12

Imprecise Uncertainty Modelling of Air Pollutant PM_{10}

Danni Guo[1], Renkuan Guo[2], Christien Thiart[2] and Yanhong Cui[2]
[1]*Climate Change and Bioadapation Division, South African National Biodiversity Institute, Cape Town*
[2]*Department of Statistical Sciences, University of Cape Town, Cape Town, South Africa*

1. Introduction

Particulate matter (PM) refers to solid particles and liquid droplets found in air. Many manmade and natural sources produce PM directly, or produce pollutants that react in the atmosphere to form PM. The resultant solid and liquid particles come in a wide range of sizes, and particles that are 10 micrometers or less in diameter (PM_{10}) can be inhaled into and accumulate in the respiratory system and are believed to pose health risks (Environmental Protection Agency, 2010). Particulate matter is one of the six primary air pollutants the Environmental Protection Agency (EPA) regulates, due to exposure to high outdoor PM_{10} concentrations causes increased disease and death (Environmental Protection Agency, 2010). Therefore, PM_{10} concentrations, amongst many other air pollutants, are sampled and measured in various places in California, United States.

The general trend of PM air pollutant concentrations in the air in California are on the decrease, but it continues to be monitored and observed. The California standards for annual PM_{10} concentrations is that the annual arithmetic mean is 20 $\mu g/m^3$, and the national standard is 50 $\mu g/m^3$ before 2006 (California Environmental Protection Agency Air Resources Board, 2010, Environmental Protection Agency, 2010). The State of California sets very high standards for their air quality, and air pollutants are carefully monitored.

However, in reality, it is too costly in terms of time, finance, and manpower to keep all the 213 sites to be monitoring and recording. In Fig. 1, a complete map of all 213 sample locations for PM_{10} are shown. However, one must note that these sample sites are never all used at any given year, PM_{10} samples are taken at different locations each year. At best, a maximum of 102 PM_{10} samples are collected during some years, and at worst, 61 PM_{10} samples are collected at that year. Therefore comparisons of PM_{10} between years are difficult, due to missing data at sample sites. It is difficult to construct kriging maps in terms of actual observations annually since the air pollutants were measured in different locations each year although the site design originally planned was quite delicate statistically.

Each year, approximately 40% of the 213 sites were actually observed. We call a site that does not have a recorded PM_{10} value as "missing value", and since there are no patterns so that serious problems would twist the kriging map constructions. In Fig. 2, this is clearly demonstrated. In 1989, there are 61 PM_{10} samples collected (29% of 213 locations), and in 2000, there are 94 PM_{10} samples collected (44%).

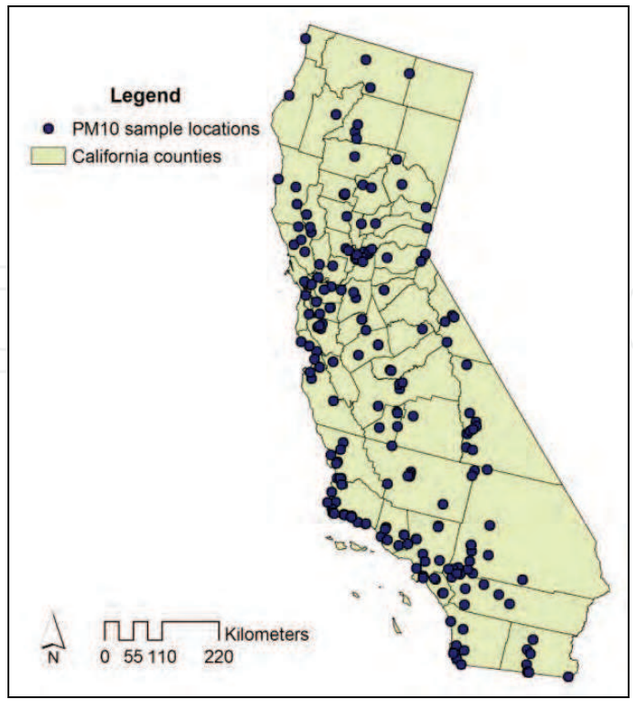


Fig. 1. Complete 213 Observational Sites in the California State

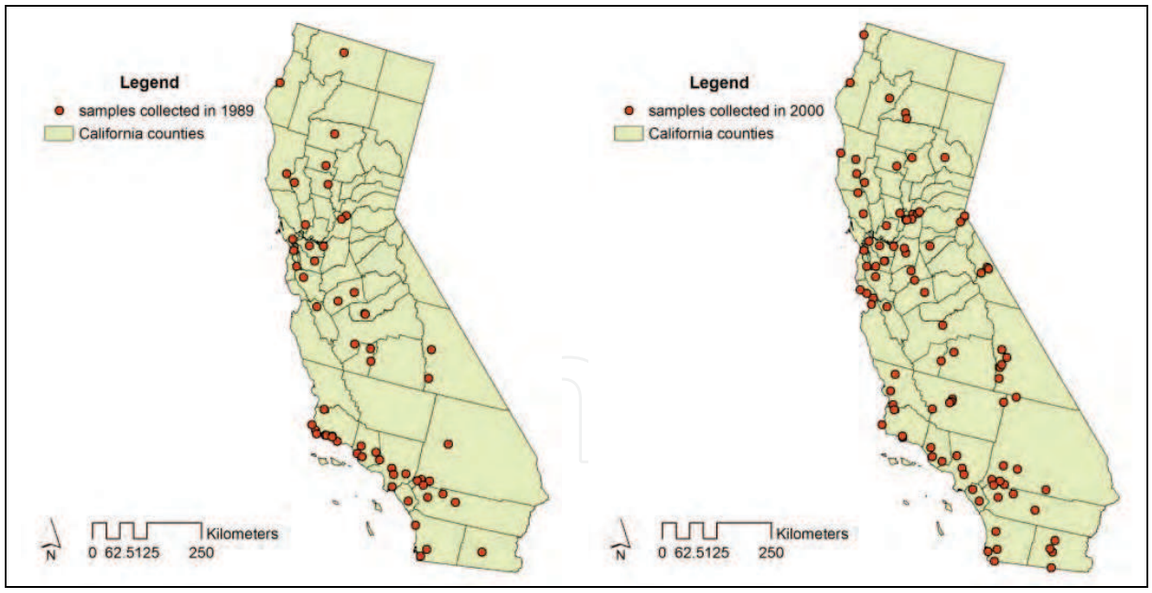


Fig. 2. PM_{10} Samples Collected in California in 1989 (61 sites) and 2000 (94 sites)

The data scarcity brings in a series of (five) fundamental issues into the spatial-temporal modelling and prediction practices for California PM_{10} data, namely:

1. The necessity to recognize the impreciseness in analyzing the spatial-temporal pattern in terms of California PM_{10} records, which inevitably acts the solidness of a geo-statistical analysis;

2. Which theoretical foundations are appropriate for modelling impreciseness uncertainty;
3. How to fill up the "missing value" sites so that the "complete" records are available, which is either an original annual average from the original observations (40%) recorded on the site or a or predicted value by "neighbourhood sites" (60%), i.e., to facilitate spatial-temporal imprecise PM_{10} value by interpolations and extrapolations;
4. How to estimate the parameters of uncertain processes (temporal patterns), particularly the rate of change parameter α_i, $i = 1,2,\cdots,213$;
5. Create annual kiging maps (19 maps) under spatially isotropy and stationarity assumptions so that the changes between annual maps can be analyzed by kriging map difference between 2007 and 1989 and kriging map of location rate of change.

These issues will be addressed in the remaining sections sequentially.

2. The necessity of modelling impreciseness in California PM_{10} spatial-temporal analysis

Impreciseness is a fundamental and intrinsic feature in the PM_{10} spatial-temporal modelling, due to the observational data shortage and incompleteness. Spatially, there are 213 sites involved, and temporally, PM_{10} observations were collected from 1989 to 2007, over a 19-year period. During the 19-year period, there are only two sites (Site 2125 and Site 2804) having complete 19 year records. There are 16 sites having only 1 record (8%) and 70 sites having 10 or above records. To have a statistically significant time-series analysis, 50 data points are minimal requirement for each site, so classical time-series analysis (probabilistic analysis) cannot be performed. In order to have a quick overall evaluation of PM_{10} records on each site, we borrow the statistical quality control idea here (Electric, 1956, Montgomery, 2001). But we do not carry on traditional 6-sigma rule, rather, classify the PM_{10} records into four groups: 1-$(5,20]$, 2-$(20,35]$, 3-$(35,50]$, 4-$(50,160]$. These four-group limits in Table 1 reflect the national standard, (50 $\mu g/m^3$) and California state standard (20 $\mu g/m^3$) respectively. For example, 1 -$(5,20]$ is for a location whose PM_{10} fall in 5 to 20 ($\mu g/m^3$).

County name	1-$(5,20]$	2-$(20,35]$	3-$(35,50]$	4-$(50,160]$	No. of Sites
Los Angeles			3	7	10
Kern	1	2	1	5	9
Riverside	3		3	3	9
San Diego			5	3	8
Imperial	3			4	7
Lake		3			3
Inyo	2	2	4		8

Table 1. PM_{10} Hazard level evaluation over selected 7 counties

One must be aware that the classification is not in absolute sense, rather, additional rules are adding (similar to quality control chart pattern analysis (Electric, 1956):
(1) if a single point, then, classify the site hazard level according to which group it falls in;
(2) if a sequence of records, some of them, particularly early points may fall in higher (or lower) hazard level, but if last three points fall in a lower (or higher) hazard level, the later level would be chosen for the site.

The additional rule 1 can attribute to expert's knowledge confirmation, while the additional rule 2 can be regarded as an expert's decision based on trend pattern.

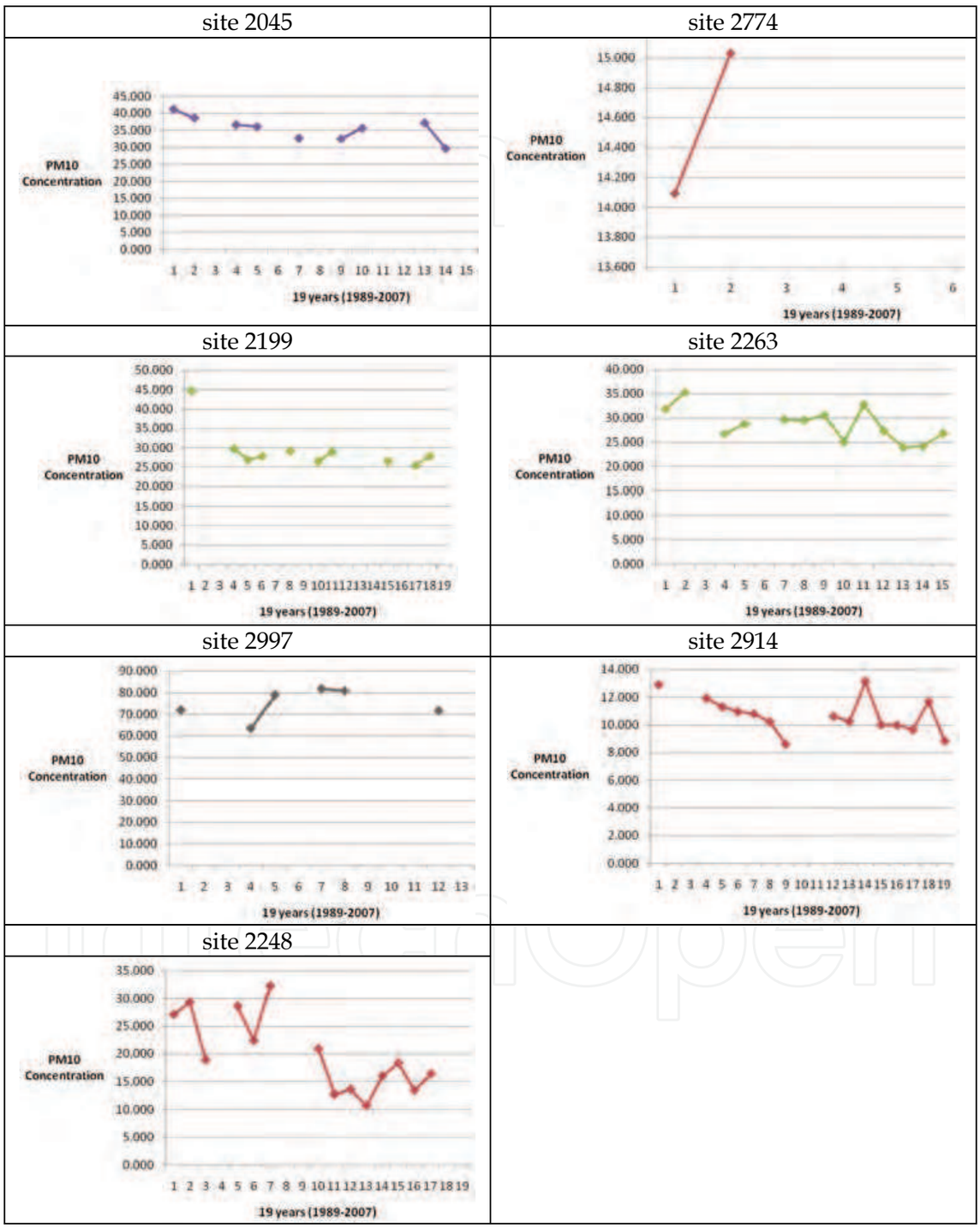


Fig. 3. The 7 sites from the selected 7 counties with original PM_{10} data plots and the hazard level classifications

Fig. 3 shows the classifications of a seven sites from the selected 7 counties in Table 1, each county one site is picked up for illustration purpose. The red coloured plot means the hazard level 1 $(5,20]$; the green coloured plot means the hazard level 2 $(20,35]$; the purple coloured plot means the hazard level 3 $(35,50]$; and the black coloured plot means the hazard level 4 $(50,160]$.

It is evident that facing the impreciseness caused by incomplete data recording, one has to rely on expert's knowledge to compensate the inadequacy and accuracy in collected observational data. Impreciseness is referred to a term with a connotation specified by an uncertain measure or an uncertainty distribution for each of the actual or hypothetical members of an uncertainty population (i.e., collection of expert's knowledge). An uncertain process is a repeating process whose outcomes follow no describable deterministic pattern, but follow an uncertainty distribution, such that the uncertain measure of the occurrence of each outcome can be only approximated or calculated.

The uncertainty modelling without a measure specification will not have an rigorous mathematical foundations and consequently the modelling exercise is baseless and blindness. In other words, measure specification is the prerequisite to spatial-temporal data collection and analysis. For example, without Kolmogrov's (1950) three axioms of probability measure, randomness is not defined and thus statistical data analysis and inference has no foundation at all.

Definition 2.1: Impreciseness is an intrinsic property of a variable or an expert's knowledge being specified by an uncertain measure.

It is therefore inevitably to seek appropriate form of uncertainty theory to meet the impreciseness challenges. In the theoretical basket, interval uncertainty theory (Moore, 1966), fuzzy theory (Zadeh, 1965, 1978), grey theory (Deng, 1984), rough set theory (1982), upper and lower provisions (or expectations) (Walley, 1991), or Liu's uncertainty theory (2007, 2010) may be chosen.

While imprecise probability theory (Utikin and Gurov, 1998) may be a typical answer to address the observational data inaccuracy and inadequacy. However the imprecise probability based spatial modelling requires too heavy assumptions. Just as Utikin and Gurov (2000) commented, "the probabilistic uncertainty model makes sense if the following three premises are satisfied: (i) an event is defined precisely; (ii) a large amount of statistical samples is available; (iii) probabilistic repetitiveness is embedded in the collected samples. This implies that the probabilistic assumption may be unreasonable in a wide scope of cases." Guo et al. (2007) and Guo (2010) did attempt to address the spatial uncertainty from the fuzzy logic and later Liu's (2007) credibility theory view of point.

Nevertheless, Liu's (2007, 2010) uncertainty theory is the only one built on an axiomatic uncertain measure foundation and fully justified with mathematical rigor. Therefore it is logical to engage Liu's (2007, 2010, 2011) uncertainty theory for guiding us to understand the intrinsic character of imprecise uncertainty and facilitate an accurate mathematical definition of impreciseness in order to establish the foundations for uncertainty spatial modelling under imprecise uncertainty environments.

3. Uncertain measure and uncertain calculus foundations

Uncertainty theory was founded by Liu in 2007 and refined in 2010, 2011. Nowadays uncertainty theory has become a branch of mathematics.

A key concept in uncertainty theory is the uncertain measure, which is a set function defined on a sigma-algebra generated from a non-empty set. Formally, let Ξ be a nonempty set (space), and $\mathfrak{A}(\Xi)$ the σ-algebra on Ξ. Each element, let us say, $A \subset \Xi$, $A \in \mathfrak{A}(\Xi)$ is called an uncertain event. A number denoted as $\lambda\{A\}$, $0 \le \lambda\{A\} \le 1$, is assigned to event $A \in \mathfrak{A}(\Xi)$, which indicates the uncertain measuring grade with which event $A \in \mathfrak{A}(\Xi)$ occurs. The normal set function $\lambda\{A\}$ satisfies following axioms given by Liu (2011):

Axiom 1: (Normality) $\lambda\{\Xi\} = 1$.

Axiom 2: (Self-Duality) $\lambda\{\cdot\}$ is self-dual, i.e., for any $A \in \mathfrak{A}(\Xi)$, $\lambda\{A\} + \lambda\{A^c\} = 1$.

Axiom 3: (σ- Subadditivity) $\lambda\left\{\bigcup_{i=1}^{\infty} A_i\right\} \le \sum_{i=1}^{\infty} \lambda\{A_i\}$ for any countable event sequence $\{A_i\}$.

Axiom 4: (Product Measure) Let $(\Xi_k, \mathfrak{A}_{\Xi_k}, \lambda_k)$ be the k^{th} uncertain space, $k = 1,2,\cdots,n$. Then product uncertain measure λ on the product measurable space $(\Xi, \mathfrak{A}_{\Xi})$ is defined by

$$\lambda = \lambda_1 \wedge \lambda_2 \wedge \cdots \wedge \lambda_n = \min_{1 \le k \le n} \{\lambda_k\} \tag{1}$$

where

$$\Xi = \Xi_1 \times \Xi_2 \times \cdots \times \Xi_n = \prod_{k=1}^{n} \Xi_k \tag{2}$$

and

$$\mathfrak{A}_{\Xi} = \mathfrak{A}_{\Xi_1} \times \mathfrak{A}_{\Xi_2} \times \cdots \times \mathfrak{A}_{\Xi_n} = \prod_{k=1}^{n} \mathfrak{A}_{\Xi_k} \tag{3}$$

That is, for each product uncertain event $\Lambda \in \mathfrak{A}_{\Xi}$ (i.e, $\Lambda = \Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_n \in \mathfrak{A}_{\Xi_1} \times \mathfrak{A}_{\Xi_2} \times \cdots \times \mathfrak{A}_{\Xi_n} = \mathfrak{A}_{\Xi}$), the uncertain measure of the event Λ is

$$\lambda\{\Lambda\} = \begin{cases} \sup\limits_{A_1 \times \cdots \times A_n \subset \Lambda} \min\limits_{1 \le k \le n} \lambda\{\Lambda_k\} & \text{if } \sup\limits_{A_1 \times \cdots \times A_n \subset \Lambda} \min\limits_{1 \le k \le n} \lambda\{\Lambda_k\} > 0.5 \\ 1 - \sup\limits_{A_1 \times \cdots \times A_n \subset \Lambda^c} \min\limits_{1 \le k \le n} \lambda\{\Lambda_k\} & \text{if } \sup\limits_{A_1 \times \cdots \times A_n \subset \Lambda^c} \min\limits_{1 \le k \le n} \lambda\{\Lambda_k\} > 0.5 \\ 0.5 & \text{otherwise} \end{cases} \tag{4}$$

Definition 3.1: (Liu, 2007, 2010, 2011) A set function $\lambda : \mathfrak{A}(\Xi) \to [0,1]$ satisfies *Axioms 1-3* is called an uncertain measure. The triple $(\Xi, \mathfrak{A}(\Xi), \lambda)$ is called an uncertainty space.

Definition 3.2: (Liu, 2007, 2010, 2011) An uncertainty variable is a measurable function ξ from an uncertainty space $(\Xi, \mathfrak{A}(\Xi), \lambda)$ to the set of real numbers, i.e., for any Borel set B of real numbers, the set $\{\tau \in \Xi : \xi(\tau) \subset B \in \mathfrak{B}(\mathbb{R})\} \in \mathfrak{A}(\Xi)$, i.e., the pre-image of B is an event.

Remark 3.3: Parallel to revelation of the connotation of randomness in geostatistics, impreciseness occupies an fundamental position in geospatial-temporal uncertainty statistical analysis. In California PM_{10} spatial-temporal study, nearly 60% sites do not have "complete" temporal sequences so that in order to fill the "missing" observations, we have to engage expert's knowledge to pursue "complete sequences" (i.e., to have 19 PM_{10} values at each individual site), which is inevitably imprecise and incomplete. Impreciseness is referred to a term here with an intrinsic property governed by an uncertainty measure or an uncertainty distribution for each of the actual or hypothetical members of an uncertainty population (i.e., collection of expert's knowledge). An uncertainty process is a repeating process whose outcomes follow no describable deterministic pattern, but follow an uncertainty distribution, such that the uncertain measure of the occurrence of each outcome can be only approximated or calculated.

Remark 3.4: Impreciseness exists in engineering, business and research practices due to measurement imperfections, or due to more fundamental reasons, such as insufficient available information, ... , or due to a linguistic nature, because it is an unarguable fact that impreciseness exists intrinsically in expert's knowledge on the real world.

Definition 3.5: Let ξ be a uncertainty quantity of impreciseness on an uncertainty measure space $\left(\Xi,\mathfrak{A}\left(\Xi\right),\lambda\right)$. The uncertainty distribution of ξ is

$$\Psi_{\xi}\left(x\right)=\lambda\left\{\tau\in\Xi\,|\,\xi\left(\tau\right)\le x\right\} \tag{5}$$

An imprecise variable ξ is an uncertainty variable and thus is a measurable mapping, i.e., $\xi:\mathbb{D}\to\mathbb{R}$, $\mathbb{D}\subseteq\mathbb{R}$. An observation of an imprecise variable is a real number, (or more broadly, a symbol, or an interval, or a real-valued vector, a statement, etc), which is a representative of the population or equivalently of an uncertainty distribution $\Psi_{\xi}\left(\cdot\right)$ under a given scheme comprising set and σ-algebra. The single value of a variable with impreciseness should not be understood as an isolated real number rather a representative or a realization from the uncertain population.

Definition 3.6: (Lipschitz condition) Let $f\left(x\right)$ be a real-valued function, $f:\mathbb{R}\to\mathbb{R}$. If for any $x,y\in\mathbb{R}^{n}$, there exists a positive constant $M>0$, such that

$$\left|f\left(x\right)-f\left(y\right)\right|<M\left|x-y\right| \tag{6}$$

Definition 3.7: (Lipschitz continuity) Let $f:\mathbb{R}^{m}\to\mathbb{R}^{m}$

1. for $\forall B\subset\mathbb{R}^{m}$, B to open set, f is Lipschitz continuous on B if $\exists M>0$ such

$$\left\|f\left(x\right)-f\left(y\right)\right\|<M\left\|x-y\right\|,\ \forall x,y\in B \tag{7}$$

where $\left\|\cdot\right\|$ is some metric (for example, Euclidean distance in $\mathbb{R}^{m}$), such

$$d\left(f\left(x\right),f\left(y\right)\right)<Md\left(x,y\right),\ \forall x,y\in B \tag{8}$$

2. for each $z\in\mathbb{R}^{m}$, f is Lipschitz continuous locally on the open ball B of center z radius $M>0$ such

$$B_M(z) = \{y \in \mathbb{R}^m \mid d(y,z) < M\} \tag{9}$$

3. if f is Lipschitz continuous on the whole space $\mathbb{R}^m$, then the function is called globally Lipschitz continuous.

Remark 3.8: For continuity requirements, Lipschitz continuous function is stronger than that of the continuous function in Newton calculus but it is weaker than the differentiable function in Newton differentiability sense. In other words, Lipschitz-continuity does not warrant the first -order differentiability everywhere but it does mean nowhere differentiability. Lipschitz-continuity does not guarantee the existence of the first-order derivative everywhere, however, if exists somewhere, the value of the derivative is bounded since

$$\frac{|f(x)-f(y)|}{|x-y|} < M \tag{10}$$

by recalling the definition of the Newton derivative

$$\lim_{y \to x} \frac{f(x)-f(y)}{x-y} = f'(x) \tag{11}$$

Similar to the concept of stochastic process in probability theory, an uncertain process $\{\xi_t, t \geq 0\}$ is a family of uncertainty variables indexed by t and taking values in the state space $\mathbb{S} \subseteq \mathbb{R}$.

Definition 3.9: (Liu 2010, 2011) Let $\{C_t, t \geq 0\}$ be an uncertain process.

1. $C_0 = 0$ and all the trajectories of realizations are Lipschitz-continous;
2. $\{C_t, t \geq 0\}$ has stationary and independent increments;
3. every increment $C_{t+s} - C_s$ is a normal uncertainty variableb with expected value 0 and variance t^2, i.e., the uncertainty distribution of $C_{t+s} - C_s$ is

$$\Psi_{C_{t+s}-C_s}(z) = \left(1 + \exp\left(-\frac{xz}{\sqrt{3}t}\right)\right)^{-1} \tag{12}$$

Then $\{C_t, t \geq 0\}$ is called a canonical process.

Remark 3.10: Comparing to Brownian motion process $\{B_t, t \geq 0\}$ in probability theory, which is continuous almost everywhere and nowhere is differentiable, while Liu's canonical process $\{C_t, t \geq 0\}$ is Lipschitz-continuous and if $\{C_t, t \geq 0\}$ is differentiable somewhere, the derivative is bounded. Therefore $\{C_t, t \geq 0\}$ is smoother than $\{B_t, t \geq 0\}$.

Definition 3.11: (Liu, 2010, 2011) Suppose $\{C_t, t \geq 0\}$ is a canonical process, and f and g are some given functions, then

$$d\xi_t = f(t, \xi_t)dt + g(t, \xi_t)dC_t \tag{13}$$

is called an uncertain differential equation. A solution to the uncertain differential equation is the uncertain process $\{\xi_t, t \geq 0\}$ satisfying it for any $t > 0$.

Remark 3.12: Since dC_t and $d\xi_t$ are only meaningful under the umbrella of uncertain integral, i.e., the an uncertain differential equation is an alternative representation of

$$\xi_t = \xi_0 + \int_0^t f(s,\xi_s)ds + \int_0^t g(s,\xi_s)dC_s \tag{14}$$

Definition 3.13: The geometric canonical process $\{G_t, t \geq 0\}$ satisfies the uncertain differential equation

$$dG_t = \alpha G_t dt + \sigma G_t dC_t \tag{15}$$

has a solution

$$G_t = \exp(\alpha t + \sigma C_t) \tag{16}$$

where α can be called the drift coefficient and $\sigma > 0$ can be called the diffusion coefficient of the geometric canonical process $\{G_t, t \geq 0\}$ due to the roles played respectively.

4. Spatial interpolation and extrapolation via inverse distance approach

Statistically, spatial interpolation and extrapolation modeling is actually a kind of linear regression modeling exercises, say, kriging methodology. Considering the shortage of California PM_{10} data records, we will utilize a weighted linear combination approach, which was first proposed by Shepard (1968). The weights are the inverse distances between the missing value cell to the actual observed PM_{10} value cells. The weight construction is a deterministic method, which is neutral and does not link to any specific measure theory. It is widely used in spatial predictions and map constructions in geostatistics, but is not probability oriented, rather, molecular mechanics stimulated. A unique aspect of geostatistics is the use of regionalized variables which are variables that fall between random variables and completely deterministic variables. The weight of an observed PM_{10} value is inversely proportional to its distance from the estimated value. Let:

C_{ij} The j^{th} cell on the Site i, (i represent the actual site number), $i = 1,2,\cdots,213$. Note index j points to the cell where no PM_{10} value is recorded, i.e., missing value cell.

x_i Longitude of site i

y_i Latitude of site i

d_{ij} The distance between site i and site j (where missing value j^{th} cell is located)

$$d_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$$

w_{ij} Weight, $w_{ij} = \begin{cases} 0 & \text{Cell } (i,j) \text{ has no } PM_{10} \text{ obs.} \\ 1/d_{ij} & \text{Cell } (i,j) \text{ has } PM_{10} \text{ obs.} \end{cases}$

Then the inverse distance formula is,

$$C_{ij} = \frac{w_{ij}}{\sum_{i=1}^{213} w_{ij}} z_{ij} \tag{17}$$

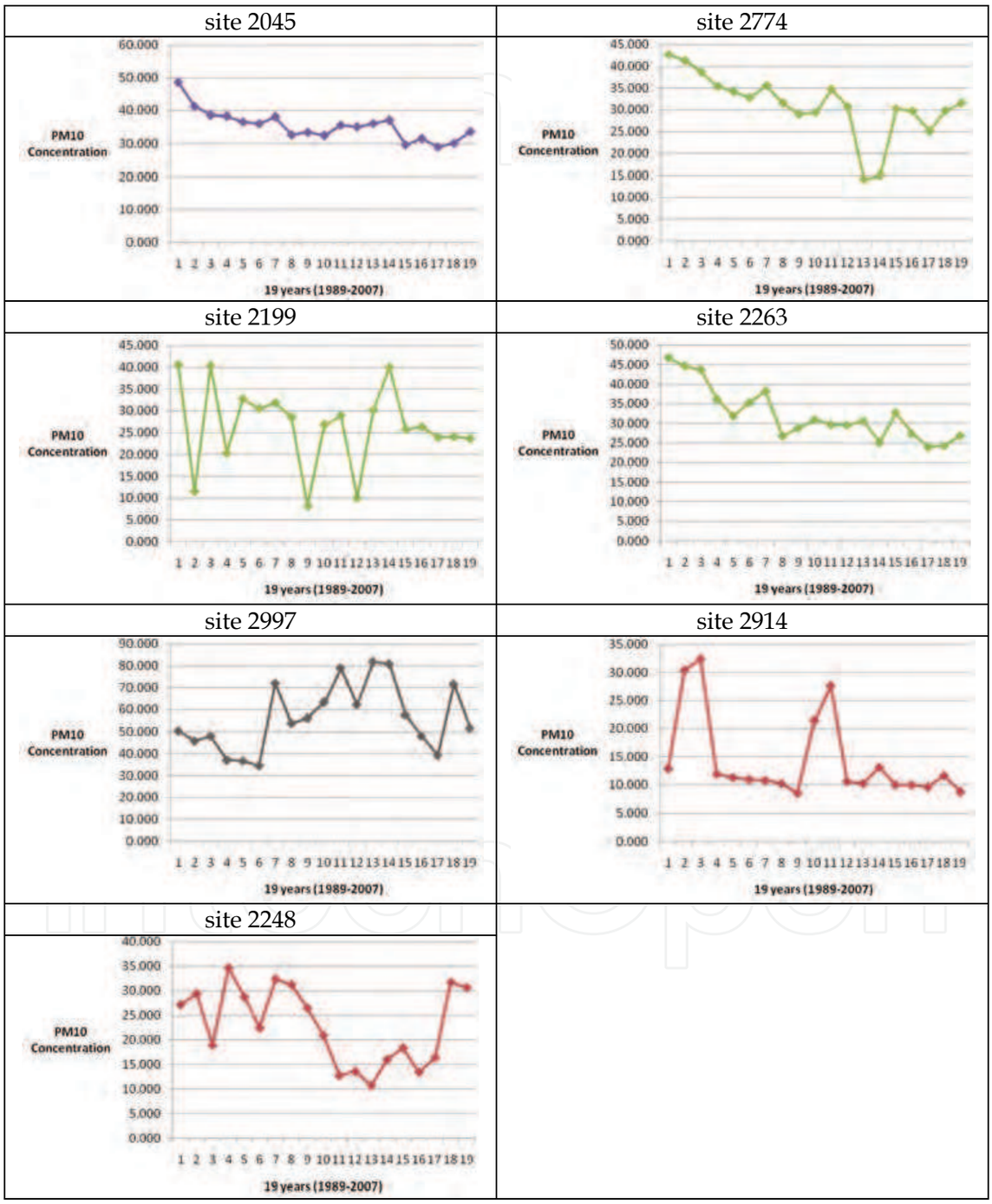


Fig. 4. The 7 sites from the selected 7 counties with completed 19 year observations of PM_{10}

We wrote a VBA Macro to facilitate the interpolations and the extrapolations to "fill" up the 2048 missing value cells in terms of the 1639 cells with PM_{10} values. With the interpolations and the extrapolations, every site has 19 PM_{10} values now. As to whether the inverse distance approach can facilitate highly accurate predictions for each cell without a observed PM_{10} value, we performed a re-interpolation and re- extrapolation scheme (by deleting a true PM_{10} record, then fill it by the remaining records one by one) to evaluate the mean square value for error evaluation, the calculated mean of sum of error squares is 59.885, which is statistically significant (asymptotically).
We plotted sites 2045, 2744, 2199, 2263, 2297, 2914, and 2248 (appeared in Fig. 3) respectively in Fig. 4. By comparing Fig. 3 and 4, it is obvious that only Site 2744 the hazard level changed (moving up to next higher hazard level), while the hazard level of other six sites are unchanged. This may give an justification of the inverse distance approach. Keep in mind, the aim of this article is investigate whether the PM_{10} level is changed over 1989 to 2007 19-year period. The change is not necessarily be accurate but reasonably calculated because of the impreciseness features of PM_{10} complete records.

5. Uncertain analysis of site temporal pattern

Once the interpolations and the extrapolations in terms of the inverse distance approach is completed, a "complete" data set is available, containing 4047 data records of 213 sites over 19 years. The next task is for a given site, how to model the uncertain temporal pattern. It is obvious that the "complete" data set contains impreciseness uncertainty due to the interpolations and the extrapolations. We are unsure that the impreciseness uncertainty is of random uncertainty, so that we still use uncertain measure theory to pursue the temporal uncertainty modelling.
Recall that the **Definition 3.13** in Section 3 facilitates a uncertain geometric canonical process, $\{G_t, t \geq 0\}$. Notice that $G_0 = 0$ may not fit the data reality so that we propose a modified uncertain geometric canonical process, $\{G_t^*, t \geq 0\}$ with $G_0 > 0$:

$$G_t^* = G_0 G_t = G_0 \exp(\alpha t + \sigma C_t) \tag{18}$$

Note that

$$\ln G_t^* = \ln G_0 + \alpha t + \sigma C_t \tag{19}$$

Let $y_t = \ln G_t^*$, $\alpha_0 = \ln G_0$, then we have

$$y_t = \alpha_0 + \alpha t + \sigma C_t \, , t = 1, 2, \cdots, 18 \tag{20}$$

Recall the relevant definitions in Section 3, we have

$$\mathrm{E}[C_t] = 0 \text{ , and } V[C_t] = t^2 \tag{21}$$

But note that for $\forall s < t$,

$$\begin{aligned} \mathrm{E}[C_t C_s] &= \mathrm{E}\left[\left(C_s + (C_t - C_s)\right) C_s\right] \\ &= \mathrm{E}\left[C_s^2\right] + \mathrm{E}\left[C_s (C_t - C_s)\right] \\ &= s^2 + \mathrm{E}\left[C_s (C_t - C_s)\right] \end{aligned} \tag{22}$$

Notice that the increment $C_t - C_s$ is independent of C_s, i.e., C_{t-s} is independent of C_s. Therefore,

$$\begin{aligned}
\mathrm{E}\left[C_{t-s}C_s\right] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z_1 z_2 d\Phi_{C_{t-s},C_s}\left(z_1,z_2\right) \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z_1 z_2 d\min\left\{\Psi_{C_{t-s}}\left(z_1\right),\Upsilon_{C_s}\left(z_2\right)\right\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z_1 z_2 d\min\left\{\left(1+\exp\left(-\frac{\pi z_1^2}{\sqrt{3}(t-s)}\right)\right)^{-1},\left(1+\exp\left(-\frac{\pi z_2^2}{\sqrt{3}s}\right)\right)^{-1}\right\}
\end{aligned} \tag{23}$$

since if ξ_1 and ξ_2 are independent uncertain variables with uncertainty distributions Ψ_{ξ_1} and Υ_{ξ_2} respectively, then the joint uncertainty distribution of (ξ_1,ξ_2) is $\Phi_{\xi_1,\xi_2}\left(z_1,z_2\right)=\min\left\{\Psi_{\xi_1}\left(z_1\right),\Upsilon_{\xi_2}\left(z_2\right)\right\}$. Hence we obtain the expression of $\sigma_{s,t}$:

$$\sigma_{s,t} = \mathrm{E}\left[C_s C_t\right] = s^2 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z_1 z_2 d\min\left\{\left(1+\exp\left(-\frac{\pi z_1^2}{\sqrt{3}(t-s)}\right)\right)^{-1},\left(1+\exp\left(-\frac{\pi z_2^2}{\sqrt{3}s}\right)\right)^{-1}\right\} \tag{24}$$

Then the i^{th} "variance-covariance" matrix for uncertain vector $\left(y_{i1},y_{i2},\cdots,y_{i19}\right)'$

$$\Gamma_i = \sigma^2\left(\sigma_{j,k}^{i}\right)_{19\times 19} \tag{25}$$

where i is the site index, $j,k=1,2,\cdots,19$ are the entry pair in Γ_i matrix. Hence we have a regression model (Draper and Smith, 1981, Guo et al., 2010, Guo, 2010).
For the i^{th} site, the regression model is

$$y_{it} = \alpha_{i0} + \alpha_i t + \sigma_i C_{i,t} \tag{26}$$

Then in terms of the weighted least square criterion we can define an objective function as

$$J_i(\alpha_{i0},\alpha_i) = \left(\underline{Y}_i - X\underline{\beta}_i\right)'\Gamma^{-1}\left(\underline{Y}_i - X\underline{\beta}_i\right) \tag{27}$$

where

$$\underline{Y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{i19} \end{bmatrix} \quad \underline{\beta} = \begin{bmatrix} \alpha_{i0} \\ \alpha_i \end{bmatrix} \quad X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & 19 \end{bmatrix} \tag{28}$$

We further notice that

$$\begin{aligned}
r_{ij} &= y_{i,j} - y_{i,j-1} \\
&= \alpha_i + \sigma_i\left(C_{ij} - C_{i,j-1}\right) = \alpha_i + \sigma_i \Delta C_{ij} \\
&j = 2,3,\cdots,19
\end{aligned} \tag{29}$$

Then it is reasonable to estimate α_i by

$$\hat{\alpha}_i = \frac{1}{18}\sum_{j=2}^{19} r_{ij} \tag{30}$$

Furthermore, we notice that

$$\hat{\sigma}_i = \sqrt{\frac{1}{18}\sum_{j=2}^{19}\left(r_{ij} - \hat{\alpha}_i\right)^2} \tag{31}$$

Also, we can evaluate

$$\sigma_{j,k}^{i} = j^2 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z_1 z_2 d\min\left\{\left(1+\exp\left(-\frac{\pi z_1}{\sqrt{3}(k-j)}\right)\right)^{-1}, \left(1+\exp\left(-\frac{\pi z_2}{\sqrt{3}j}\right)\right)^{-1}\right\} \tag{32}$$

in terms of numerical integration, Then an estimate for Γ_i matrix is obtained:

$$\hat{\Gamma}_i = \hat{\sigma}_i^2\left(\sigma_{j,k}^{i}\right)_{19\times 19} \tag{33}$$

Finally, we use the approximated objective function

$$\hat{J}_i(\alpha_{i0},\alpha_i) = \left(\underline{Y}_i - X\underline{\beta}_i\right)^{'}\hat{\Gamma}_i^{-1}\left(\underline{Y}_i - X\underline{\beta}_i\right) \tag{34}$$

to obtain a pair of estimates $(\tilde{\alpha}_{i0},\tilde{\alpha}_i)$. Repeat this estimation process until all the 213 weighted least square estimate $(\tilde{\alpha}_{i0},\tilde{\alpha}_i)$ are obtained.

Recall the definition of coefficient α_i so that the sign and the absolute value of α_i indicates the geometric change over the 19 years. Since the estimation procedure of α_i involves all the spatial-temporal information, it is reasonable to have them plotted in a kriging map to reveal the overall changes over 19-year period.

6. Kriging maps and time-change maps based on completed PM_{10} data

Kriging map presentation is vital for a geostatistian's visualization, and maps reveal hidden information or the whole picture. A sample statistic is typically condensing the wide-spread information into a numerical point. While, a kringing map is actually a map statistic (or a statistical map) which contains infinitely many information aggregated from limited "sample" information (i.e., observations). Kriging itself is not specifically probability oriented, it is another weighted linear combination prediction, but requires more mathematical assumptions. In fuzzy geostatistics, the fuzzy kriging scheme has also been developed (Bardossy et al., 1990).

Ordinary kriging (abbreviated as OK) is a linear predictor, see Cressie (1991) and Mase (2011). The formula is

$$Z(s_0) = \sum_{j=1}^{N}\lambda_j Z(s_j) \tag{35}$$

where s_j are spatial locations with observation $Z(s_j)$ available and the coefficients λ_j satisfy the OK linear equation system

$$\begin{cases} \sum_{j=1}^{N} \lambda_j \gamma\left(\varepsilon(s_j) - \varepsilon(s_i)\right) - \psi = \gamma\left(\varepsilon(s_0) - \varepsilon(s_i)\right),\ i = 1,2,\cdots,N \\ \sum_{j=1}^{N} \lambda_j = 1 \end{cases} \tag{36}$$

The OK system is generated under the assumptions of an additive spatial model

$$Z(s) = \mu(s) + \varepsilon(s) \tag{37}$$

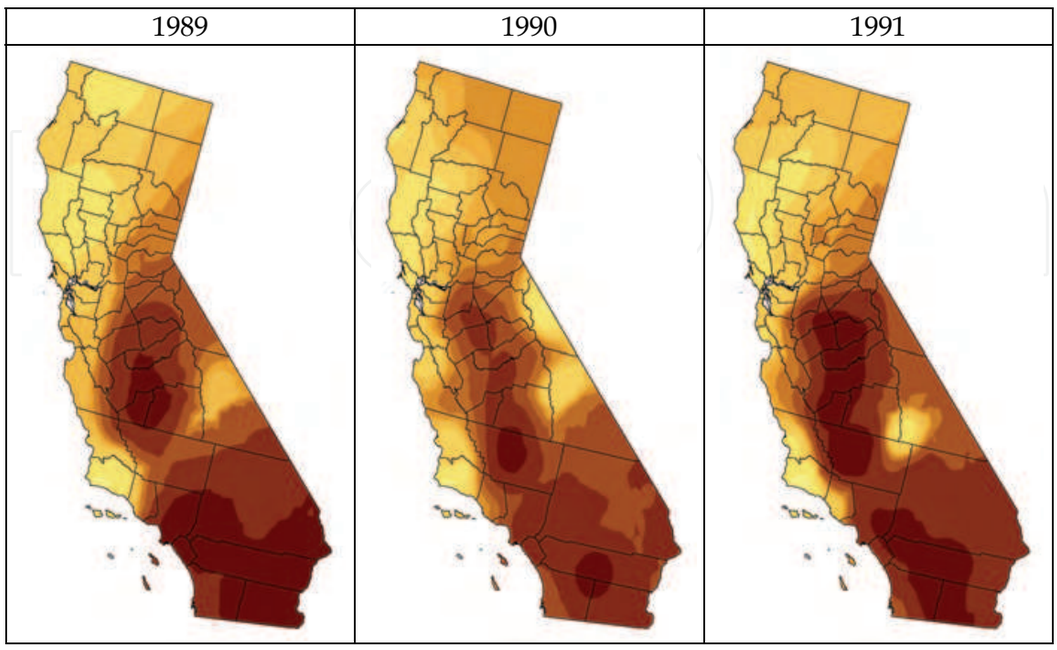
where $\mu(s)$ is the basic (expected) spatial trend and $\varepsilon(s)$ is a Gaussian error $N\left(0, \sigma^2(s)\right)$, i.e., Gaussian variable with mean and variance

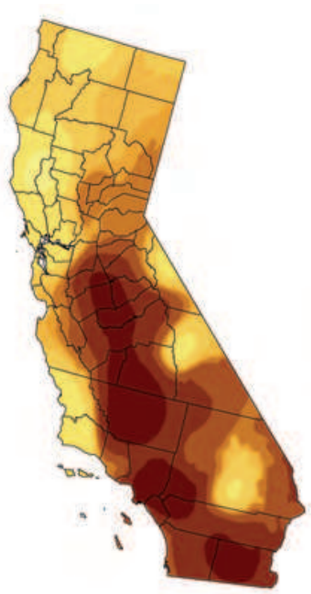
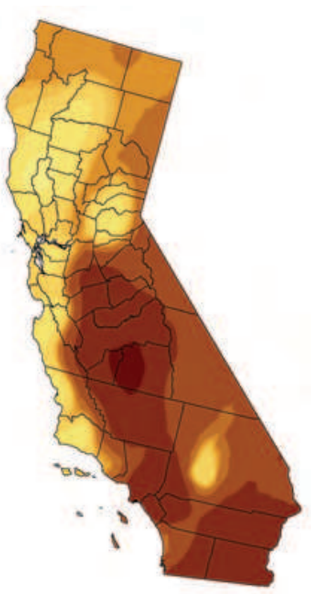
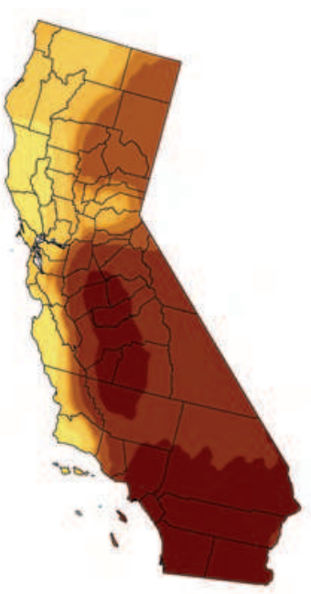
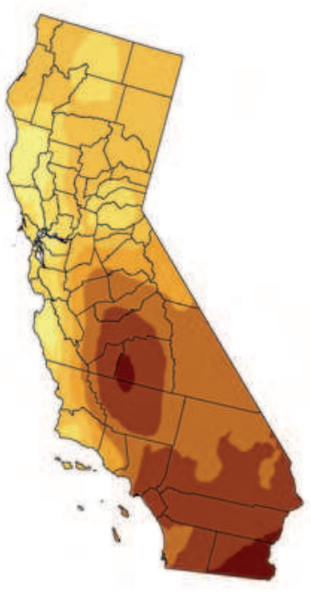
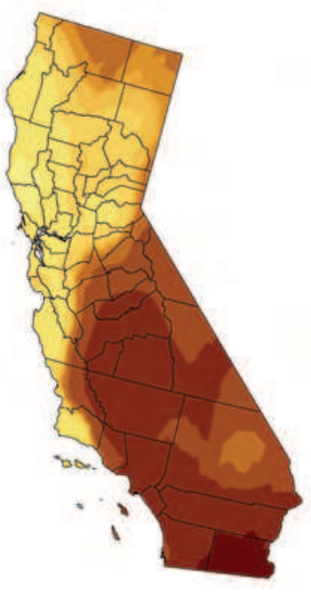
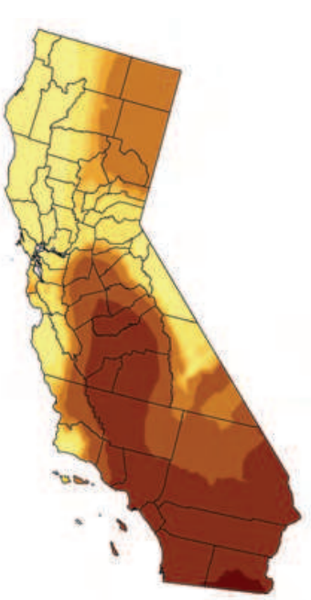
$$\mathrm{E}\left[\varepsilon(s)\right] = 0,\ V\left[\varepsilon(s)\right] = \sigma^2(s) \tag{38}$$

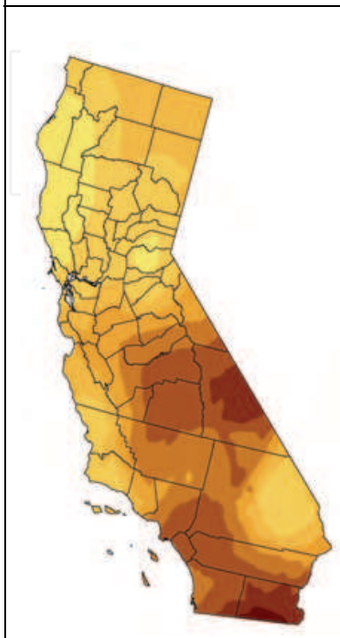
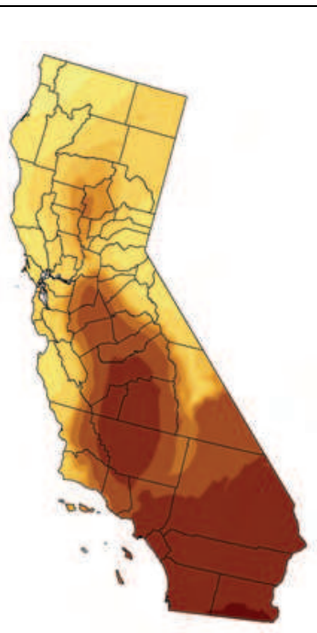
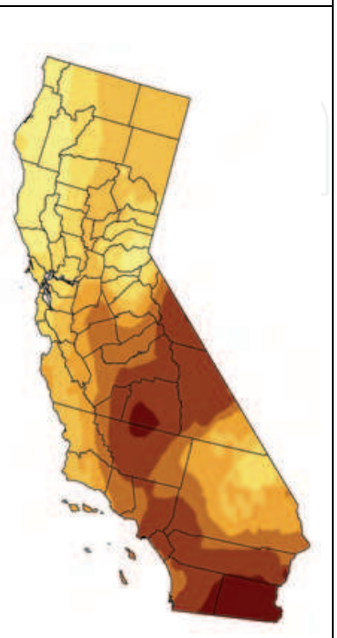
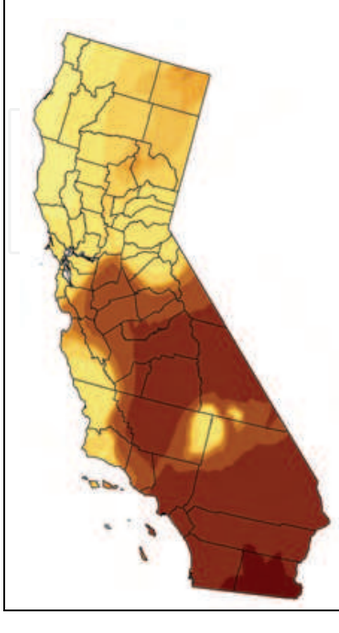
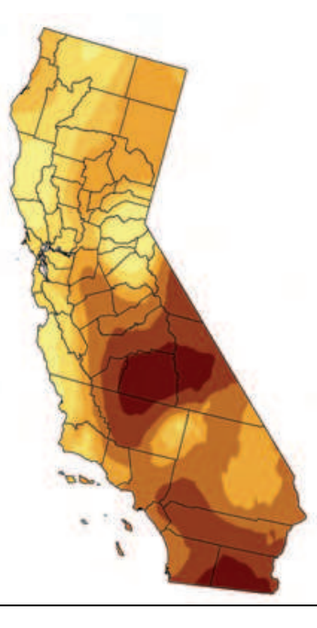
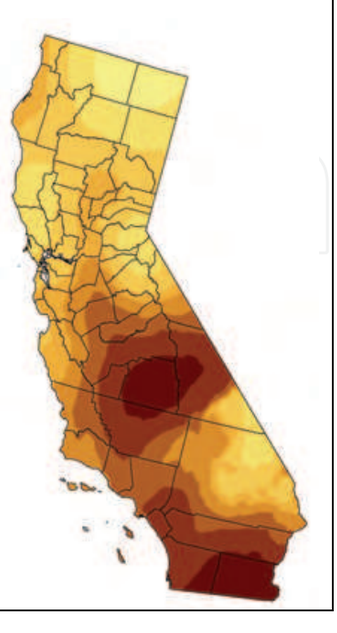
respectively. Accordingly, the variogram 2γ of the random error function $\varepsilon(\cdot)$ is just defined by

$$2\gamma(h) = \mathrm{E}\left[\left(\varepsilon(s+h) - \varepsilon(h)\right)^2\right] \tag{39}$$

where h is the separate vector between two spatial point $s+h$ and s under the isotropy assumption.



1992	1993	1994
		
1995	1996	1997
		

1998	1999	2000
		
2001	2002	2003
		

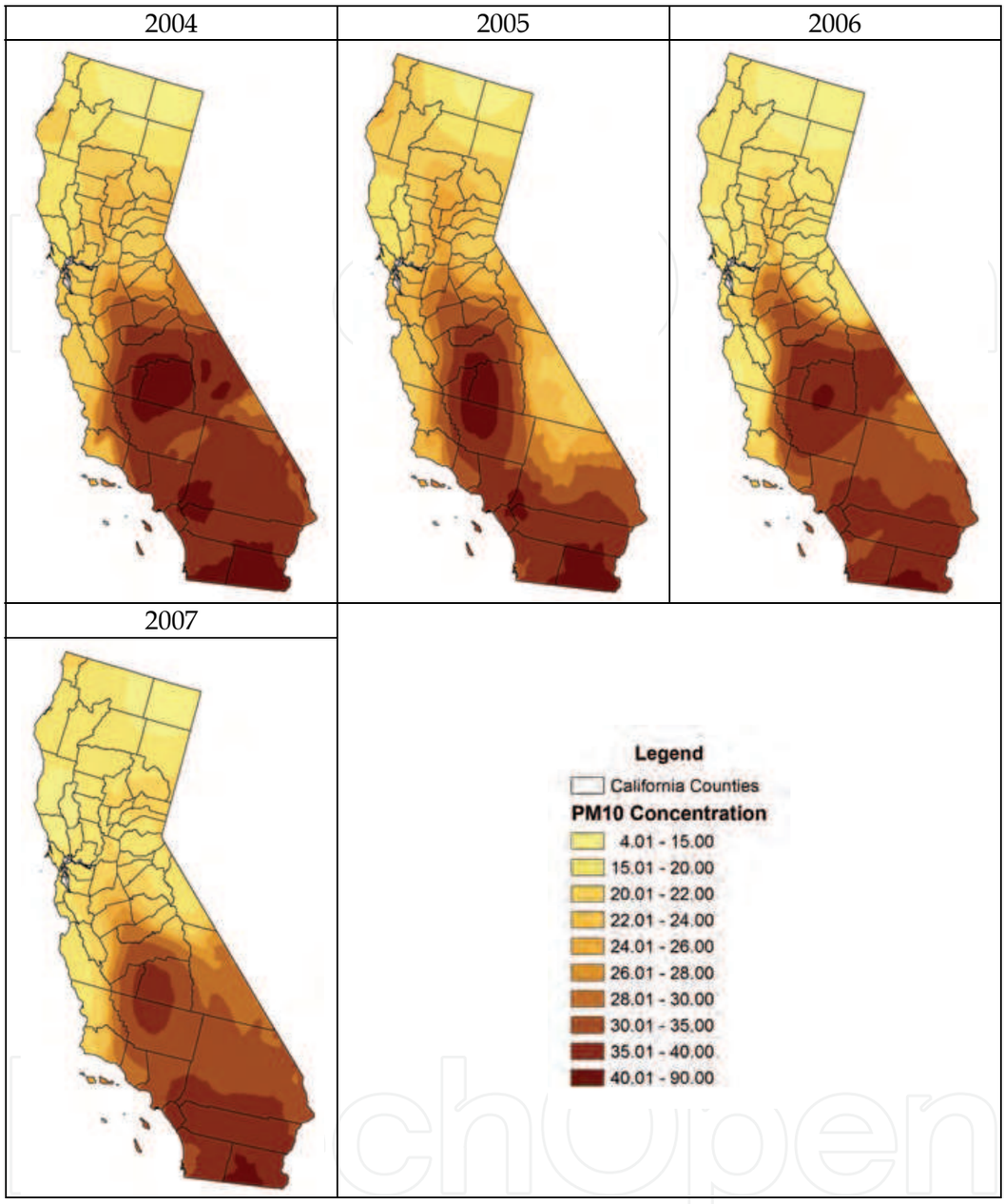


Fig. 5. Kriging Prediction Maps for PM_{10} in California 1989-2007.

The 213 observation sites now have 19-year PM_{10} values, a "complete" data set is now available, containing 4047 data records of 213 sites over 19 years, and then the 19 ordinary kriging pred4iction maps are generated for comparisons. In Fig. 5, all 19 years of PM_{10} concentration in California State are shown. It is very interesting to examine the change in PM_{10} concentrations through the 19 years, based upon the modelled complete 213 site data. In particular, 1998 shows to have an extremely low PM_{10} concentration. Although air quality is varied over the years, but in general, the PM_{10} concentration is decreasing, showing an improvement of air quality trend.

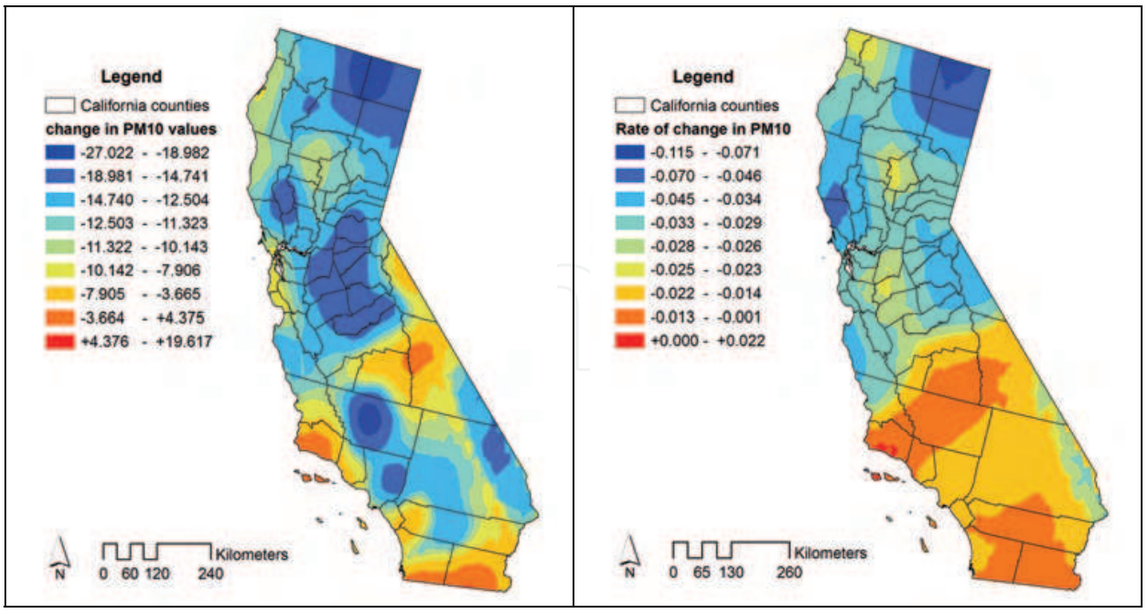


Fig. 6. Changes in PM_{10} values and the rate of change of PM_{10} in California between 1989 and 2007.

As one can clearly see from Fig. 6, that PM_{10} concentration has clearly decreased over the 19 years, and air quality has improved remarkably over the years. The blue and green colours show negative changes, and red shows positive changes or near positive changes. Counties such as San Diego, Inyo, Santa Barbara, Imperial, still show an increase in PM_{10} concentration in the air, and indicate bad air quality. While Kern, Modoc, Siskiyou counties show the most improvement in air quality. The left map in Fig. 6 is PM_{10} record difference between 2007 and 1989 at each location, in total 231 values, and then a difference map is constructed. It is obvious that the difference map only utilizes 1989 and 2007 two-year PM_{10} records, 1990, 1991, ..., 2006 seventeen years' information do not participate the change map construction. The right map in Fig. 6 show completed $\tilde{\alpha}_i$, $i = 1, 2, \cdots, 213$, the rate of change over 1989 to 2007 19-year period.

Note that the calculations of $\tilde{\alpha}_i$, $i = 1, 2, \cdots, 213$ involve all nineteen years by temporal regression, the dependent variable y are estimated form the actual PM_{10} observations cross over all the available locations. Therefore, the rate of change parameter α_i at each individual location contains all spatial-temporal information. It is reasonable to say the rate of change parameter $\tilde{\alpha}_i$ is an aggregate statistic for revealing the 19-year changes over 213 locations. $\tilde{\alpha}_i$ kriging map is thus different from 2007-1989 kriging maps. The positive sign of $\tilde{\alpha}_i$ indicates the increasing trend in PM_{10} concentration, while the negative sign f $\tilde{\alpha}_i$ indicates the decreasing trend in PM_{10} concentration. The absolute value of $\tilde{\alpha}_i$ reveals the magnitude of change of PM_{10} concentration. It is worth to report, among 213 locations, 193 locations have negative $\tilde{\alpha}_i$, while the negative $\tilde{\alpha}_i$ locations are 20 (9% approximately).

7. Discussion

Air quality and health is always a central issue to public concern on the quality of life. In this chapter, we examined PM_{10} levels over 19 years, from 1989 to 2007, in the California State.

Facing the difficult task of a lack of "complete" PM_{10} observational data, we utilised the inverse distance weight methodology to "fill" in the locations with missing values. By doing so, the impreciseness uncertainty is introduced, which is not necessarily explained by probability measure foundation. We noted the character of a regionalized variable in geostatistics and therefore engage Liu's (2010, 2011) uncertainty theory to address the impreciseness uncertainty. In this case, we developed a series of uncertain measure theory founded spatial-temporal methodology, including the inverse distance scheme, the kriging scheme, and the geometric canonical process based weighted regression analysis in order to extract the change information from the incomplete 1989-2007 PM_{10} records. The use of the rate of change parameter alpha is a new idea and it is an aggregate change index utilized all spatial-temporal data information available. It is far better than classical change treatments. However, due to the limitations of our ability, we are unable to demonstrate the detailed uncertain measure based spatial analysis model. In the future research, we plan to develop a more solid uncertain spatial prediction methodology.

8. Acknowledgements

I would like to thank the California Air Resources Board for providing the air quality data used in this paper. This study is supported financially by the National Research Foundation of South Africa (Ref. No. IFR2009090800013) and (Ref. No. IFR2011040400096).

9. References

Bardossy, A.; Bogardi, I. & Kelly, E. (1990). Kriging with imprecise (fuzzy) variograms, I: Theory. *Mathematical Geology*, Vol. 22, pp. 63–79.

California Environmental Protection Agency Air Resources Board. (2010). Ambient Air Quality Standards (AAQS) for Particulate Matter. (www.arb.ca.gov)

Cressie, N. (1991). *Statistics for Spatial Data*. Wiley-Interscience, John-Wiley & Sons Inc. New York.

Deng, J. L. (1984). *Grey dynamic modeling and its application in long-term prediction of food productions*. Exploration of Nature, Vol. 3, No. 3, pp. 7-43.

Draper, N. & Smith, H. (1981). *Applied Regression Analysis*. 2nd Edition. John Wiley & Sons, Inc. New York.

Electric, W. (1956). *Statistical Quality Control Handbook*. Western Elctric Corporation, Indianapolis.

Environmental Protection Agency (EPA). 2010. National Ambient Air Quality Standards (NAAQS). U.S. Environmental Protection Agency. (www.epa.gov)

Guo, D.; Guo, R. & Thiart, C. (2007). Predicting Air Pollution Using Fuzzy Membership Grade Kriging. *Journal of Computers, Environment and Urban Systems*. Editors: Andy P Jones and Iain Lake. Elsevier, Vol. 31, No. 1, pp. 33-51. ISSN: 0198-9715

Guo, D.; Guo, R. & Thiart, C. (2007). Credibility Measure Based Membership Grade Kriging. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*. Vol. 15, No. Supp. 2, (April 2007), pp. 53-66. B.D. Liu (Editor). ISSN 0218-4885

Guo, D. (2010). *Contributions to Spatial Uncertainty Modelling in GIS*. Lambert Academic Publishing (online.lap-publishing.com). ISBN 978-3-8433-7388-3

Guo, R.; Cui, Y.H. & Guo, D. (2010). Uncertainty Statistics. (Submitted to Journal of Uncertainty Systems, under review)

Guo, R.; Cui, Y.H. & Guo, D. (2010). Uncertainty Linear Regression Models. (Submitted to Journal of Uncertainty Systems, under review)

Guo, R.; Guo, D. & Thiart, C. (2010). Liu's Uncertainty Normal Distribution. *Proceedings of the First International Conference on Uncertainty Theory*, August 11-19, 2010, Urumchi & Kashi, China, pp 191-207, Editors: Dan A. Ralescu, Jin Peng, and Renkuan Guo. International Consortium for Uncertainty Theory. ISSN 2079-5238

Kolmogorov, A.N. (1950) *Foundations of the Theory of Probability*. Translated by Nathan Morrison. Chelsea, New York.

Liu, B.D. (2007). *Uncertainty Theory: An Introduction to Its Axiomatic Foundations*. 2nd Edition. Springer-Verlag Heidelberg, Berlin.

Liu, B.D. (2010). *Uncertainty Theory*: *A Branch of Mathematics of Modelling Human Uncertainty*. Springer-Verlag, Berlin.

Liu, B.D. (2011). *Uncertainty Theory,* 4th Edition, 17 February, 2011 drafted version.

Liu, S.F. & Lin, Y. (2006). *Grey Information*. Springer-Verlag, London.

Mase, S. (2011). GeoStatistics and Kriging Predictors, In: *International Encyclopedia of Statistical Science*. Editor: Miodrag Lovric, 1st Edition, 2011, LVIII, pp. 609-612, Springer.

Montgomery, D.C. (2001). *Introduction to Statistical Quality Control*. 4th Edition. John Wiley & Sons, Now York.

Moore, R.E. (1966). *Interval Analysis*. Prentice-Hall, Englewood Cliff, NJ. ISBN 0-13-476853-1

Pawlak, Z. (1982). Rough Sets. *International Journal of Computer and Information Sciences*, Vol. 11, pp. 341-356.

Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 ACM National Conference*, pp. 517–524.

Utkin, L.V. & Gurov, S.V. (1998). New reliability models on the basis of the theory of imprecise probabilities. *Proceedings of the 5th International Conference on Soft Computing and Information Intelligent Systems*, Vol. 2, pp. 656-659.

Utkin, L.V. & Gurov, S.V. (2000). New Reliability Models Based on Imprecise Probabilities. *Advanced Signal Processing Technology, Soft Computing*. Fuzzy Logic Systems Institute (FLSI) Soft Computing Series - Vol. 1, pp. 110-139, Charles Hsu (editor). Publisher, World Scientific. November 2000. ISBN 9789812792105

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall. ISBN 0412286602

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, Vol. 8, pp. 338-353.

Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, Vol. 1, pp. 3-28.

Advanced Air Pollution

Edited by Dr. Farhad Nejadkoorki

ISBN 978-953-307-511-2
Hard cover, 584 pages
Publisher InTech
Published online 17, August, 2011
Published in print edition August, 2011

Leading air quality professionals describe different aspects of air pollution. The book presents information on four broad areas of interest in the air pollution field; the air pollution monitoring; air quality modeling; the GIS techniques to manage air quality; the new approaches to manage air quality. This book fulfills the need on the latest concepts of air pollution science and provides comprehensive information on all relevant components relating to air pollution issues in urban areas and industries. The book is suitable for a variety of scientists who wish to follow application of the theory in practice in air pollution. Known for its broad case studies, the book emphasizes an insightful of the connection between sources and control of air pollution, rather than being a simple manual on the subject.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen