduces errors into these inputs, which propagate through the model and contributes in uncertainty of the results. The interpolation errors are not likely to be spatially homogeneous, but will depend on the number and proximity of the data points used for each grid point interpolation. As it is shown in Figure 3-3, the higher interpolation error (i.e., higher uncertainty in input parameter with green probability density function) results in more uncertain results (comparing to the input with lower uncertainty and the same sensitivity coefficient shown in blue). In this step, the variance vector of initial concentrations is determined by associating it with spatial interpolation error (i.e., Kriging variance) as an accuracy measure. Before kriging, the uncertainty of ozone concentrations is only available at measuring points due to the instrumentation errors (consideration of measurement errors in kriging procedure will be

discussed in section 3.3.3). After applying the kriging method, variance of the estimated concentrations for different grid cells would be provided as a measure of uncertainty over the entire domain.

- Step B: Sensitivity analysis

The chemical transport models require a large number of inputs (e.g., initial concentrations of different species at each grid cell). Uncertainty analysis should focus on

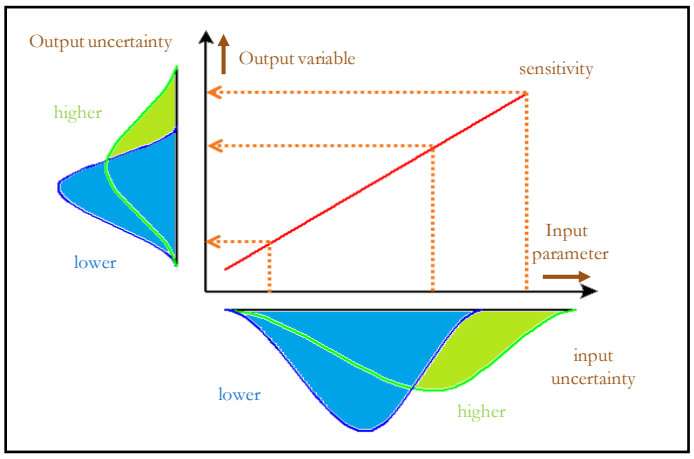


Figure 3-3: Contribution of input uncertainties in the output uncertainty (given the same sensitivity coefficient)

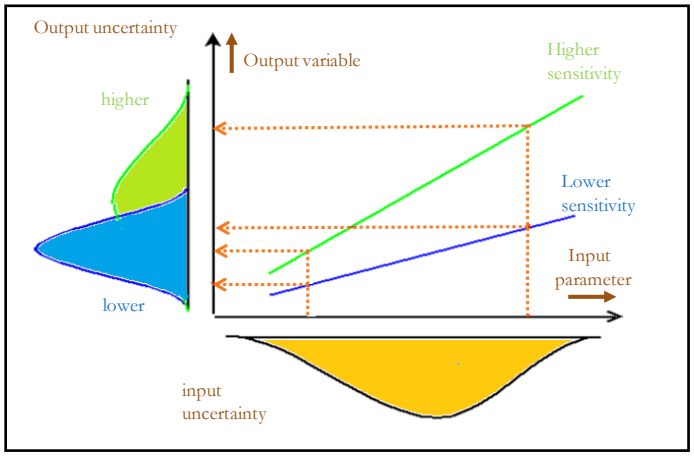


Figure 3-4: Contribution of an uncertain input with two levels of sensitivity in the output uncertainty

those inputs that are likely to contribute most to the uncertainties related to model predictions. Sensitivity Analysis (SA) methods are applied to determine the contribution of a specific, uncertain variable to the overall uncertainty in model results.

Figure 3-4 illustrates the impact of sensitivity on model output uncertainty. The input propagated with higher sensitivity (slope of the green line) contributes more to the uncertainty of model prediction. In step B, the adjoint of CMAQ is employed to estimate the sensitivity of health benefits from reduced short-term mortality with respect to ozone concentration reduction all over the model domain.

- Step C: Uncertainty propagation

For any model, it is often desirable to provide some measure of uncertainty in results. The methods for calculating the uncertainty in model outputs induced by the uncertainties in its inputs are referred to as uncertainty propagation. Uncertainty propagation allows an assessment of how the output variance will decline for a given reduction in the variance from one or more inputs. Clearly, the output variance will mainly improve from a reduction of the input variance that contributes most to the uncertainty in the model output. In this step, the perturbation method is applied which is based on the Taylor series expansion of the model response around the mean value of the input parameters.

The aforementioned main steps are described in detail in the following sections. In section 3.3, Kriging interpolation method is introduced as a mean of quantifying input uncertainties. Details on adjoint sensitivity approach are provided in section 3.4 and the last section, 3.5, is dedicated to perturbation method which is selected for uncertainty propagation.

3.3 Quantification of input uncertainty

Step one in this spatial uncertainty analysis is quantification of input uncertainty by applying kriging method to interpolate the input data.

3.3.1 Observation Data

Hourly ground level ozone observations for 2007 are retrieved from U.S. EPA Air Quality System (AQS) (U.S. EPA, 2014) (Figure 1-1). More than a thousand ozone monitoring stations (urban, suburban, and rural) are now active in AQS that cover the majority of the continental U.S.

Uncertainty associated with the ozone data is not available for each individual monitoring site. Thus, the instrument error (noise) is considered as measurement error. Two measurement methods for ozone are approved in the U.S. by the U.S. EPA: one is based on the chemiluminescence that occurs when ozone and ethylene react, and the other on the attenuation of Ultraviolet (UV) radiation by ozone. Ozone exhibits a strong absorption band in the ultraviolet region at 254 nm and this feature is the basis of the photometric measurement method for ozone. This method is used in almost all of the state and local air monitoring stations that reported data to EPA AQS from 2005 to 2009.

U.S. Federal Regulations (40 CFR Part 53) require that O_3 monitors (based on UV photometry) should report hourly average concentrations with a precision of 10 ppb and noise of 5 ppb (U.S. EPA, 2013). In this study, all monitors are regarded to have the same measurement error which is equal to the maximum allowable noise (5 ppb).

After obtaining all ground-level ozone measurements, the spatial distribution of this pollutant is analyzed for the ozone season and ozone levels are estimated at other locations where direct measurements were not carried out. It is recognized that the statistical approach, geostatistical methods or kriging, enjoys several advantages over other interpolation techniques (Liu et al., 1996). The most important feature which sets kriging apart from the deterministic models is the ability to estimate prediction error for each interpolation point, i.e., the kriging variance. Secondly, kriging interpolates observations using weights that do not depend upon data values. Third, it is an exact interpolator, which means that the estimate at any observational point is the observation itself (Barnes, 1980). Last but not least, ozone which is the pollutant of concern is

amenable to kriging concept because of its spatially correlated and constant variance across well-defined geographic regions (Lefohn et al., 1987).

As previously discussed in background, various kriging methods have been introduced and applied in the literature. The choice of kriging method depends on the characteristics of the data and the type of spatial model desired. Simple Kriging (SK) is based on "knowledge of the mean and the covariance", which are not available for the current problem. Also, ozone data does not display trends at the scale of the entire United States that can be modeled by simple functions. Because of this fact, Universal Kriging (UK) is not chosen for this study either. Ordinary Kriging (OK), which is the most commonly kriging method in practice and is suitable for the dataset of this study, is employed for every grid cell separately to estimate the concentration at that cell.

3.3.2 Searching neighborhood

Before embarking on an ordinary kriging modeling, it is important to build a dataset (i.e., observed concentrations) for the grid cell of interest. It is assumed that concentrations at locations which are further have less spatial autocorrelation with the prediction grid cell. Therefore, distant observations that have little influence are eliminated by establishing search neighborhoods. Not only is there less relationship with distant locations, but it is also possible that these locations have detrimental influence if they are located in an area much different than the prediction location. Another reason to use search neighborhoods is for computational speed. The smaller the search neighborhood, the faster the predictions can be made.

By the assumption that there are no directional influences on the spatial autocorrelation of the data (isotropic assumption), the shape of the neighborhood is considered as a circle (Figure 3-5). The measured concentrations within this circle are used in each kriging procedure. Initially, a fixed search radius is set for all grid nodes (e.g., 200 km) without considering the spatial autocorrelation. This radius will then be modified in step 2 of the

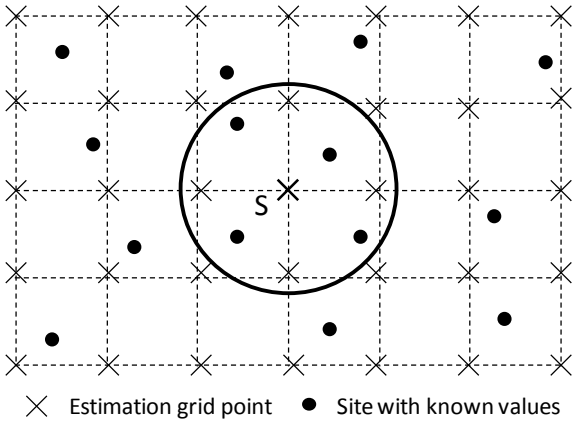


Figure 3-5: Estimation grid with search radius around grid point S

kriging method by deriving the range parameter, a**,** of the semivariogram. The range indicates the distance over which data are correlated; therefore it is used as a measure of suitability for the assumed radius of neighborhood.

3.3.3 The ordinary kriging model

Ordinary kriging is a common type of kriging in practice. OK assumes that there is a constant but unknown mean across the spatial domain of interest. Suppose concentration is a random field Z (.) which is intrinsically stationary. For any location s

$$E[Z(s)] = \mu \tag{3-2}$$

$$Var[Z(s) - Z(s+h)] = 2\gamma(h)$$

where $\gamma(h)$ is the semivariogram and a function of the distance h separating two locations. For the kriging system, the concentration at location s_0 is predicted as weighted average of the surrounding data:

$$\hat{Z}(s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i) \tag{3-3}$$

where $Z(\boldsymbol{s}_i)$ is the measured concentration at the ith location, λ_i is an unknown weight for the measured concentration at the ith location (Figure 3-6).

To ensure the predictor is unbiased for the unknown point, the sum of the weight λ_i must equal one. Kriging minimizes the mean squared error of prediction:

$$min\ \sigma_e^2 = E\left[Z(s_0) - \hat{Z}(s_0)\right]^2 \qquad or \qquad min\ \ \sigma_e^2 = E[Z(s_0) - \sum_{i=1}^{N} \lambda_i Z(s_i)]^2$$

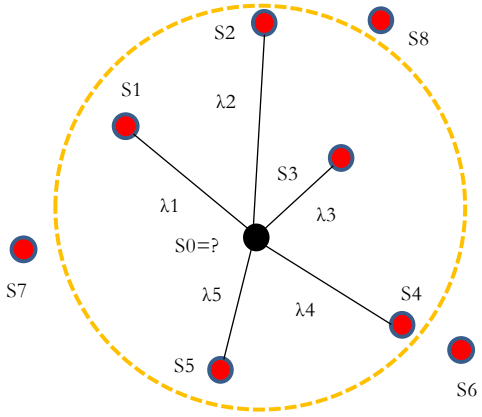


Figure 3-6: Kriging spatial interpolation

For intrinsically stationary process the last equation can be written as (Le and Zidek, 2006):

$$\sigma_e^2 = 2\sum_{i=1}^{n} \lambda_i \gamma(s_0 - s_i) - \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\, \lambda_j \gamma(s_i - s_j) \qquad (3\text{-}4)$$

The minimization is carried out over $(\lambda_1, \lambda_2, \dots, \lambda_n)$, subject to the constrain $\sum_{i=1}^{n} \lambda_i = 1$. Therefore, the minimization problem can be written as:

$$\min\ 2\sum_{i=1}^{n} \lambda_i \gamma(s_0 - s_i) - \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\, \lambda_j \gamma\left(s_i - s_j\right) - 2\mu\left(\sum_{i=1}^{n} \lambda_i - 1\right) \qquad (3\text{-}5)$$

where μ is the Lagrange multiplier. After differentiating (3-5) with respect to $\lambda_1, \lambda_2, \dots, \lambda_n$, and μ and setting the derivatives equal to zero:

$$-\sum_{j=1}^{n} \lambda_j \gamma\left(s_i - s_j\right) + \gamma(s_0 - s_i) - \mu = 0, \quad i = 1, \dots, n \qquad (3\text{-}6)$$

and

$$\sum_{i=1}^{n} \lambda_i = 1 \qquad (3\text{-}7)$$

Using matrix notation the previous system of equations can be written as:

$$\boldsymbol{\Gamma} * \boldsymbol{\lambda} = \boldsymbol{g} \tag{3-8}$$

Or

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \cdots & \gamma_{1n} & 1 \\ \gamma_{21} & \gamma_{22} & \gamma_{23} & \cdots & \gamma_{2n} & 1 \\ \cdots & \cdots & \ddots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \gamma_{n3} & \cdots & \gamma_{nn} & 1 \\ 1 & 1 & \cdots & \cdots & 1 & 0 \end{bmatrix} * \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \\ \vdots \\ \vdots \\ \gamma_{n0} \\ 1 \end{bmatrix}$$

The goal is to solve the equations for all of the λ_is (the weights), so the predictor and its variance can be formed. The elements of the matrices and vectors can be filled if the semivariogram is known. The way to calculate the semivariogram is explained in the kriging steps below. The gamma matrix $\boldsymbol{\Gamma}$ contains the modeled (i.e., theoretical) semivariogram values between all pairs of sample locations, where γ_{ij} denotes the theoretical semivariogram values based on the distance between the two samples identified as the i^{th} and the j^{th} locations. The vector $\boldsymbol{g}$ contains the theoretical semivariogram values between each measured location and the prediction location, where γ_{i0} denotes the modeled semivariogram values based on the distance between the i^{th} sample location and the prediction location. The unknown μ in the vector $\boldsymbol{\lambda}$ is also estimated and it arises because of the unbiasedness constraint.

The following steps are followed in ordinary kriging method:

1) The first step is calculating the empirical semivariogram. For creating the empirical semivariogram, the distance and squared difference between each pair of locations are needed. The distance between two locations is calculated by using the Euclidean distance:

$$h = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \tag{3-9}$$

The empirical semivariance between two data points is an important concept in geostatistics and is defined as:

$$\gamma(s_i, s_j) = \gamma(h) = \frac{1}{2}\mathbb{E}\left[Z(s_i) - Z(s_j)\right]^2 \quad (3\text{-}10)$$

Here the empirical semivariogram $\gamma(h)$ is equal to half of the square difference between pairs of data points. A plot of γ(h) against h is known as semivariogram. Kriging is built on the assumption that things that are close to one another are more alike than those farther away (i.e., spatial autocorrelation). The empirical semivariogram is a mean to explore this relationship.

With larger datasets, the number of pairs of locations increases and semivariogram becomes difficult to manage and interpret. Therefore, the pairs of locations are grouped. This treatment is referred to as binning. There are some guidelines on the optimal size of the bins. We use the recommendation by U.S. EPA (2004) based on the similarity of the process of generating an empirical variogram to the process of generating a histogram. According to this suggestion, the number of bins is set equal to the square root of the number of data points. It represents a compromise between the number of bins and the number of pairs in each bin. After grouping pairs into each lag class or bin based on their locations, the average semivariance between sample points seprated by lag h for each group needs to be calculated as:

$$\gamma(h) = \frac{1}{2N}\sum_{i=1}^{N}[Z(\boldsymbol{s}_i) - Z(\boldsymbol{s}_i + h)]^2 \quad (3\text{-}11)$$

where N is the number of pairs of sample points sorted by direction in the bin.

2) The second step is fitting a model. The empirical semivariogram is modeled with a continuous function that represents a theoretical semivariogram. First, the average semivariance versus average distance of the bins should be plotted (i.e., empirical semivariogram). Then a curve that provides the best fit through the points in the semivariogram needs to be defined. This entails finding a curve such that the square difference between each point and the curve is as small as possible. This is referred to as

the least-squares fit. Once the model is fit, the derived theoretical semivariogram helps us determine semivariogram values for various distances.

The most common models used in the variogram modeling process are linear, spherical, exponential, and Gaussian. The linear model is tried first. In case a good fit with linear regression cannot be obtained (i.e., the coefficient of determination, or R^2 of the regression is not close to one), then the spherical model (the most common correlation function for estimating ozone concentrations) is applied to represent the theoretical semivariogram:

$$\gamma(h) = \begin{cases} C_0 + C_1\left[\frac{3h}{2a} - \frac{h^3}{2a^3}\right] & h \le a \\ C_0 + C_1 & h > a \end{cases} \tag{3-12}$$

where h is the lag distance and C_0 is the nugget (i.e., the semivariance at a lag distance zero). $C_0 + C_1$ is the sill (i.e., upper bound of the semivariogram), and a is the range (i.e., lag distance at which the semivariogram reaches the sill). By the least-squares method, the parameters (C_0, C_1, a) can be extracted for the data set of each grid cell. These model features (or parameters) are unique for each grid cell.

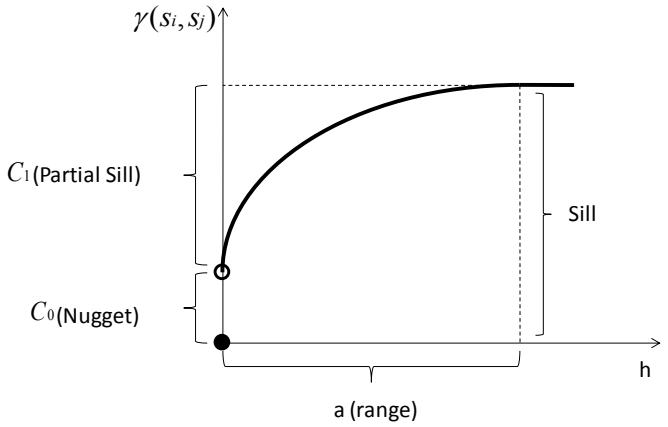


Figure 3-7: Semivariogram parameters nugget, sill, and range

The nugget provides an indication of the amount of measurement errors that are specific to the measuring methods and devices. The diagonal of the Γ matrix contains $\gamma(0)$. If we trust our data completely, i.e., $\gamma(0) = 0$, then the kriging predictions go exactly through

the observations. The kriging variance is zero at the observations, because with no residual, the value of the spatial process is exactly known. An example of this is shown in Fig. 3-8 (top), with prediction in one dimension.

Alternatively, if we acknowledge measurement or sampling error in our data, i.e., $\gamma(0) = C_0$, kriging interpolates depend on the residual variation which is considered to be instrument (measurement) error (i.e. 5 ppb), which one would want to smooth over. In Fig. 3-8 (bottom), the predictions and variances under measurement error is shown. Now the residual is treated as error, and the prediction at the observation is not the observation itself, because the observation is contaminated with error.

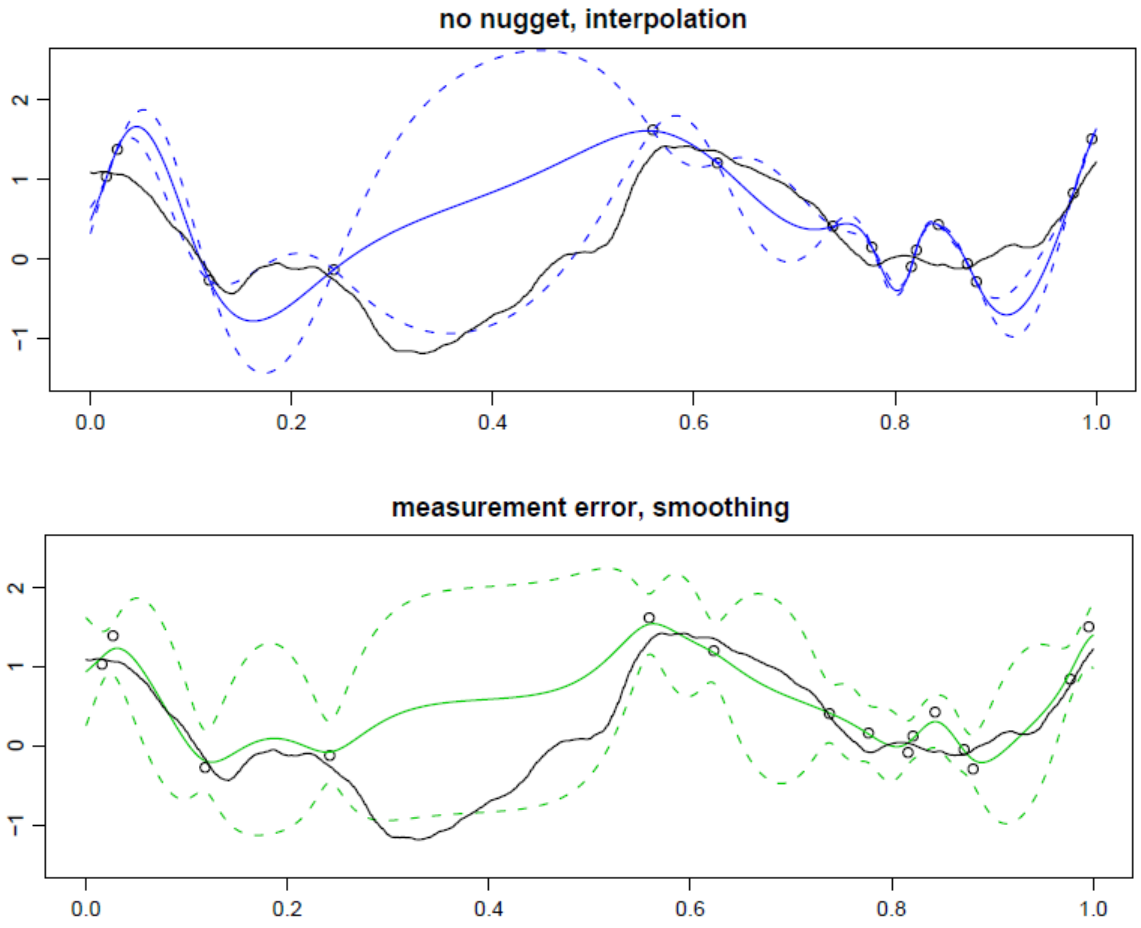


Figure 3-8: Kriging predictions under two scenarios (Cressie, 1993)

Figure 3-8 presents Kriging predictions under two scenarios. In each scenario, the black line is the true underlying smooth process, generated from the sum of a very smooth

process and a more local, less smooth process, while the observations (the dots) also include a component of independent errors. The solid colored lines are the kriging predictions and the dashed lines are point-wise confidence intervals. In (top), we assume a nugget of zero, which fits an interpolator. In (bottom), the residual is assumed to be instrument error, so kriging serves as a smoother.

For nonlinear models, it is impractical to calculate R-squared as a measure of goodness of fit. Instead, the standard error of estimate, S value, can be calculated by:

$$S = \sqrt{\frac{\sum_{i=1}^{N}(\gamma(h)-\hat{\gamma}(h)^2}{N}} \tag{3-13}$$

The S value provides an overall measure of how well the spherical model fits the data. Smaller values are better since it indicates that the observations are closer to the fitted curve.

3) The third step is solving the equations. The equations for ordinary kriging are contained in matrices and vectors that depend on spatial autocorrelation among the measured sample locations and prediction location ($\boldsymbol{\Gamma}$ and $\boldsymbol{g}$). The autocorrelation values come from the theoretical semivariogram described in step 2. After creating $\boldsymbol{\Gamma}$ matrix and $\boldsymbol{g}$ vector**,** $\boldsymbol{\lambda}$**,** which contains the weights to assign to the measured values around the prediction location, can be derived by the following formula:

$$\boldsymbol{\lambda} = \boldsymbol{\Gamma}^{-1} * \boldsymbol{g} \tag{3-14}$$

where $\boldsymbol{\Gamma}^{-1}$ is the inverse matrix of $\boldsymbol{\Gamma}$. By performing LU factorization method, the inverse of $\boldsymbol{\Gamma}$ is obtained.

4) The fourth and last step is making a prediction and calculating the associated uncertainties with the predicted value. By finding the weights, the estimator can be computed by Equation 3-3. For finding the variance of the estimator, Equation 3-6 is multiplied by λ_i and summed over i = 1, …, n to get:

$$-\sum_{i=1}^{n}\lambda_i\sum_{j=1}^{n}\lambda_j\,\gamma(s_i-s_j)+\sum_{i=1}^{n}\lambda_i\gamma(s_0-s_i)-\sum_{i=1}^{n}\lambda_i\mu=0$$

or

$$-\sum_{j=1}^{n}\lambda_i\lambda_j\,\gamma(s_i-s_j)+\sum_{i=1}^{n}\lambda_i\gamma(s_0-s_i)-\sum_{i=1}^{n}\lambda_i\mu=0$$

Therefore,

$$\sum_{j=1}^{n}\lambda_i\lambda_j\,\gamma(s_i-s_j)=\sum_{i=1}^{n}\lambda_i\gamma(s_0-s_i)-\sum_{i=1}^{n}\lambda_i\mu$$

Substituting this result into Equation 3-4 results in:

$$\sigma_e^2=\sum_{i=1}^{n}\lambda_i\gamma(s_i-s_0)+\mu \qquad (3\text{-}15)$$

Using matrix notation, the kriging variance can be written as:

$$\sigma_e^2=\boldsymbol{g}\,.\boldsymbol{\lambda} \qquad (3\text{-}16)$$

The square root of the kriging variance is called the kriging Standard Deviation (SD). Kriging standard deviations are used to quantify the spatially varying interpolation uncertainties.

3.4 Sensitivity analysis

Step two in the uncertainty analysis is to perform sensitivity analysis. In order to efficiently assess uncertainties associated with health metrics, knowledge of how the variation in the short-term O_3 mortality can be apportioned to different sources is required. Adjoint (backward) sensitivity analysis, as discussed in section 2.5.3, is one of the best options to show how a model output defined in a single grid cell or across the domain is affected due to a perturbation in numerous inputs. Adjoint equation results in the adjoint variables, which contain information about the sensitivity of health metrics with respect to ozone concentrations at various locations and times.

The adjoint of CMAQ used in this study, was developed for gas-phase processes and was implemented in CMAQ version 4.5.1 (Hakami et al., 2007). The Community Multiscale Air Quality (CMAQ) model is a Chemical Transport Model (CTM) which is developed by the U.S. EPA (Byun and Schere, 2006) to meet the needs of both environmental managers and scientists to evaluate the impact of air quality management practices and simulate chemical and physical interactions in the atmosphere. CMAQ model has been widely used by the air quality community worldwide to model multiple air quality issues, such as tropospheric ozone, fine particles, toxics, acid deposition, and visibility degradation. This model also has capabilities to solve air quality problems in multiple scales including the urban and regional scales (Sandu and Chai, 2011).

3.4.1 Modeling episode and domain

The CMAQ domain in this study is the United States with a 36 km grid size, and 34 vertical layers. The simulation period is from July 1st to September 30th of 2007, and these months are chosen because it is favorable to ozone formation throughout the U.S. The U.S. domain is selected for this study, due to the availability of comprehensive emission inventories and meteorological data for this domain.

3.4.2 CMAQ platform

The CMAQ modeling system has three main components (Figure 3-9):

- *The emission model* provides CMAQ-ready gridded and allocated emissions files from inventory data. In this thesis, the Sparse Matrix Operator Kernel Emission (SMOKE) version 2.4 (CEP, 2009) is used as the emission model. The SMOKE model converts the resolution of the emission inventory data to the resolution needed by the air quality model. The emissions inventories are typically made up of total annual emissions prepared for each emission source (e.g., stationary, mobile, etc). Air quality models; on the other hand, usually need hourly emission data for each species in each grid cell. Consequently, an emission inventory is

transformed through temporal/spatial allocation and chemical speciation to achieve the input requirements of the CMAQ.

- *The meteorological model* simulates the atmospheric state and motion. We use the Weather Research and Forecasting (WRF, version 3.1) model as the meteorological model. WRF includes two main parts: the WRF Preprocessing System (WPS) and the dynamic simulation model. The WPS prepares inputs to the dynamic model; then the dynamic model vertically interpolates the data. The modeling domain of WRF used here (version 3.1) covers North America with 165 columns, 129 rows and 35 vertical layers. As the WRF output file cannot be directly used as input in CMAQ, it needs to be preprocessed by Meteorology Chemistry Interface Processor (MCIP) (Byun and Schere, 2006). MCIP is able to change the domain into a smaller domain horizontally and vertically. In this thesis MCIP projects the WRF files into a smaller domain with 148 columns, 112 rows, and 34 layers which is used for the CMAQ simulations.
- *The CMAQ Chemical Transport Model (CCTM)* includes advection, diffusion, gas-phase chemical reactions, aqueous-phase reactions, aerosol dynamics and thermodynamics, and dry and wet deposition (Byun and Schere, 2006).

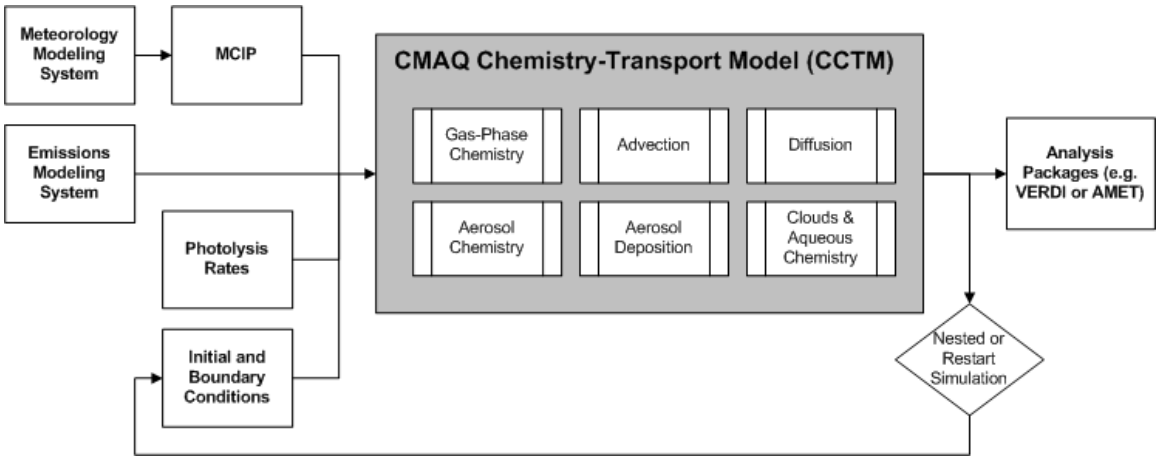


Figure 3-9: CMAQ chemistry-transport model and associated preprocessors (CEP, 2012)

Initial and boundary conditions are the other required inputs of CTM, as well as photolysis rate coefficients. In the beginning of a model simulation, the first time step is

produced by an initial condition processor (ICON). Afterwards, the boundary conditions processor (BCON) is responsible for providing the concentrations of all the species at the boundaries of the model domain. The JPROC processor calculates the clear sky photolysis rate for different times and latitudes. The photolysis rates affect the rate of chemical reactions and depend on the time of day, season, atmospheric composition, and latitude.

In adjoint of CMAQ, the sequence of processes is similar to the forward CMAQ, but model integration is run backward in time. The adjoint CMAQ also requires the concentration of species for nonlinear operators (for calculation of the transpose of the Jacobian in Equation 2-27) such as the chemistry solver. A discrete adjoint scheme is employed in adjoint CMAQ for chemistry, diffusion, and vertical advection and a continuous adjoint is employed for horizontal advection. This adjoint model uses the Kinetic Pre-Processor (KPP) version 2.2 for chemistry integration, (Damian et al., 2002; Daescu et al., 2000), the piecewise parabolic method (Colella and Woodward, 1984) for the horizontal advection process, and an upwind first order finite difference method for vertical diffusion.

3.4.3 CMAQ-Adjoint performance evaluation

In air quality models, the performance evaluation can be done by comparing the modeled concentrations and observation data. However, in an adjoint model, there is no observation to be compared with because the outputs are derivatives. This shortcoming is a general limitation of sensitivity methods. The adjoint model used in this work has been validated extensively against the forward model (Hakami et al., 2007). The uncertainties and limitations existing in CMAQ remain unaddressed in the adjoint of CMAQ as well.

As explained in section 2.5.3, the adjoint model calculates the derivatives of a metric known as the cost function with respect to model inputs (Equation 2-28). The forcing term is the local derivative of the cost function with respect to concentrations. In this work, the adjoint cost function is defined as the monetary value of mortality (M) resulting from short-term exposure to O_3 which will be discussed in detail below.

3.4.4 Adjoint cost function and forcing term

Linkage between epidemiological models and adjoint calculations is established through appropriate definition of the adjoint cost function. A change in mortality valuation (ΔM) associated with a change in pollutant concentration (ΔC) is often given by:

$$\Delta M = M_0 \,.P.V_{SL}\,(1 - e^{-\beta \Delta C}) \tag{3-17}$$

Where M_0 is the baseline non-accidental mortality rate, P is the population, V_{SL} is the Value of a Statistical Life (VSL), and β is the concentration response factor based on epidemiological models. From Bell et al. (2004), β value is taken as 4.27×10^{-4} ppb^{-1} for 8-hr average O_3. Total population and baseline mortality rates for the U.S. are obtained for each county from the Centers for Disease Control and Prevention (CDC). Non-accidental mortality rates are calculated from International Classification of Disease (ICD)-10 codes A-R as in Bell et al. (2004).

VSL is the most common mortality valuation metric and is a measure of an individual's willingness-to-pay to reduce their probability of death (Alberini et al., 2006). VSLs in 2006 equivalents (adjusted using the Consumer Price Index) of $6.8M USD are applied for the U.S. (U.S. EPA, 2010).

As β is usually a small number, Equation 3-17 can be linearized using a Taylor expansion (i.e., $\Delta M = M_0 \,.P.V_{SL}.\beta.\Delta C$). The adjoint cost function is introduced to the adjoint model through the adjoint forcing term (φ) defined as (Pappin and Hakami, 2013):

$$\varphi = \frac{\partial M}{\partial C} \approx \frac{\Delta M}{\Delta C} \approx M_0 \,.P\,.V_{SL}\,.\beta \tag{3-18}$$

The forcing terms in adjoint model can be regarded as source of influences on the adjoint cost function. Forcing terms are injected into the adjoint model and derivatives evolve backward in time through each process and towards the originating sources. More details on using the adjoint model for ozone mortality cost valuation can be found elsewhere (Pappin and Hakami, 2013).

3.5 Uncertainty propagation

Step three of the uncertainty analysis involves transforming the uncertainties in model inputs and parameters into the uncertainty in model outputs. There are three main types of methods for uncertainty propagation depending on the application (Sudret and Der Kiureghian, 2000)

- *Second moment analysis* methods can be used when the mean value and standard deviation of the response are of interest. These methods provide the mean representing the central or expected tendency of the response, and the variance-covariance representing the amount of scattering or variation around the mean. Methods such as the perturbation method, the weighted integral method or the quadrature method belong to this category.

- *Structural reliability analysis* methods are applied when the tail of the response PDF (Probability Density Function) is of interest. This method helps in computing a probability of failure, where "failure" is defined as the event "Y is exceeding a specified threshold". FORM (First-order Reliability Method)/SORM (Second-order Reliability Method), importance sampling, or directional simulation methods can be used in this case.

- *Spectral* methods are employed when the whole PDF of the response is desired. Monte Carlo Simulation (MCS) is the basic approach to solve this type of problems.

It is often difficult to obtain the probability distributions of the response and in some cases the whole PDF or even the tail of it is not required. In this study, calculating the standard deviation of the model output around its mean value is of interest; hence second moment analysis is the most suitable approach for this purpose. As far as second moment analysis is concerned, the perturbation method remains the best compromise between accuracy and efficiency, especially if the coefficient of variation of the input random variables is not too large and the model is not too non-linear (Sudret, 2007).

In this section, the perturbation method is presented which is based on Taylor series expansion, also known as first-order approximation. It greatly simplifies the process of propagating uncertainty through the model and yields analytical expressions for the mean and variance of the model output.

3.5.1 Estimating the mean value of the response

Suppose $\boldsymbol{X}$ is a vector of n uncertain inputs. The Taylor series expansion of a function $Y = g(\boldsymbol{X})$ about the mean values of $\boldsymbol{X}$, i.e., $\mu_{\boldsymbol{X}} = \{\mu_{X1}, \mu_{X2}, \dots, \mu_{Xn}\}^T$ is given by:

$$E[Y] = E[g(\boldsymbol{X})] \approx g(\mu_{\boldsymbol{X}}) + \sum_{i=1}^{n} E[X_i - \mu_{Xi}] \left[\frac{\partial g(\boldsymbol{X})}{\partial X_i}\right]_{\mu_{\boldsymbol{X}}}$$

$$+ \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} E\left[(X_i - \mu_{Xi})(X_j - \mu_{Xj})\right] \left[\frac{\partial^2 g(\boldsymbol{X})}{\partial X_i \partial X_j}\right]_{\mu_{\boldsymbol{X}}} + \cdots \quad (3\text{-}19)$$

Here the derivatives are evaluated at the mean values of $\boldsymbol{X}$. By definition, $E[(\boldsymbol{X} - \mu_{\boldsymbol{X}})] = 0$ and $E\left[(X_i - \mu_{Xi})(X_j - \mu_{Xj})\right] = Cov(X_i, X_j)$ is the covariance matrix of X. Thus the approximation of the mean value is:

$$E[Y] \approx g(\mu_{\boldsymbol{X}}) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j) \left[\frac{\partial^2 g(\boldsymbol{X})}{\partial X_i \partial X_j}\right]_{\mu_{\boldsymbol{X}}} \quad (3\text{-}20)$$

In Equation 3-20, the first term in the right hand side is the first order approximation of the mean output value $E[Y]$. It is obtained as the output of the model evaluated at the mean value of the input parameters. Moreover, the second term depends on the covariance matrix and the Hessian matrix of the model computed at the mean value point. In the case of independent input parameters, the covariance matrix is diagonal and contains the variance of each input parameter. Equation 3-20 reduces in this case to:

$$E[Y] \approx g(\mu_{\boldsymbol{X}}) + \frac{1}{2}\sum_{i=1}^{n} Var(X_i) \left[\frac{\partial g(\boldsymbol{X})}{\partial X_i}\right]^2_{\mu_{\boldsymbol{X}}} \quad (3\text{-}21)$$

3.5.2 Estimating the variance of the response

From Equation 3-20, the variance of the response may be computed as follows:

$$Var[Y] = E[(Y - E[Y])^2] \approx E[\left(Y - g(\mu_{\boldsymbol{X}})\right)^2] \tag{3-22}$$

From the first order expansion in Equation 3-19 carried out at the mean value $\mu_{\boldsymbol{X}}$, the latter equation simplifies into:

$$Var[Y] \approx E[\left(\sum_{i=1}^{n}(X_i - \mu_{X\boldsymbol{i}})\left[\frac{\partial g(\boldsymbol{X})}{\partial X_i}\right]_{\mu_{\boldsymbol{X}}}\right)^2]$$

$$\approx \sum_{i=1}^{n}\sum_{j=1}^{n} E[(X_i - \mu_{X\boldsymbol{i}})\left(X_j - \mu_{X\boldsymbol{j}}\right)\left[\frac{\partial g(\boldsymbol{X})}{\partial X_i}\right]_{\mu_{\boldsymbol{X}}}\left[\frac{\partial g(\boldsymbol{X})}{\partial X_j}\right]_{\mu_{\boldsymbol{X}}} \tag{3-23}$$

Hence, the first order expansion of the variance of the response leads to:

$$Var(Y) = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j)\left[\frac{\partial g(\boldsymbol{X})}{\partial X_i}\right]_{\mu_{\boldsymbol{X}}}\left[\frac{\partial g(\boldsymbol{X})}{\partial X_j}\right]_{\mu_{\boldsymbol{X}}} \tag{3-24}$$

In case of uncorrelated input variables, Equation 3-24 further simplifies into (Ayyub and Klir, 2006):

$$Var(Y) \approx \sum_{i=1}^{n} Var(X_i)\left[\frac{\partial g(\boldsymbol{X})}{\partial X_i}\right]^2_{\mu_{\boldsymbol{X}}} \tag{3-25}$$

Thus, the total uncertainty in the output, expressed as variance, is explicitly decomposed as the sum of contributions of each input parameter, each contribution being a mix of the variance of this input parameter, $Var(X_i)$ and the squares of the sensitivity coefficients for the output Y with respect to this input parameter, $(\frac{\partial g(\boldsymbol{X})}{\partial X_i})^2$. Equation 3-25 is the basis of many techniques for error analysis widely used in the physical sciences and engineering (Morgan and Henrion, 1990).

In this work, the variance of each input variable (i.e., concentration in each grid cell) is assumed to be introduced by interpolation. Kriging Standard Deviation (SD) is used for quantifying the interpolation uncertainties (section 3.3). The sensitivity of the health metric with respect to different inputs is also determined by adjoint sensitivity analysis (section 3.4). Equation 3-25 allows us to examine which uncertain input concentrations are the main contributors to uncertainty in the health metric. It also presents how the health metric variance will decrease for a given reduction in the concentration variance of one or more grid cells. The output variance will mostly decline from a reduction of the input variance that largely contributes to the uncertainty in model output. However, this may not correspond to the grid concentration with the largest variance because the model will display different sensitivities to each uncertain input.

It should be noted that Equation 3-25 is based on the assumption of linearity in g and zero correlation between input uncertainties. In some applications in air quality, where these assumptions are deemed unacceptable, more advanced analysis techniques, often based on Monte Carlo simulation, may be employed (Moore and Londergan, 2001; Borrego et al., 2008).

3.6 Shannon information content

The objective of our network evaluation procedure is to determine the spatial station pattern that will result in maximum information return. The existing ozone monitoring network in the U.S. is examined in this work to determine the extent of information gain resulting from the addition of one or more monitoring stations. The best locations for such addition are also identified.

Metrics from information theory are used to quantify uncertainty-based information (i.e., decrease in uncertainty) when the system state is known. Some of the information metrics are employed to measure the decrease in the (co-)variance of the error including the Shannon information, the trace of the Fisher information matrix, and the degrees of freedom for signal). These three types of information metrics were reviewed briefly in

section 2.3. We represent the information content of observations through the measures of Shannon Information Content.

As previously discussed in section 2.3.2, the Shannon information is related to the scaling of the Gaussian probability densities. It can be estimated via Equation 2-5 by constructing both the prior and posterior covariance matrices. Around each grid cell, a region is defined where ozone concentrations are correlated and the prior covariance matrix is generated for each region before adding a new monitor. Center of grid cells are nominated as locations for new stations, therefore the posterior covariance matrices are calculated for the regions after addition of new monitors at the centers.

For the neighborhood of each grid cell, the prior covariance matrix of the predicted (i.e., kriged) ozone data is constructed firstly. The predicted concentrations result from interpolation of the existing ozone observational data.

If maximum 8-hr ozone concentration is considered to be a $(n \times 1)$ random vector, with n independent pieces of data (number of grid cells in each region), $X_1, X_2, \dots, X_n$, then the covariance matrix $COV[X]$ is a $(n \times n)$ matrix whose (i , j) -th entry is the covariance between X_i and X_j:

$$\text{cov}_{ij} = \text{cov}[X_i, X_j] = E[(X_i - \mu_i)(X_j - \mu_j)] \tag{3-26}$$

where:

$$\mu_i = E(X_i) \tag{3-27}$$

is the expected value of the ith entry in the vector X over ozone season (90 days). In other words we have:

$$COV[X] = E\begin{bmatrix} (X_1 - \mu_1)(X_1 - \mu_1) & (X_1 - \mu_1)(X_2 - \mu_2) & \dots & (X_1 - \mu_1)(X_n - \mu_n) \\ (X_2 - \mu_2)(X_1 - \mu_1) & (X_2 - \mu_2)(X_2 - \mu_2) & \dots & (X_2 - \mu_2)(X_n - \mu_n) \\ \vdots & \vdots & \ddots & \vdots \\ (X_n - \mu_n)(X_1 - \mu_1) & (X_n - \mu_n)(X_2 - \mu_2) & \dots & (X_n - \mu_n)(X_n - \mu_n) \end{bmatrix}.$$

Since $\mathrm{cov}[X_i, X_j] = \mathrm{var}(X_i)$ when $i = j$, the diagonal entries of the covariance matrix are equal to the variances of the individual components of X. Therefore, $COV[X]$ is simplified to (Wasserman, 2004):

$$COV[X] = \begin{bmatrix} \mathrm{var}[X_1] & \dots & \mathrm{cov}[X_1, X_n] \\ \vdots & \ddots & \vdots \\ \mathrm{cov}[X_n, X_1] & \dots & \mathrm{var}[X_n] \end{bmatrix} \tag{3-28}$$

Similarly, the posterior covariance matrices are computed for the area around the grid cells where the new monitors are added. As ozone concentrations are not available at the new points, CMAQ model results are utilized as measurement data in these grid cells. Afterwards, kriging method is employed again by the new set of observations (previous monitors plus the new one) and the autocorrelation in ozone concentration data of different grid cells in the region can be estimated. Finally, the posterior covariance of domain is constructed for each candidate monitoring station.

By having the prior and posterior covariance matrices, the information gained is calculated at each of the nodes from Equation 2-5. The nodal gained information is ranked from highest to lowest, with the highest chosen as the optimum location for network addition.

4 Chapter: Results and Discussion

4.1 Overview

As discussed earlier, the main objective of this study is to determine the locations of new ozone monitoring stations to result in the greatest information gain in the network with regards to policy relevant metrics such as mortality .To accomplish this, uncertainty analysis has been employed to evaluate reduction in uncertainty of ozone-associated mortality rates resulting from the addition of new monitors at different locations.

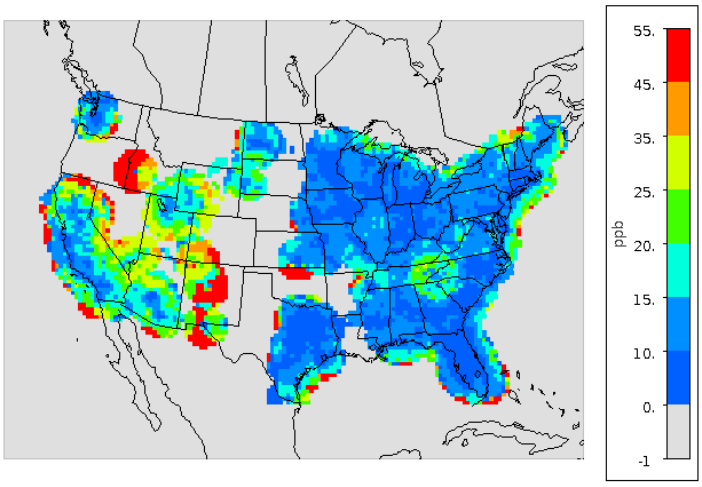
In the first step of uncertainty analysis, uncertainties associated with the predicted ozone concentrations have been quantified. In other words, the kriging variance field for the existing network is generated. In the second step, an estimate of the monetary value of changes in mortality risk is calculated. Sensitivity analyses of mortality risk reduction benefits from ozone reductions have been carried out using the Value of Statistical Life (VSL). In the last step, ozone input uncertainties are transformed to uncertainties in monetized estimates for mortality risk reduction benefits associated with ozone reductions. Process of propagating uncertainty has been done through the perturbation method.

In addition, the change in information content resulting from the addition of monitoring stations, are evaluated by Shannon entropy. The nodal gained information is calculated and ranked from highest to lowest.

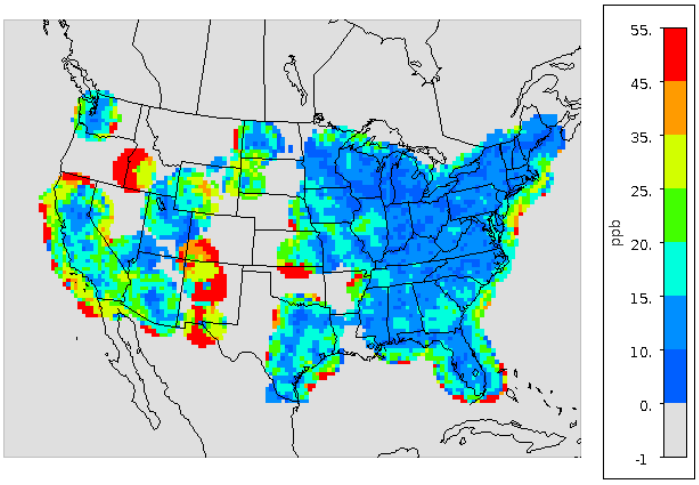
The previous chapter described the methods used for our analysis, while the results obtained from each step of uncertainty analysis and Shannon information metric are presented in this chapter. Geographic locations of the cities, which are used in this chapter, are shown in the U.S. map in Appendix A.

4.2 Interpolation uncertainty

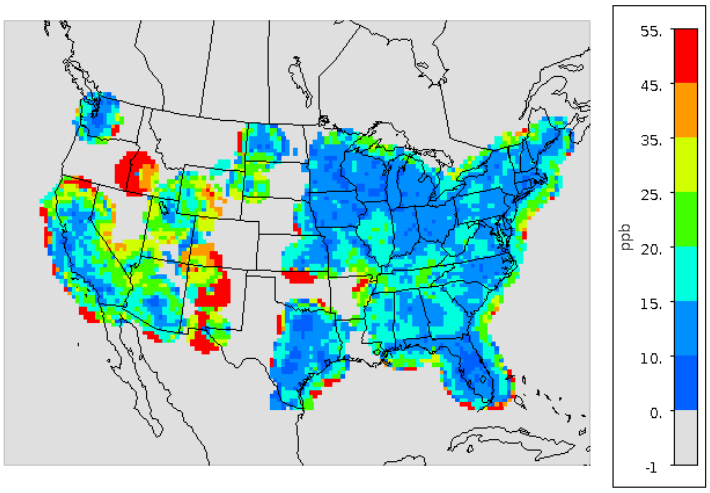
As discussed previously in section 3.3, spatial interpolation (e.g., kriging) is required to obtain ozone concentrations at each grid cell (i.e., each 36- by 36-km box). Estimate of



(a)



(b)



(c)

Figure 4-1: Standard deviation of ozone level at a) 5:00 b) 15:00 c) 21:00 GMT

the uncertainty in the predicted ozone estimates is based on the prediction errors produced by the kriging approach. The kriging variance field is generated for the existing network as a measure of estimated ozone uncertainties. The field standard deviations (the square root of kriging variance) has then been calculated and presented in ppb (Figure 4-1).

The average of standard deviations for kriging estimates is taken over July 1, 2007 to September 30, 2007 and shown for 3 different times of day (GMT). It is clearly visible from the plots that spatial distribution of SD varies at different hours of day (Figure 4-1a to 4-1c). The diurnal variation is considered later in section 4.5.

Kriging standard deviations reflect the uncertainty about station representativeness and the uncertainty of the monitoring measurement. By observing the contrast between Figure 4-1 and 4-2, we come to the conclusion that scarcity of monitors around measuring stations usually leads to higher SDs, i.e., less reliable kriging estimation. For example SD values in the cities of Albuquerque, NM, and Boise, ID are very high since they have only 7 and 3 stations in their radius of neighborhood respectively. In contrast, New York and Los Angeles have comparatively smaller standard deviations because of the abundant monitoring stations in their vicinity (Table 4-1).

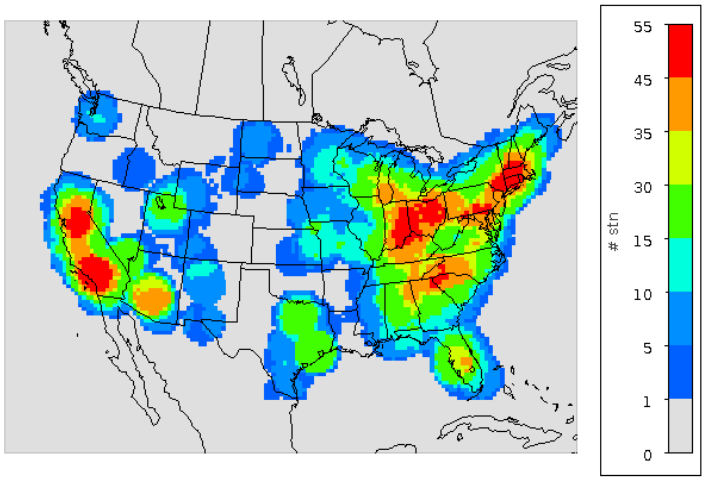


Figure 4-2: Number of stations in the radius of neighborhood of each grid

For the grid cells where no monitors exist in the surroundings, kriging method can predict neither the concentration nor the variance. There are no negative values for standard deviation and the range between -1 and 0 in grey color in Figure 4-1 is allocated to the cells with scarce monitors around.

In some locations, SD values are high even though enough monitoring stations exist in their radius of neighborhood (Figure 4-2). This can be due to non-optimal distribution of observations or high measurement errors of them. For instance, Charlotte, NC and Greenville, SC are two cities with relatively high ozone uncertainties in spite of the large number of stations in their neighborhood (35-40).

Table 4-1: Average daily standard deviation values at four different cities

City	# of stations	Radius of neighborhood (m)	SD (ppb)
Albuquerque, NM	7	21783.5	46
Boise, ID	3	20069.2	39
New York, NY	30	350271	5
Los Angeles, CA	50	359150	6

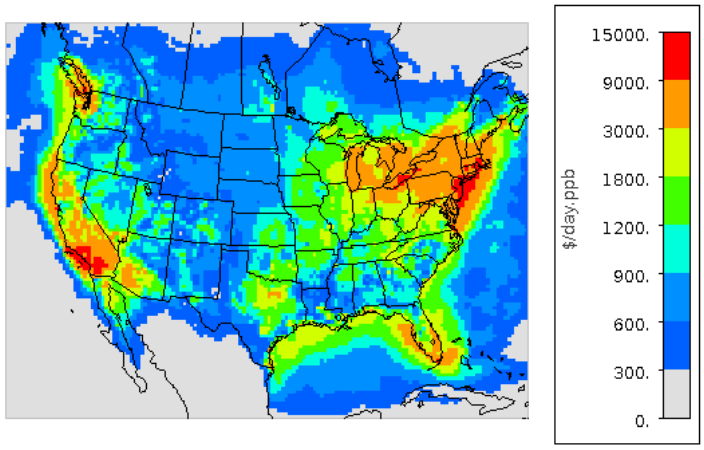
4.3 Sensitivity coefficient

The backward sensitivity analysis (adjoint of CMAQ) has been employed to estimate mortality risk reduction benefits from ozone reductions in each location or grid cell across the U.S. (Figure 4-3). Estimated health benefits are reported in ($/day) for one ppb change in ozone concentrations. For example, in Figure 4-3, a value of $1000/day.ppb in a grid cell indicates that a ppb reduction in ozone concentrations from that cell would benefit the U.S. by $1000/day in reduced mortality (Pappin and Hakami, 2013).

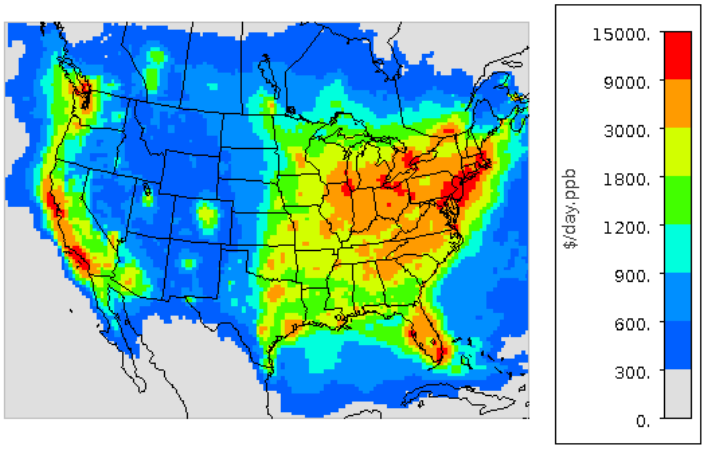
Sensitivity of mortality risk reduction benefits is averaged out over July 1, 2007 to September 30, 2007 and shown for 3 different times of day. Sensitivity coefficients present different spatial distributions at different hours of day (Figure 4-3a to 4-3c). Diurnal variation of sensitivity coefficients is discussed later in section 4.5.

In densely populated areas, there are many people affected by the air quality and it can be observed in Figure 4-3 that the nation-wide health benefits are more sensitive to changes in ozone concentrations in these locations. For example New York and Los Angeles have large influences, daily average at $12,630/day and $14,305/day for a ppb reduction in O_3 in each city, respectively. Albuquerque and Boise, on the other hand, are less sensitive to ozone reduction with low populations.

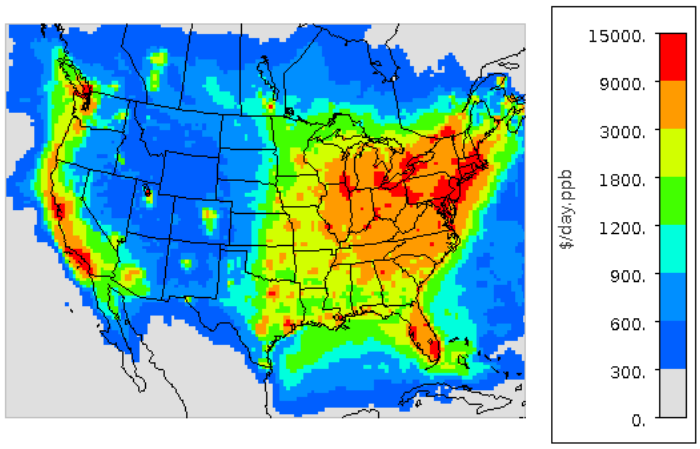
By sensitivity analysis, we determine the contribution of uncertain ozone concentrations to the overall uncertainty in health benefits. To reduce the uncertainty related to mortality risks the most and protect human health, it is reasonable to consider the contribution of higher sensitivity coefficients in our network design.



(a)



(b)



(c)

Figure 4-3: Mortality risk reduction benefits gained by ozone reductions at a) 5:00 b) 15:00 c) 21:00 GMT

4.4 Hourly propagated uncertainty

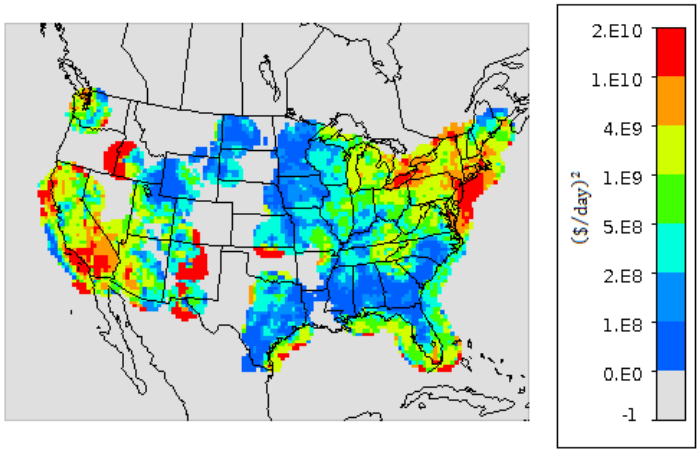
To evaluate improvement in monetized health benefit uncertainty by a reduction in ozone concentration variance of one or more grid cells, we have applied the perturbation method of uncertainty propagation based on Taylor series expansion. According to Equation 3-25, the propagated uncertainty in mortality risks consists of several hourly uncertainties (i.e., the propagated uncertainty terms) at different locations:

$$U_{hourly} = Var(C)\left(\frac{\partial M}{\partial C}\right)^2 = (SD(C))^2\left(\frac{\partial M}{\partial C}\right)^2 \quad (4\text{-}1)$$

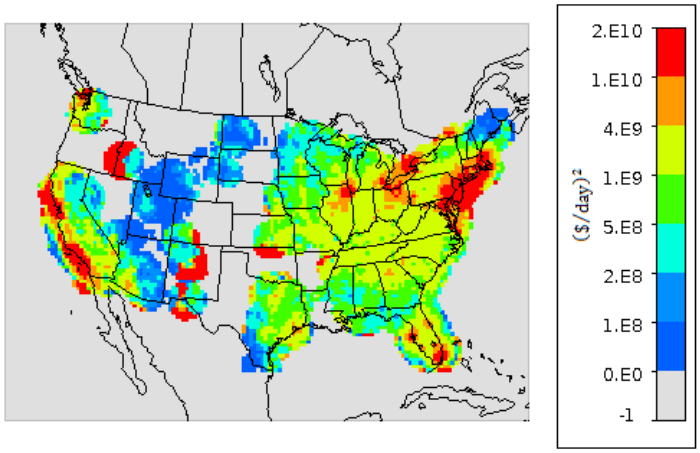
U_{hourly} is hourly uncertainty $(\$/day)^2$ and $\frac{\partial M}{\partial C}$ is sensitivity of health benefits with respect to reductions in ozone levels (obtained in section 4.3). $Var(C)$ is the variance of ozone concentrations and is equal to the square of standard deviation, $SD(C)$, which was presented in section 4.2.

High hourly uncertainties are the main contributors to health benefit uncertainty. Therefore, total uncertainty will mainly decline from a reduction in the concentration variance at locations with high hourly uncertainties. New York and Los Angeles are two examples of these locations with hourly uncertainties higher than 10^{10} $(\$/\text{day})^2$.

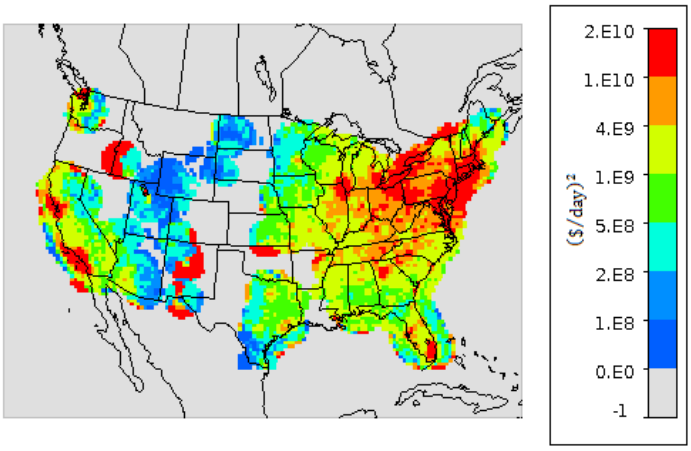
Figure 4-4 shows that high hourly uncertainty may not be necessarily related to the grid cell with large ozone uncertainty (i.e., SD value) because the model displays different sensitivities to each uncertain concentration. This fact can be examined by comparing two different cities of New York and Albuquerque in Figures 4-1, 4-3, and 4-4. Although New York has low estimated ozone uncertainty (at daily average of 5 ppb), propagated hourly uncertainty is high which is due to high sensitivity values in this city. On the other hand, large propagated uncertainty in Albuquerque is owing to high SD values (at daily average of 46 ppb). As a result, both ozone uncertainty and sensitivity play active roles in propagated uncertainty.



(a)



(b)

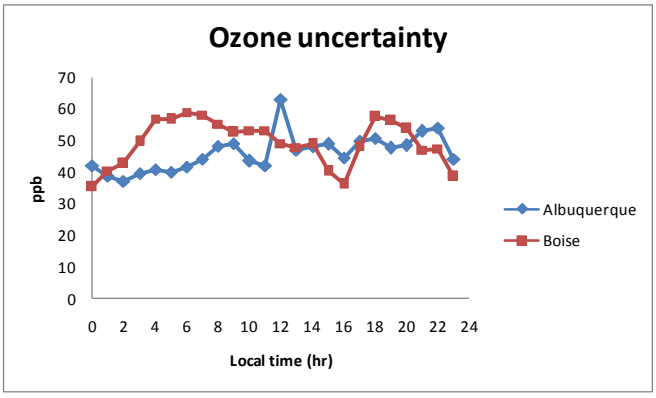


(c)

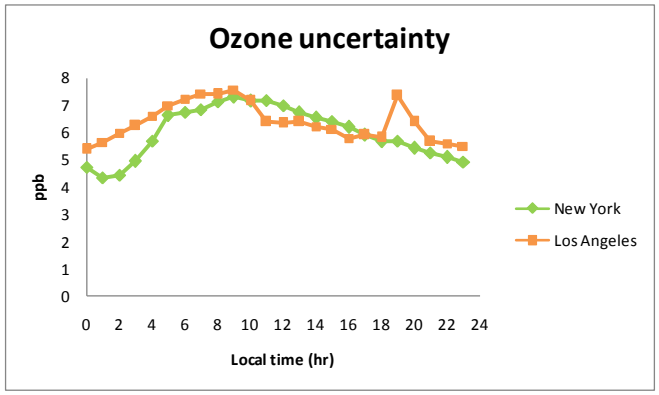
Figure 4-4: Hourly propagated uncertainty of mortality risk reduction benefits associated with ozone reductions at a) 5:00 b) 15:00 c) 21:00 GMT

4.5 Local propagated uncertainty

Figures 4-5 and 4-6 present diurnal variation of ozone uncertainty and sensitivity coefficient at four different cities in the study area. Significant hour-to-hour fluctuations in the figures are evident which can be due to the dependence of atmospheric pollutant transformation and transport on meteorological conditions and level of ozone concentration. Ozone concentrations also vary over different scales of time. Downwind of urban areas where ozone precursors are emitted, ozone concentrations tend to peak in the middle of the day and dip to their lowest concentrations during the middle of the night. A similar pattern is observed in rural areas except that the peak O_3 is delayed by several hours relative to the peak observed at urban areas.

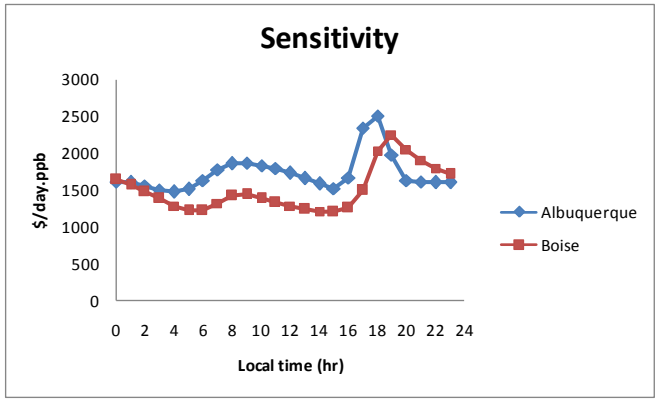


(a)

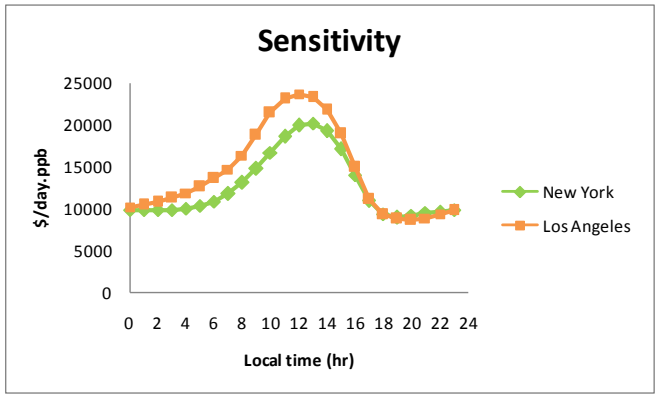


(b)

Figure 4-5: Diurnal variation of ozone standard deviation for four cities: a) Albuquerque & Boise b) New York & Los Angeles

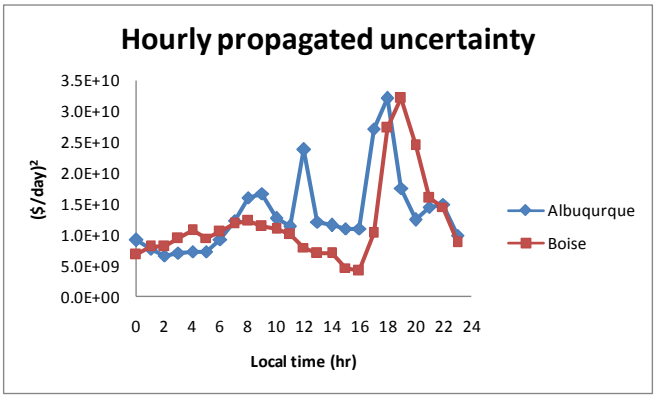


(a)

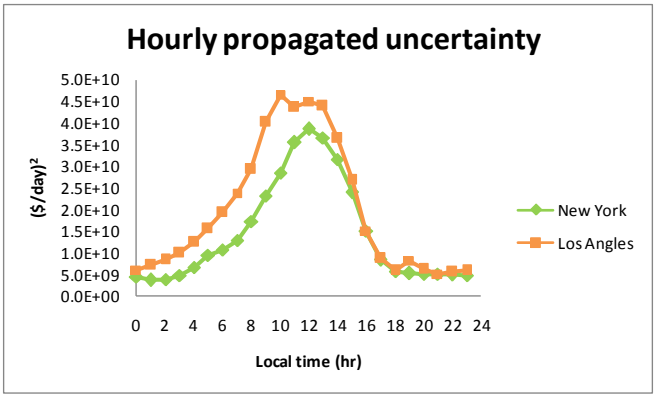


(b)

Figure 4-6: Diurnal variation of health benefits from ozone reductions for four cities: a) Albuquerque & Boise b) New York & Los Angeles



(a)



(b)

Figure 4-7: Diurnal variation of hourly propagated uncertainty for four cities: a) Albuquerque & Boise b) New York & Los Angeles

The changes of ozone uncertainty and sensitivity coefficient at different hours of a day lead to diurnal variation in hourly propagated uncertainty (Figure 4-7). In order to account for this variation, local propagated uncertainty is defined for each grid cell, which is derived by averaging hourly propagated uncertainties over a day:

$$U_{local} = \frac{1}{24}\sum_{t=0}^{23} Var(C^t)\left(\frac{\partial M}{\partial C^t}\right)^2 \qquad (4\text{-}2)$$

C^t indicates ozone concentration at time t of the day. Local propagated uncertainty of health benefits associated with ozone reductions, U_{local}, is presented in $(\$/\text{day})^2$ in Figure 4-8.

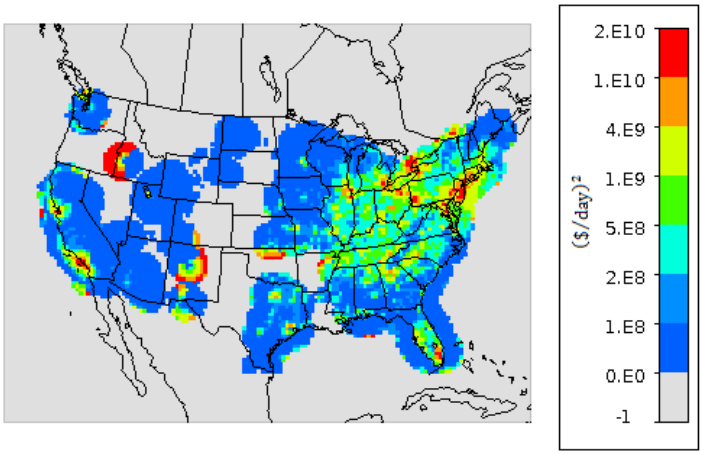


Figure 4-8: Local propagated uncertainty of mortality risk reduction benefits associated with ozone reductions

To have the uncertainty of monetary value estimate of mortality risks for the interested region given the monitoring network, overall propagated uncertainty can be defined by adding up local propagated uncertainty over the entire domain:

$$U_{overall} = \sum_{i,j} U_{local}(i,j) \tag{4-3}$$

By adding a new station in a grid cell, the kriging estimated uncertainty at that grid cell and those cells locating in its neighborhood will reduce. This causes reduction in the overall uncertainty of health benefits:

$$\Delta U_{overall} = \sum_{i=1}^{n} ((U_{local})_{prior} - (U_{local})_{posterior})$$

$$= \left(U_{local} - \frac{1}{T}\sum_{t=1}^{T} Var(meas)\left(\frac{\partial M}{\partial c^t}\right)^2\right) + \sum_{i=2}^{n}((U_{local})_{prior} - (U_{local})_{posterior}) \tag{4-4}$$

where n is the number of stations in the grid radius of neighborhood including the new station in the grid center. Subscripts *prior* and *posterior* refer to the states before and after adding a station.

Addition of a monitor causes reduction in kriging uncertainty of ozone level from the prior value to measurement error at the monitoring location (5 ppb). In addition,

uncertainty of ozone will improve from the prior to posterior values in the other grid cells situated in the radius of neighborhood of the new monitor. Consequently, the uncertainty of health benefit will decrease from local propagated uncertainty to sensitivity coefficient multiplied by measurement uncertainty at the grid point of interest (the first term in Equation 4-4). The second order term is also related to the contribution of other grid cells in decline of health benefit uncertainty.

The effect of adding a monitor at each grid cell on overall propagated uncertainty has been calculated through Equation 4-4 and is shown in Figure 4-9. Overall propagated uncertainties are presented in $ per day. For instance, in Figure 4-9, a value of $1000/day in a grid cell indicates that by adding a new station at that cell, the uncertainty of health benefit estimates would improve by $1000/day. If $1000/day is significant in comparison with overall mortality risk reduction benefits, it will be worth adding a station at that location. Otherwise, the benefits of adding a monitor are outweighed by costs of a new station.

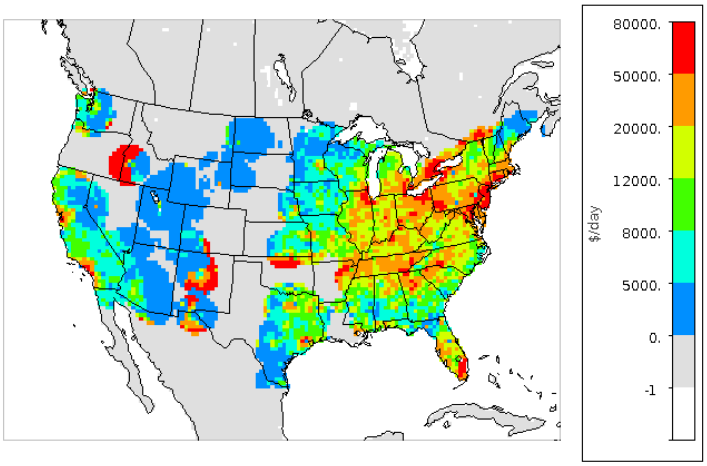


Figure 4-9: Contribution of a new monitor at each grid in reducing the overall uncertainty in mortality risk reduction benefits associated with ozone reductions

In spite of the fact that our ozone observations are limited to the U.S., adding a new monitoring station in some regions in Canada has a contribution in reducing the overall uncertainty in health benefits (Figure 4-9). This contribution is due to the kriging approach which considers a radius of neighborhood around each candidate location.

Therefore, adding a new monitor in Canada near the U.S. border can have a role in reducing the overall uncertainty.

4.6 Shannon information

Spatial variation of ozone concentrations over the domain of interest has been computed by kriging method. As a result, the ozone data is reconstructed at different locations and the Shannon entropy of the existing monitoring network is derived.

Inclusion of additional monitors in places where ozone cannot be adequately inferred from the existing monitors leads to decrease in Shannon entropy (i.e., increase in the information content) of the monitoring network. The extent of reduction depends on the location of adding a monitor. Therefore, each grid cell (grid center) is examined as a potential monitoring site and the entropy of the network after adding a monitor is calculated. The change in entropy of the network before and after adding a monitor reveals the ozone information added by each of these candidate locations.

Figure 4-10 shows the information gain from a new monitor at different grid cells over the U.S. region. Changes in entropy (information content) are described in nat which is a logarithmic unit of information (i.e., base e in Equation 2-1). In Figure 4-10, a value of 1 nat in a grid cell shows that addition of a new station would increase the information content or decrease the entropy (uncertainty) of the network by $\log_2 \mathrm{e}\ (\approx 1.443)$ bits. Locations which make significant contributions in the information increase of the network, such as west of Colorado & east of Kansas State are the most suitable candidates for new monitoring stations (from the viewpoint of Shannon information).

It is impractical to evaluate covariance matrix for the grid cells with small number of monitors in their surrounding since there is no estimated kriging concentration in these cells and/or their neighborhood. Hence, the increase in information due to the placement of an additional monitor cannot be evaluated for these grid cells (i.e., grey color in Figure 4-10).

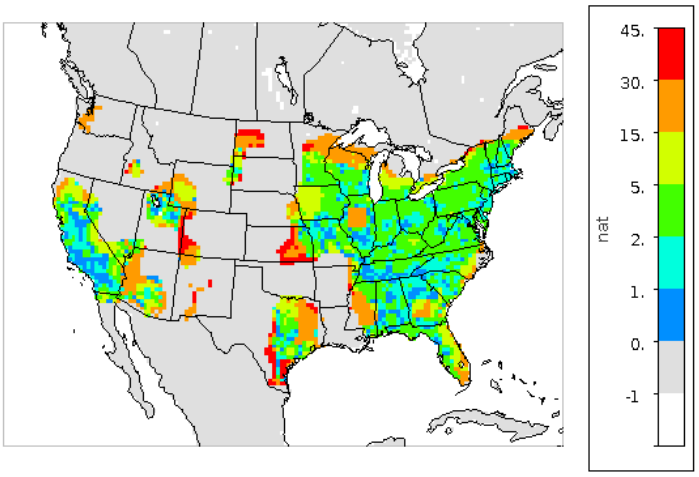


Figure 4-10: Contribution of a new monitor at each grid in increasing the information content of the monitoring network

Comparing Figure 4-8 and 4-10 shows that additional ozone monitoring locations suggested by Shannon entropy differ from those suggested by uncertainty analysis. Uncertainty analysis focuses on the locations that contribute significantly in reducing the uncertainty of health benefits associated with ozone reductions; something that requires both high ozone uncertainty and sensitivity coefficients. However, locations of interest based on Shannon entropy are those which decrease the entropy (uncertainty) maximally without accounting for the sensitivity of health benefits.

5 Chapter: Conclusions and Future work

"Neither the road ends here nor is it the destination, keep walking my friend the journey has just begun…"

Protection of human health and the environment from pollutant effects is the primary goal of air pollution control programs. In order to evaluate the effects caused by air pollution sources, air quality monitoring network is essential to monitor and control the atmospheric pollution. Determination of location and number of monitoring stations is among the most important factors to be taken into account when designing these networks. The network design objective can be to provide maximum information about the air quality for a given area. In view of the fact that monitoring costs including installation, maintenance, and management costs are substantial and ever-increasing, optimal network design plays a key role in air quality management programs.

This thesis has investigated methods for identifying the best locations for addition of monitoring stations to maximize information about the ozone level and mortality risk reduction benefits associated with ozone reductions over the continental U.S. region. The following key findings have been made:

- Uncertainty analysis is practical in providing the uncertainty associated with health benefits from reduced O_3 short-term mortality in this region. It involves the propagation of ozone uncertainties to uncertainties in health benefits metrics.

- Adjoint of CMAQ is a powerful tool in calculating sensitivity of health benefits with respect to ozone reductions. Spatial variability in sensitivity coefficients shows different contribution of uncertain ozone concentrations to the overall uncertainty in health benefits. The largest sensitivities are mostly related to heavily populated areas such as New York and Los Angeles.

- Kriging interpolation is effective in predicting the ozone concentrations and their associated uncertainty (i.e., kriging variance) at non-monitored locations. As expected, the spatial scarcity of monitors over long distances leads to higher ozone uncertainties. This point is observed in two cities of Albuquerque and Boise with scarce monitors in their vicinity. On the other hand, large number of

stations around a location does not guarantee low ozone uncertainties; Charlotte and Greenville are examples of two cities of this type.

- Diurnal variations in ozone uncertainty and sensitivity coefficients dictate that all hourly health benefits should be considered and combined into one metric for each location. Local propagated uncertainty is defined for this purpose.

- The effects of adding a monitor at different grid cells on the overall uncertainty of health benefits metric reveals the optimum locations for monitoring stations. In locations where the uncertainty contribution of a monitor is significant in comparison to the overall mortality risk reduction benefits, addition of a new monitor is considered worthwhile. Otherwise, the benefits of adding a new monitor may be outweighed by costs of a new station.

- Shannon entropy provides a measure of the change in information content resulting from the addition of one or more monitoring stations at different locations. This approach prioritizes the potential sites based on the information gained about the state of the atmosphere from each of them, as opposed to information regarding the health metric.

- Both uncertainty analysis and Shannon entropy suggest locations for adding new monitors. As expected, these locations are not the same. Uncertainty analysis attempts to minimize uncertainty (i.e., maximize uncertainty-based information) of health benefits, while Shannon entropy focuses on maximizing the information content of the network (just about ozone levels) without considering health benefits of ozone reductions.

It should be noted that this research has been conducted to demonstrate the effectiveness of uncertainty analysis in quantifying the information content of monitoring networks. Therefore, the results should be regarded as a proof-of-concept and no policy decisions are intended to be made from them. More conclusive quantification of information requires further research with high-resolution, multi-year, and multi-pollutant simulations with adequate consideration for uncertainties.

The limitations and potential topics for future research that can improve our understanding of monitoring network design are discussed further below:

- Extend the research to a longer period. As the study period is limited to O_3 season in 2007, it does not capture inter-annual variability in ozone estimated uncertainty and health benefits.

- Evaluate the assumption of isotropy for ozone concentrations. The covariance functions used in this study for kriging are developed under the hypotheses of homogeneity and isotropy. It means that we assume that the covariance between the ozone concentrations depend only on distance. This assumption can be used for short ranges; however, anisotropy of ozone concentrations tend to occur at large ranges and the wind direction plays an important role (i.e., covariance is a function of distance and direction). Directional examination by small number of sample sites makes it possible to check the validity of isotropic assumption.

- Consider non-stationarity assumption for covariance matrix. In stationary processes the mean and covariance do not vary with respect to time. We construct the prior and posterior covariance matrices for calculating Shannon information based on the assumption that the time series are stationary. This assumption can be justified by the fact that maximum ozone concentrations do not change dramatically over the ozone season. To have a more realistic assumption, non-stationarity should be investigated which involves more elaborate computations.

- Combine the information content of ozone and its precursors, NOx and VOC monitoring networks. As there is a likely association between NOx and VOC emissions, and ozone concentrations, their contribution in ozone related health effects should be considered. Thus, the effect of new NOx and VOC monitors in the information gained about health benefits of ozone reductions is worth taking into account.

- Consider elimination and relocation of monitoring stations. The efficiency of the existing monitoring network should be evaluated and the result of this assessment may include a redefinition of the size and scope of the network which can lead to

the elimination (due to redundancy or uselessness of the collected data) or relocation of some stations.

- Account for uncertainties associated with sensitivity estimates. Sensitivity coefficients are affected by various uncertainties in epidemiological values, mortality valuation, and atmospheric modeling (e.g., representation of complex atmospheric chemistry).

- Include morbidity and long-term mortality. Without consideration of morbidity and long-term mortality in sensitivity analysis, health benefits are underestimated.

- Expand the research work with cost-benefit analysis. Costs of a new station (installation, maintenance, and management costs) should be considered and compared with benefits of its information gain.

Appendices

Appendix A

In this appendix, we present the U.S. map with locations of the cities which were mentioned in this work.

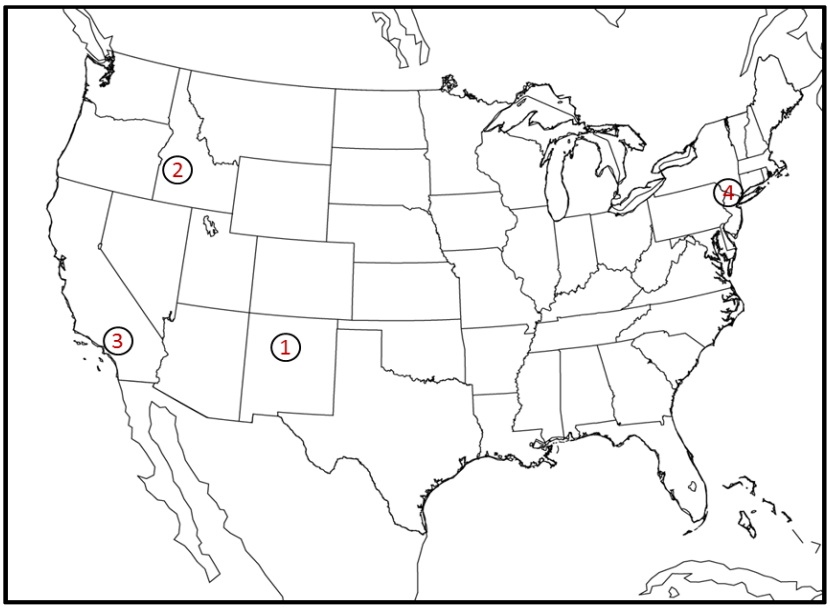


Figure A-1: Locations of some cities in the U.S.: 1) Albuquerque, 2) Boise, 3) Los Angeles, and 4) New York

References

Ainslie, B., Reuten, C., Steyn, D. G., Le, N. D., & Zidek, J. V. (2009). Application of an entropy-based Bayesian optimization technique to the redesign of an existing monitoring network for single air pollutants. Journal of environmental management, 90(8), 2715-2729.

Alberini, A., Tonin, S., Turvani, M., & Chiabai, A. (2006). Paying for permanence: Public preferences for contaminated site cleanup. Milano: Fondazione Eni Enrico Mattei.

Anderson, H. R., Atkinson, R. W., Peacock, J. L., Marston, L., & Konstantinou, K. (2004). Meta-analysis of time-series studies and panel studies of particulate matter (PM) and ozone (O_3). Report of a WHO Task Group. Copenhagen: World Health Organization.

Ayyub, B. M., & Klir, G. J. (2006). Uncertainty modeling and analysis in engineering and the sciences. Boca Raton, FL: Chapman & Hall/CRC.

Baldauf, R. W., Lane, D. D., & Marote, G. A. (2002). Ambient air quality monitoring network design for assessing human health impacts from exposures to airborne contaminants. Environmental monitoring and assessment, 66(1), 63-76.

Bell, M. L., McDermott, A., Zeger, S. L., Samet, J. M., & Dominici, F. (2004). Ozone and short-term mortality in 95 US urban communities, 1987-2000. Jama, 292(19), 2372-2378.

Bonn, G. (2003). Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide. Copenhagen: WHO Regional Office for Europe.

Borrego, C., Monteiro, A., Ferreira, J., Miranda, A. I., Costa, A. M., Carvalho, A. C., & Lopes, M. (2008). Procedures for estimation of modelling uncertainty in air quality assessment. Environment International, 34(5), 613-620.

Brown, J. D., & Heuvelink, G. (2005). Assessing uncertainty propagation through physically based models of soil water flow and solute transport. Encyclopedia of hydrological sciences.

Buell, C. E. (1975). Objective procedures for optimum location of air pollution observation stations. US Environmental Protection Agency, Office of Research and Development, Environmental Sciences Research Laboratory.

Burnett, R. T., Brook, J. R., Yung, W. T., Dales, R. E., & Krewski, D. (1997). Association between ozone and hospitalization for respiratory diseases in 16 Canadian cities. Environmental research, 72(1), 24-31.

Burrough, P. A., McDonnell, R., Burrough, P. A., & McDonnell, R. (1998). Principles of geographical information systems (Vol. 333). Oxford: Oxford university press.

Byun, D., & Schere, K. L. (2006). Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. Applied Mechanics Reviews, 59(2), 51-77.

Cacuci, D. G., Ionescu-Bujor, M., & Navon, I. M. (2005). Sensitivity and Uncertainty Analysis: Applications to large-scale systems (Vol. 2). CRC Press.

Cameron, K., & Hunter, P. (2002). Using spatial models and kriging techniques to optimize long-term ground-water monitoring networks: a case study. Environmetrics, 13(5-6), 629-656.

Cardelino, C. A., & Chameides, W. L. (1995). An observation-based model for analyzing ozone precursor relationships in the urban atmosphere. Journal of the Air & Waste Management Association, 45(3), 161-180.

Caselton, W. F., Kan, L., & Zidek, J. V. (1992). Quality data networks that minimize entropy. Statistics in the environmental and earth sciences, 10-38.

Caselton, W. F., & Zidek, J. V. (1984). Optimal monitoring network designs. Statistics & Probability Letters, 2(4), 223-227.

CEP. (2012). Community Multiscale Air Quality (CMAQ) v5.0 User's Manual. Environment, Carolina Environmental program, Chapel hill, NC.

CEP. (2009). Sparse Matrix Operator Kernel Emissions Modeling System (SMOKE) v2.4 User's Manual. Environment, Carolina Environmental program, Chapel hill, NC.

Chappelka, A. H., & Samuelson, L. J. (1998). Ambient ozone effects on forest trees of the eastern United States: a review. New Phytologist, 139(1), 91-108.

Cohan, D.S., Hakami, A., Hu, Y. & Russell, A.G. (2005). Nonlinear response of ozone to emissions: source apportionment and sensitivity analysis. Environmental science & technology, 39, 17, pp. 6739-6748.

Colella, P., & Woodward, P. R. (1984). The piecewise parabolic method (PPM) for gas-dynamical simulations. Journal of computational physics, 54(1), 174-201.

Cressie, N. A., & Cassie, N. A. (1993). Statistics for spatial data (Vol. 900). New York: Wiley.

CRS 2010, McCarthy. J. E. (2010). Ozone Air Quality Standards: EPA's Proposed. January 2010 Revisions, Congressional Research Service, 7-5700, R41062, from www.crs.gov

Daescu, D., Carmichael, G. R., & Sandu, A. (2000). Adjoint implementation of Rosenbrock methods applied to variational data assimilation problems. Journal of Computational Physics, 165(2), 496-510.

Damian, V., Sandu, A., Damian, M., Potra, F., & Carmichael, G. R. (2002). The kinetic preprocessor KPP-a software environment for solving chemical kinetics. Computers & Chemical Engineering, 26(11), 1567-1579.

De Smith, M. J., Goodchild, M. F., & Longley, P. (2007). Geospatial analysis: a comprehensive guide to principles, techniques and software tools. Leicester: Matador.
Diem, J. E., & Comrie, A. C. (2002). Predictive mapping of air pollution involving sparse spatial observations. Environmental pollution, 119(1), 99-117.

Diem, J. E., & Comrie, A. C. (2002). Predictive mapping of air pollution involving sparse spatial observations. Environmental pollution, 119(1), 99-117.

Dunker, A. M. (1984). The decoupled direct method for calculating sensitivity coefficients in chemical kinetics. The Journal o f Chemical Physics, 5/(5), 2385- 2393. doi: 10.1063/1.447938

Elbern, H. & Schmidt, H. (1999). A four-dimensional variational chemistry data assimilations scheme for Eulerian chemistry transport modeling, J. Geophys. Res., 104, 18583–18598.

Elbern, H., Schmidt, H., Talagrand, O., Ebel, A. (2000). 4D-variational data assimilation with an adjoint air quality model for emission analysis. Environ Model Software, 15, 539-548.

Elbern, H., Schmidt, H., & Ebel, A. (1997). Variational data assimilation for tropospheric chemistry modeling, J. Geophys. Res., 102, 15967–15985.

Environment Canada. (2012). Air Quality Health Index. from http://www.ec.gc.ca/cas-aqhi/default.asp?lang=En&n=CB0ADB16-1.

Environment Canada. (2000). Canada-wide standards for particulate matter and ozone. Canadian Council of Ministers of the Environment (CCME), June 5-6, 2000, Quebec City.

Farhat, N., Ramsay, T., Jerrett, M., & Krewski, D. (2013). Short-Term Effects of Ozone and PM2. 5 on Mortality in 12 Canadian Cities. Journal of Environmental Protection, 4, 18.

Fedorov, V., & Mueller, W. (1989). Comparison of two approaches in the optimal design of an observation network. Statistics: A Journal of Theoretical and Applied Statistics, 20(3), 339-351.

Fiore, A. M., Dentener, F. J., Wild, O., Cuvelier, C., Schultz, M. G., Hess, P. & Zuber, A. (2009). Multimodel estimates of intercontinental source-receptor relationships for ozone pollution. Journal of Geophysical Research: Atmospheres (1984–2012), 114(D4).

Fisher, M. (2003). Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems. ECMWF Technical Memoranda, 2003; 397.

Fisher, M. & Lary, D. J. (1995). Lagrangian four-dimensional variational data assimilation of chemical species, Quart. J. Roy. Meteorol. Soc., 121, 1681–1704.

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. Philosophical Transactions of the Royal Society of London, 1922; Series A, 222, 309-368

Fleming, S. W. (2007). An information theoretic perspective on mesoscale seasonal variations in ground-level ozone. Atmospheric Environment, 41(27), 5746-5755.

Gao, D., Stockwell, W. R., & Milford, J. B. (1995). First-order sensitivity and uncertainty analysis for a regional-scale gas-phase chemical mechanism. Journal of Geophysical Research: Atmospheres (1984–2012), 100(D11), 23153-23166.

Haas, T.C. (1992). Redesigning continental-scale monitoring networks. Atmospheric Environment 26A (18), 3323–3333.

Hakami, A., Henze, D. K., Seinfeld, J. H., Singh, K., Sandu, A., Kim, S., & Li, Q. (2007). The adjoint of CMAQ. Environmental science & technology, 41(22), 7807-7817.

Hakami, A., Henze, D. K., Seinfeld, J. H., Chai, T., Tang, Y., Carmichael, G. R., & Sandu, A. (2005). Adjoint inverse modeling of black carbon during the Asian Pacific Regional Aerosol Characterization Experiment, J. Geophys. Res., 110, D14301, doi:10.1029/2004JD005671.

Hakami, A., Odman, M. T., Russell, A. G. (2003). High-order, direct sensitivity analysis of multidimensional air quality models. Environmental Science and Technology, 37, pp. 2442-2452.

Hanna, S. R., Lu, Z., Christopher Frey, H., Wheeler, N., Vukovich, J., Arunachalam, S., & Alan Hansen, D. (2001). Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. Atmospheric Environment, 35(5), 891-903.

Harmanec, D. (1999). Measures of uncertainty and information. Imprecise Probability Project, from http://www.sipta.org/documentation/summary_measures/ippinfom.pdf

Heck, W. W., Taylor, O. C., Adams, R., Bingham, G., Miller, J., Preston, E., & Weinstein, L. (1982). Assessment of crop loss from ozone. Journal of the Air Pollution Control Association, 32(4), 353-361.

Hengl, T., Heuvelink, G., & Stein, A. (2004). A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma, 120(1), 75-93.

Henze, D. K., Hakami, A., & Seinfeld, J. H. (2007). Development of the adjoint of GEOS-Chem. Atmospheric Chemistry and Physics, 7(9), 2413-2433.

Holland, D. M., Saltzman, N., Cox, L. H., & Nychka, D. (1998). Spatial prediction of sulfur dioxide in the eastern United States. In geoENV II—Geostatistics for Environmental Applications (pp. 65-76). Springer Netherlands.

Hubbell, B. J., Hallberg, A., McCubbin, D. R., & Post, E. (2005). Health-related benefits of attaining the 8-hr ozone standard. Environmental Health Perspectives, 113(1), 73.

Husain, T., & Khan, S. M. (1983). Air monitoring network design using Fisher's information measures—A case study. Atmospheric Environment (1967), 17(12), 2591-2598.

Ito, K., De Leon, S. F., & Lippmann, M. (2005). Associations between ozone and daily mortality: analysis and meta-analysis. Epidemiology, 16(4), 446-457.

Jerrett, M., Burnett, R. T., Pope III, C. A., Ito, K., Thurston, G., Krewski, D., & Thun, M. (2009). Long-term ozone exposure and mortality. New England Journal of Medicine, 360(11), 1085-1095.

Jin, L., Tonse, S., Cohan, D.S., Mao, X., Harley, R.A. & Brown, N.J. (2008). Sensitivity analysis of ozone formation and transport for a Central California air pollution episode. Environmental Science and Technology, 42,10, pp. 3683- 3689.

Judek, S., Jessiman, B., Stieb, D., & Vet, R. (2005). Estimated number of excess deaths in Canada due to air pollution. document non publié.

Junninen, H., Niska, H., Tuppurainen, K., Ruuskanen, J., & Kolehmainen, M. (2004). Methods for imputation of missing values in air quality data sets. Atmospheric Environment, 38(18), 2895-2907.

Kanaroglou, P. S., Jerrett, M., Morrison, J., Beckerman, B., Arain, M. A., Gilbert, N. L., & Brook, J. R. (2004). Establishing an air pollution monitoring network for intra-urban population exposure assessment: A location-allocation approach. Atmospheric Environment, 39(13), 2399-2409.

Kim, S., Byun, D.W. & Cohan, D. (2009). Contributions of inter-and intra-state emissions to ozone over Dallas-Fort Worth, Texas. Civil Engineering and Environmental Systems, 26, no. 1, pp. 103-116.

Klir, G. J. (2006). Uncertainty and information foundations of generalized information theory. Hoboken, N.J.: Wiley-Interscience.

Lamb, R. G., Chen, W. H., & Seinfeld, J. H. (1975). Numerico-empirical analysis of atmospheric diffusion theories. J. Atmos. Sci., 32, 1794–1807.

Langstaff, J., Seigneur, C., Mei-Kao, L., Behar, J., & McElroy, J. L. (1986). Design of an optimum air monitoring network for exposure assessments. Atmospheric Environment (1967), 21(6), 1393-1410.

Le, N. D., & Zidek, J. V. (2006). Statistical analysis of environmental space-time processes. New York: Springer.

Lee, Y. M., & Ellis, J. H. (1997). On the equivalence of kriging and maximum entropy estimators. Mathematical geology, 29(1), 131-152.

Lefohn, A.S.; Simpson, J.; Knudsen, H.P.; Bhumralkar, C.; Logan, J.A. (1987). An evaluation of the Kriging method to predict 7-h seasonal mean ozone concentrations for estimating crop losses. J. Air Pollut. Contr. Assoc. 37: 595-602.

Levy, J. I., Chemerynski, S. M., & Sarnat, J. A. (2005). Ozone exposure and mortality: an empiric bayes metaregression analysis. Epidemiology, 16(4), 458-468.

Liu, L. J. S., & Rossini, A. J. (1996). Use of kriging models to predict 12-hour mean ozone concentrations in metropolitan Toronto—a pilot study. Environment International, 22(6), 677-692.

Lu, G. Y., & Wong, D. W. (2008). An adaptive inverse-distance weighting spatial interpolation technique. Computers & Geosciences, 34(9), 1044-1055.

Ludwig, F. L., Javitz, H. S., & Valdes, A. (1983). How Many Stations Are Required to Estimate the Design Value and the Expected Number of Exceedances of the zone Standard in an Urban Area?. Journal of the Air Pollution Control Association, 33(10), 963-967.

Manning, M., & Petit, M. (2003). A concept paper for the AR4 cross cutting theme: uncertainties and risk. Intergovernmental Panel on Climate Change Concept Paper, http://www. ipcc. ch/activity/cct1. pdf (last accessed on 31 July 2006), Geneva, Switzerland.

Marchuk, G. (1974). Numerical solution of the problems of the dynamics of the atmosphere and the ocean. (In Russian), Gidrometeoizdat.

Martien, P. T. & Harley, R. A. (2006). Adjoint sensitivity analysis for a three-dimensional photochemical model: Application to Southern California. Environ. Sci. Technol., 40, 4200–4210.

Matheron, G. (1969). Le Krigeage Universel: Fascicule 1. Cahiers du CMM., 82 p

Matheron, G. (1962). Traite de geostatistique appliqué. Tome I: Memoires du Bureau de Recherches Geologiques et Minieres, no. 14, Editions Technip, Paris, 333 p

Mazzeo, N., & Venegas, L. (2010). Development and application of a methodology for designing a multi-objective and multi-pollutant air quality monitoring network for urban areas. In: Air Quality, A. Kumar (Ed.), 23 47 , Sciyo, Rijeka, Croatia. www.sciyo.com

McConnell, R., Berhane, K., Gilliland, F., London, S. J., Islam, T., Gauderman, W. J., & Peters, J. M. (2002). Asthma in exercising children exposed to ozone: a cohort study. The Lancet, 359(9304), 386-391.

Meirink, J. F., Eskes, H. J., & Goede, A. P. H. (2006) Sensitivity analysis of methane emissions derived from SCIAMACHY observations through inverse modeling. Atmos. Chem. Phys., 6,1275–1292.

Menut, L., Vautard, R., Beekmann, M., & Honore, C. (2000). Sensitivity of photochemical pollution using the adjoint of a simplified chemistry-transport model. J. Geophys. Res., 105, 15379–15402.

Menut, L. (2003). Adjoint modeling for atmospheric pollution process sensitivity at regional scale. Journal of Geophysical Research

Milford, J. B., Gao, D., Russell, A. G., & McRae, G. J. (1992). Use of sensitivity analysis to compare chemical mechanisms for air-quality modeling. Environmental science & technology, 26(6), 1179-1189.

Morgan, M. G., & Henrion, M. (1990). Uncertainty: a Guide to dealing with uncertainty in quantitative risk and policy analysis Cambridge University Press. New York, New York, USA.

Moore, G. E., & Londergan, R. J. (2001). Sampled Monte Carlo uncertainty analysis for photochemical grid models. Atmospheric Environment, 35(28), 4863-4876.

Mulholland, J. A., Butler, A. J., Wilkinson, J. G., Russell, A. G., & Tolbert, P. E. (1998). Temporal and spatial distributions of ozone in Atlanta: regulatory and epidemiologic implications. Journal of the Air & Waste Management Association, 48(5), 418-426.

Nester, K. & Panitz, H. J. (2006). Sensitivity analysis by the adjoint chemistry transport model DRAIS for an episode in the Berlin Ozone (BERLIOZ) experiment. Atmos. Chem. Phys., 6, 2091–2106.

Noll, K. E., & Mitsutomi, S. (1983). Design methodology for optimum dosage air monitoring site selection. Atmospheric Environment (1967), 17(12), 2583-2590.

Nowak, D. J., McHale, P. J., Ibarra, M., Crane, D., Stevens, J. C., & Luley, C. J. (1998). Modeling the effects of urban vegetation on air pollution. In Air pollution modeling and its application XII (pp. 399-407). Springer US.

Pappin, A. J., & Hakami, A. (2013). Source attribution of health benefits from air pollution abatement in Canada and the United States: an adjoint sensitivity analysis. Environmental health perspectives, 121(5), 572.

Pickett, E. E., & Whiting, R. G. (1981). The design of cost-effective air quality monitoring networks. Environmental Monitoring and Assessment, 1(1), 59-74.

Rabitz, H., Kramer, M., & Dacol, D. (1983). Sensitivity analysis in chemical kinetics. Annual review of physical chemistry, 34(1), 419-461.

Robinson, L. A. (2008). Valuing the Health Impacts of Air Emissions. Prepared for DALYs, QALYs, and other Community Measures of Environment Health

Rodgers, C. D. (2000). Inverse methods for atmospheric sounding: Theory and practice (Vol. 2, p. 256). Singapore: World scientific.

Rouhani, S. (1985). Variance reduction analysis. Water Resources Research, 21(6), 837-846.

Sacks, J., & Schiller, S. (1988). Spatial designs. Statistical decision theory and related topics IV, 2(S S), 385-399.

Saltelli, A., Chan, K., & Scott, E. M. (Eds.). (2000). Sensitivity analysis (Vol. 134). New York: Wiley.

Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). Sensitivity analysis in practice: a guide to assessing scientific models. John Wiley & Sons.

Sandu, A., Singh, K., Jardak, M., Bowman, K., & Lee, M. (2012). A Practical Method to Estimate Information Content in the Context of 4D-Var Data Assimilation. SIAM/ASA Journal on Uncertainty Quantification, 106-138.

Sandu, A., & Chai, T. (2011). Chemical data assimilation—an overview. Atmosphere, 2(3), 426-463.

Sandu, A., & Zhang, L. (2008). Discrete second order adjoints in atmospheric chemical transport modeling. Journal of Computational Physics, 227(12), 5949-5983.

Sandu, A., Liao, W., Carmichael, G. R., Henze, D. K., & Seinfeld, J. H. (2005). Inverse modeling of aerosol dynamics using adjoints: Theoretical and numerical considerations, Aerosol Sci.Technol., 39, 677–694.

Schmidt, H. & Martin, D. (2003) Adjoint sensitivity of episodic ozone in the Paris area to emissions on the continental scale. Journal of Geophysical Research., 108, 8561–8577.

Seefeld, S., & Stockwell, W. R. (1999). First-order sensitivity analysis of models with time-dependent parameters: an application to PAN and ozone. Atmospheric Environment, 33(18), 2941-2953.

Seinfeld, J. H. (1972). Optimal location of pollutant monitoring stations in an airshed. Atmospheric Environment (1967), 6(11), 847-858.

Seinfeld, J. H., & Pandis, S. N. (2006). Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. Atmospheric Chemistry and Physics (Vol. 2nd, p. 1203). Wiley-Interscience.

Shannon, C. E. (1948). The mathematical theory of communication. The Bell System Technical Journal, 27(3&4), pp. 379–423, 623–656.

Silva, C., & Quiroz, A. (2003). Optimization of the atmospheric pollution monitoring network at Santiago de Chile. Atmospheric Environment, 37(17), 2337-2345.

Singh, V. P. (2013). Entropy theory and its application in environmental and water engineering. John Wiley & Sons.

Sluiter, R. (2008). Interpolation methods for climate data: literature review. De Bilt, Royal Netherlands Meteorological Institute(KNMI).

Sudret, B. (2007). Uncertainty propagation and sensitivity analysis in mechanical models–Contributions to structural reliability and stochastic spectral methods. Habilitationa diriger des recherches, Université Blaise Pascal, Clermont-Ferrand, France.

Sudret, B. & A. Der Kiureghian (2000). Stochastic finite elements and reliability: a state-of-the-art report. Technical Report UCB/SEMM-2000/08, University of California, Berkeley.

Tager, I. B., Balmes, J., Lurmann, F., Ngo, L., Alcorn, S., & Künzli, N. (2005). Chronic exposure to ambient ozone and lung function in young adults. Epidemiology, 16(6), 751-759.

Tan, W. N., Li, Z. A., Zou, B., & DING, Y. Z. (2005). The application of geostatistics to soil science. Tropical Geography, 25(4), 307.

Trujillo-Ventura, A., & Hugh Ellis, J. (1990). Multiobjective air pollution monitoring network design. Atmospheric Environment. Part A. General Topics, 25(2), 469-479.

U.S. EPA. (2014). Technology Transfer Network, Air Quality System. from http://www.epa.gov/ttn/airs/airsaqs/index.htm

U.S. EPA. (2013). Integrated Science Assessment of Ozone and Related Photochemical Oxidants (Final Report). EPA/600/R-10/076F. Washington, DC: U.S. Environmental Protection Agency.

U.S. EPA. (2010). Guidelines for Preparing Economic Analyses. EPA/240/R-10/001. Washington, DC: U.S. Environmental Protection Agency.

U.S. EPA. (2004). Developing Spatially Interpolated Surfaces and Estimating Uncertainty. EPA/454/R-04/004. Washington, DC: U.S. Environmental Protection Agency.

U.S. EPA. (1999). Ozone Monitoring, Mapping, and Public Outreach. EPA/625/R-99/007. Washington, DC: U.S. Environmental Protection Agency.

Vautard, R., Beekmann, M., & Menut, L. (2000). Applications of adjoint modelling in atmospheric chemistry: sensitivity and inverse modelling, Environ Modell Softw, 15, 703–709.

Venkatram, A. (1988). On the use of kriging in the spatial analysis of acid precipitation data. Atmospheric Environment (1967), 22(9), 1963-1975.

Vukicevic, T. & Hess, P. (2000) Analysis of troposhperic transport in the Pacific Basin using the adjoint technique, J. Geophys. Res., 105, 7213–7230.

Wackernagel, H. (2003). Multivariate Geostatistics: An Introduction with Applications. Springer, Berlin, 387 pp.

Wasserman, L. (2004). All of statistics: a concise course in statistical inference. New York: Springer.

Webster, R., & Oliver, M. A. (2001). Geostatistics for environmental scientists. Chichester, England: John Wiley & Sons.

World Health Organization (WHO). (2008). Health risks of ozone from long-range trans-boundary air pollution. World Health Organization, Regional Office for Europe.

Wu, S., & Zidek, J. V. (1992). An entropy-based analysis of data from selected NADP/NTN network sites for 1983–1986. Atmospheric Environment. Part A. General Topics, 26(11), 2089-2103.

Yang, Y. J., Wilkinson, J. G., & Russell, A. (1997). Fast, Direct Sensitivity Analysis of Multidimensional Photochemical Models. Environmental Science & Technology, 27(10), 2859-2868.

Zhang, X. (2002). Valuing mortality risk reductions using the contingent valuation methods: evidence from a survey of Beijing Residents in 1999. Centre for Environment and Development, Chinese Academy of Social Sciences, Beijing, China.

Zlatev, T. O. I. D. R. G. Z. (2011). Air pollution modelling, sensitivity analysis and parallel implementation. International Journal of Environment and Pollution, 46(1), 83-96.