Journal of Uncertain Systems
Vol.2, No.1, pp.31-53, 2008
Online at: www.jus.org.uk

PDEMR Modelling of the Protea Rare Species Spatial Patterns

Renkuan Guo[1], Danni Guo[2], Guy F. Midgley[2], and Anthony G. Rebelo[2]

[1]*Department of Statistical Sciences*
University of Cape Town
Private Bag, Rondebosch 7701, Cape Town, South Africa
[2]*Global Change Research Group*
South African National Biodiversity Institute
Private Bag X7, Claremont 7735, Cape Town, South Africa
rguo@stats.uct.ac.za

Received 8 August 2007; Accepted 12 October 2007

Abstract

Ecological data is very costly and difficult to collect, and quite often the sampled data are insufficient for further spatial analysis. Today, we as spatial modellers are often presented with the situation whereby a set of data is collected already, although from the viewpoint of spatial analysis the data is insufficient, but re-sampling is impossible because of the cost and time limits. In this paper, we are dealing with a two spatial problems whereby: the data is just species presence only and no numerical data; and also the data sampling is not well spread over the study area. These are two very common problems that spatial modellers face everyday, and in this paper we provide some simple techniques to deal with these problems. We firstly use frequency counts to deal with species presence data, then use the recently developed *partial differential equation motivated regression* (PDEMR) model to predict the unknown locations, and finally combine these data to produce a kriging prediction map. These techniques are fairly new, but very effective in dealing with ecological data problems. For illustration, Protea rare species (i.e., the population size of 10 to 100), in the Cape Floristic Region, from 1992 to 2002, South Africa, are used.

Keywords: PDEMR model, credibility measure, random fuzzy variable, Maximum Average Chance Principle

1 Introduction

Today, ecological data is very costly and difficult to collect, and quite often the sampled data are insufficient for further spatial analysis. We as spatial modellers are often presented with the situation whereby a set of data is collected already, although from the viewpoint of spatial analysis the data is insufficient, but re-sampling is impossible because of the cost and time limits. Since the samples are not designed for spatial predictions, the samples are not well spread over the study area, and can not be used for spatial predictions such as kriging. Another problem is that quite often the data collected are just species presence data or categorical data, and this makes very difficult to model the plants, and impossible to do a kriging prediction map. These are two very common problems that spatial modellers face everyday, and in this paper we will provide some simple techniques to deal with these problems.

In this paper, we will model the Protea species in the population size of 10 to 100, in the Cape Floristic Region, from 1992 to 2002, in South Africa. We firstly use frequency counts to deal with species presence data, then use the recently developed *partial differential equation motivated regression* (PDEMR) model to predict the unknown locations, and finally combine these data to produce a kriging prediction map. These techniques are fairly new, but very effective in dealing with ecological data problems.

2 Proteas in the Cape Floristic Region

The Cape Floristic Region is located at the southern tip of the Africa, and it covers parts of Western and Eastern Cape provinces of South Africa. It is home to some 9030 plant species, and nearly 70% of which are found nowhere else. Fynbos is the predominate ecosystem in the Cape Floristic Region, and it is under serious threat (Freeth *et al.* [2]).

The Protea Atlas Project collected samples of Fynbos's flowering Proteas in the Cape Floristic Region, South Africa, from 1992 to 2002. These sample data provide valuable information on the distribution and change in the Proteas. In this case, we are focusing on the category of Proteas that has the estimated population size from 10 to 100, per sample site.

Figure 1 below shows the location of the Cape Floristic Region within South Africa, and Figure 2 shows the locations of Proteas occurrence of the population size of 10 to 100, in the Cape Floristic Region, from 1992 to 2002.

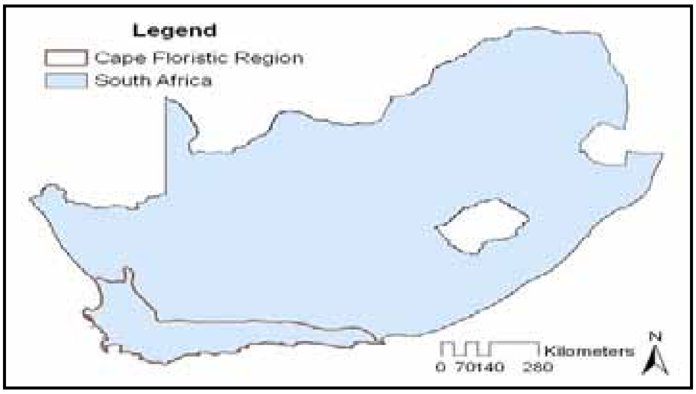
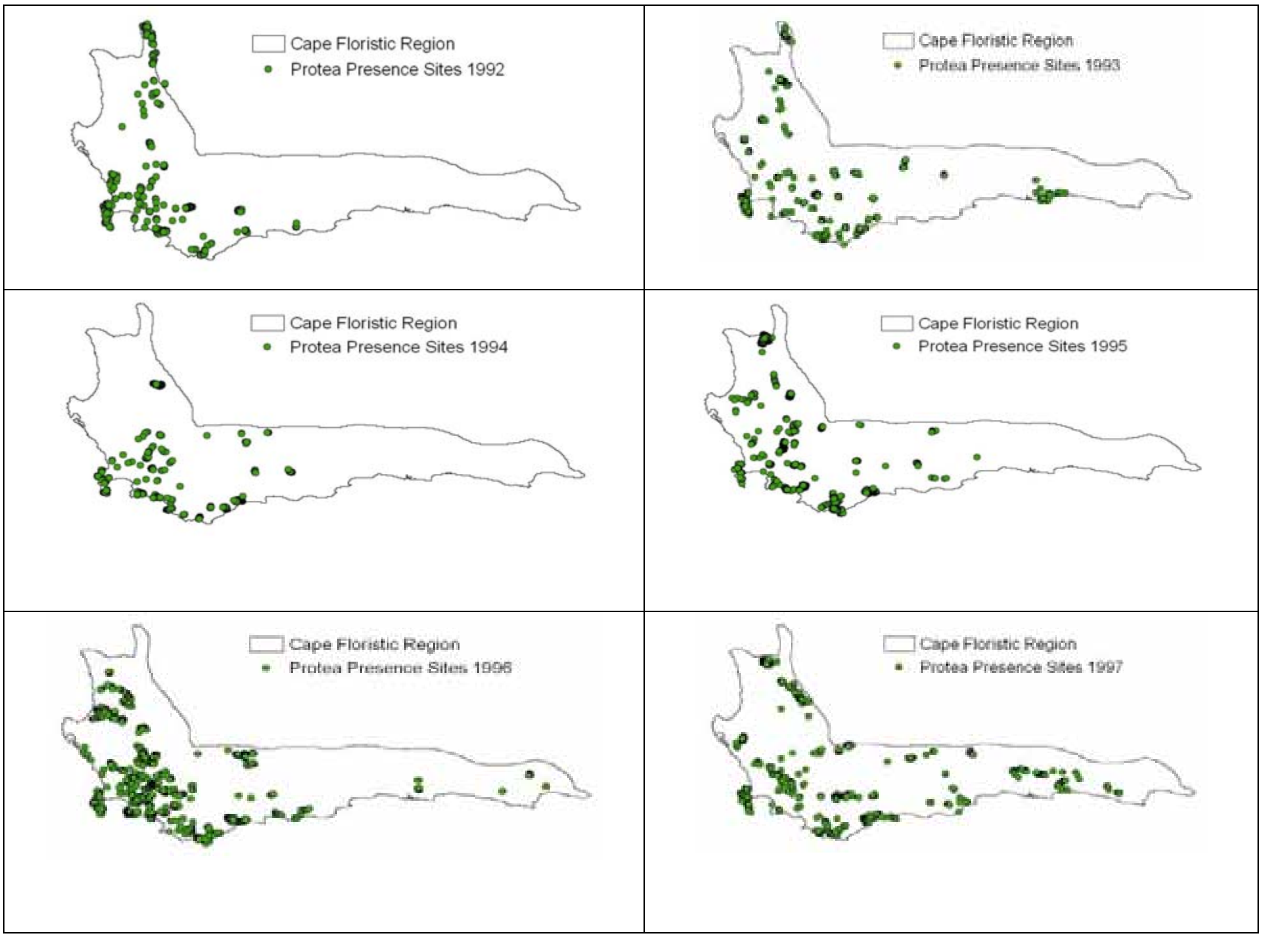


Figure 1. The Cape Floristic Region within South Africa



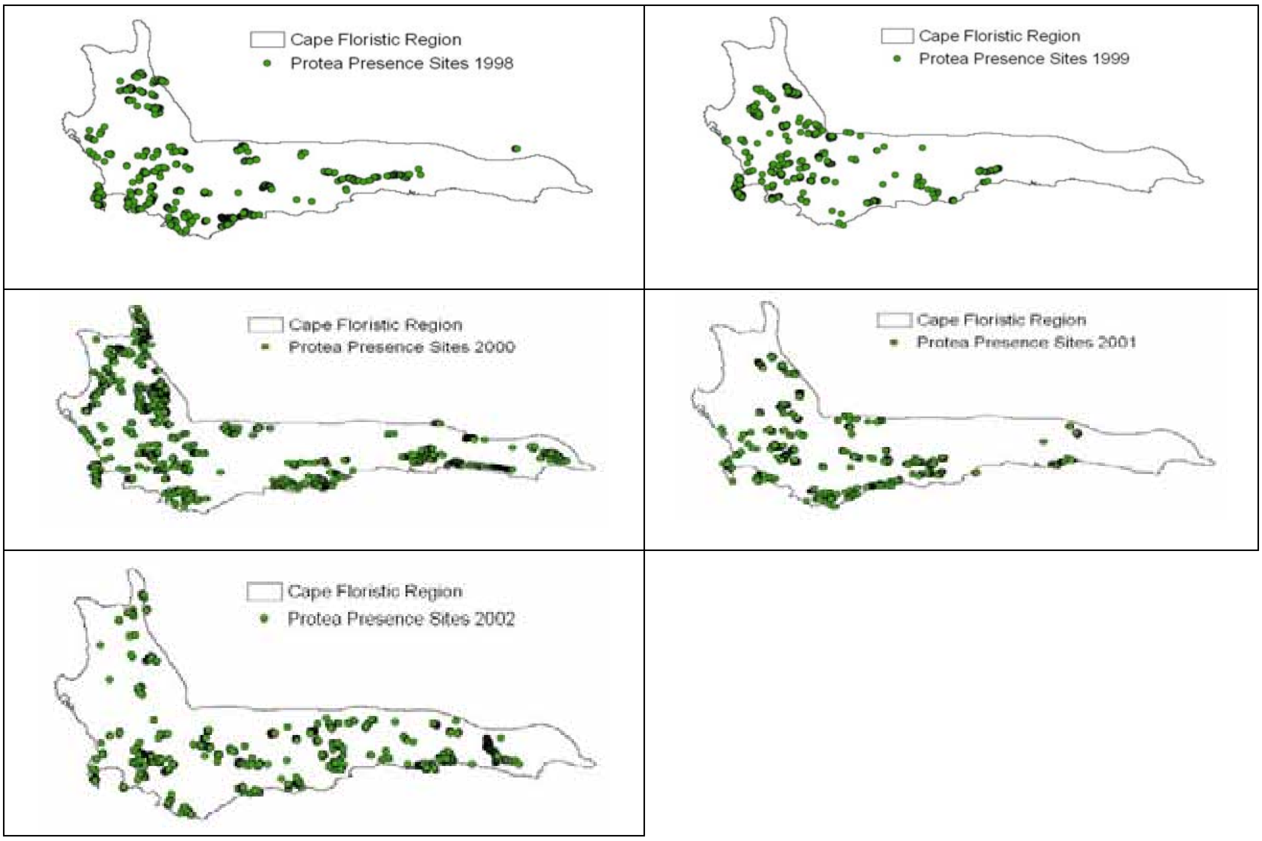


Figure 2. The sample locations of proteas in the population size of 10-100, in the Cape Floristic Region, 1992-2002

As one can see from Figure 2, the sample locations are not well spread, since its original purpose was spatial predictions, but for scientific and biodiversity knowledge. The samples tended to focus in certain areas, while other areas are entirely un-sampled. This creates a problem for kriging predictions. The Protea data are presence only data, and not numerical, which creates another problem for spatial analysis.

3 Frequency Counts of the Occurrence of Proteas

To solve the problem of presence data only, this being a categorical data issue, a simple technique of using frequency counts is used. The Cape Floristic Region is divided into 243 grid cells, and within each cell, the presence of Protea species is counted, and the resulting value is attached to each centroid point of each cell. The centroid point is needed in order for kriging prediction maps to be produced. See Figure 3.

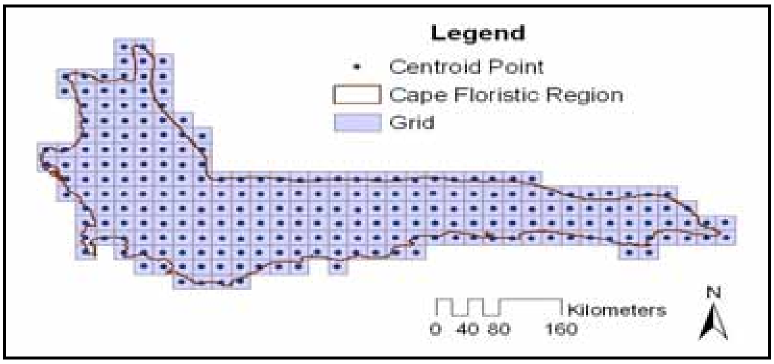
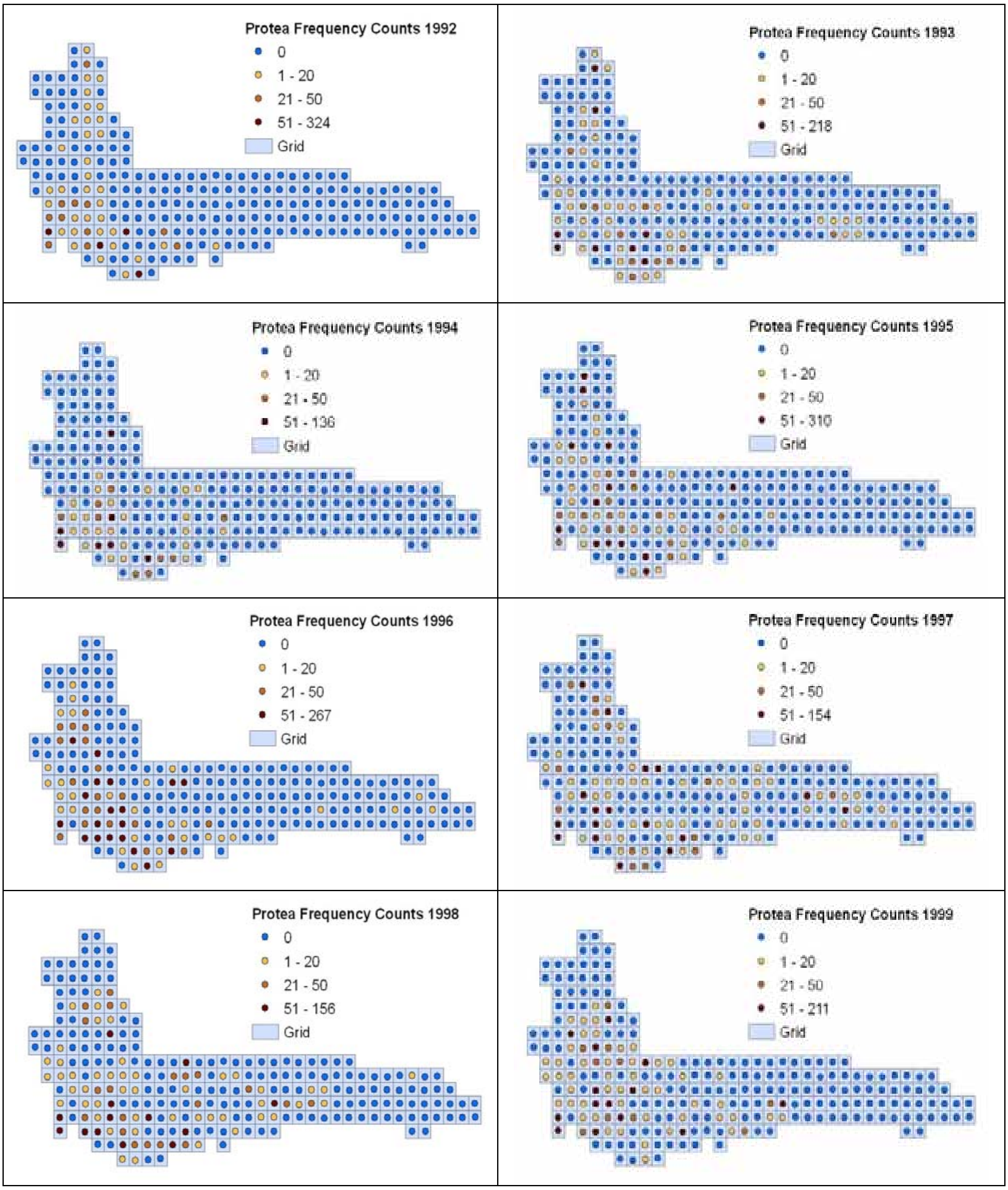


Figure 3. The grid cell division of Cape Floristic Region

In Figure 4, the blue color are 0 in value, it shows the cells that does not have any frequency counts at all. In other words, the blue point cells show the un-sampled locations within the Cape Floristic Region. It is clear that a lot of the areas are un-sampled, and these locations vary from year to year. In order for an accurate kriging prediction map to be produced, the missing cells must be filled. This means that the PDEMR model must be used in order to predict the un-sampled cells.



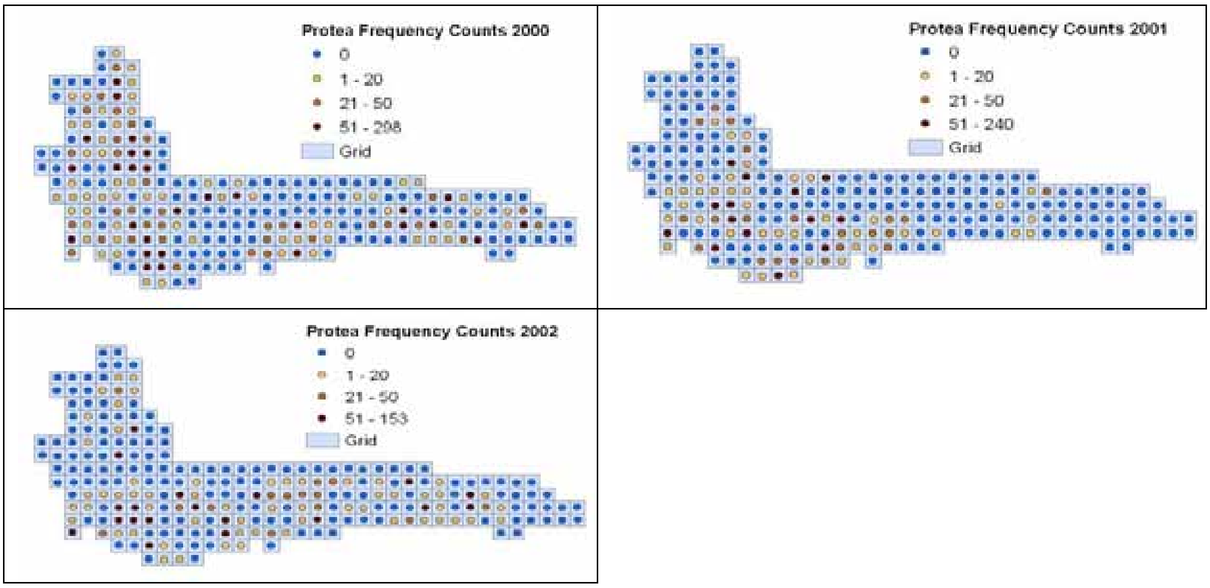


Figure 4. The sampled frequency counts of proteas in the population size of 10-100, in the Cape Floristic Region, 1992-2002

4 The Concept of DEMR Model and the Coupling Principle

In engineering theory, particularly, in modern control theory, it is often convenient to utilize a differential equation to describe the dynamic law of a continuous system. However, the unknown parameter vector $\underline{\theta}$ associated with system

Definition 4.1. *A pair of equations*

$$\begin{cases} \dfrac{d^{(p)}x}{dt^{p}} = \varphi\left(\dfrac{d^{(p-1)}x}{dt^{p-1}}, \dfrac{d^{(p-2)}x}{dt^{p-2}}, \cdots, x; \underline{\theta}\right) & (a) \\ \dfrac{1}{h^{p}}\Delta x^{(p)}(k) = \varphi\left(\dfrac{1}{h^{p-1}}\Delta x^{(p-1)}(k), \dfrac{1}{h^{p-2}}\Delta x^{(p-2)}(k), \cdots, \hat{x}(k); \underline{\theta}\right) + \varepsilon_k,\ k = 2,3,\cdots,n & (b) \end{cases} \tag{1}$$

is called the p[th]-order univariate differential equation motivated regression model, abbreviated as (p[th]-order univariate) DEMR model. Eq. (1a) is called the motivated differential equation and Eq. (1b) is called the (first) coupled regression model, where h is the grid size for the first difference Δ. As to the term $\hat{x}(k)$ is the approximation to primitive function $x(t)$ at $t = k$.

If the observation on the system is at the first difference level, denoted as $\Delta X = \{\Delta x(1), \Delta x(2), \cdots, \Delta x(n)\}$, where $\Delta x(k) = x(k) - x(k-1)$. Note that the relation between summation operator Σ and difference operator Δ, define $\Delta x(1) \triangleq x(1)$, *then*

$$x(k) = \sum_{i=1}^{k}\left(\Delta x(i)\right). \tag{2}$$

It is often using

$$\hat{x}(k) = \frac{1}{2}\left[x(k) + x(k-1)\right] \tag{3}$$

as a first approximation to $x(t)$ at $t = k$. Finally, $\{\varepsilon_k, k = 2,3,\cdots,n\}$ is the error terms of the coupled regression model in Eq. (1b) paired in the above equation system Eq. (1). The nature of errors in Eq. (1) will be discussed later. For a better understanding, let us examine a simple example.

Example 4.2. *Equation system*

$$\begin{cases} \dfrac{dx}{dt} = \alpha + \beta x & (a) \\ \Delta x(k) = \alpha + \beta \hat{x}(k) + \varepsilon_k,\ k = 2,3,\cdots,n & (b) \end{cases} \quad (4)$$

is the simplest first-order univariate DEMR model. Eq. (4b) is called as the coupled regression (abbreviated as CREG) model because its form strictly follows a "translation rule" based on the form of the motivated differential equation. We call this translation rule as the coupling principle in DEMR.

For an overall intuitive picture of DEMR model, we list the components and the translation rule in terms of the coupling principle in Table 1.

Table1. Coupling rule in univariate first-order DEMR model

Term	Motivated DE	Coupled REG
Translation rule between MDE and CREG		
Intrinsic feature	Continuous	Discrete
Independent variable	t	k
1st-order derivative	$dx(t)/dt$	$\Delta x(k) = x(k) - x(k-1)$
pst-order derivative	$d^{(p)}x(t)/dt^p$	$\Delta^n x(k) = \Delta^{n-1}x(k) - \Delta^{n-1}x(k-1)$
Primitive function	$x(t)$	$x(k)$
Model formation	$\frac{dx(t)}{dt} = \alpha + \beta x(t)$	$\Delta x(k) = \alpha + \beta \hat{x}(k) + \varepsilon_k$
PARAMETER COUPLING		
Parameter	(α,β)	(a,b)
Dynamics (Solution)	$x(t) = \left[x(0) - \frac{\alpha}{\beta}\right]e^{\beta t} + \frac{\alpha}{\beta}$	$\hat{x}(k+1) = \left[x(1) - \frac{a}{b}\right]e^{bk} + \frac{a}{b}$
Filtering (Prediction)	$dx(t)/dt = \left[\alpha - \beta\, dx(0)/dt\right]e^{\beta t}$	$\Delta\hat{x}(k+1) = \hat{x}(k+1) - \hat{x}(k)$

A fundamental note is made here that the original observations are treated as the approximated derivatives of the dynamic law x(t), however, after the rule finding, the modelling is still required to return back to the derivative level because that is the observational one.

5 Partial Differential Equation Model

It is often the case that a variable (or a group of variables, i.e., vector) under investigation relates to multi-factors and the functional relationships are specified by a system of partial equations. Similar to DEMR modelling cases, we may also face the sparse data availability. Therefore, it is necessary to investigate the partial differential equation (system) motivated (multivariate) regression (abbreviated as PDMR) modelling. As a necessary, let us review the partial differential equation (system) theory.

5.1 A Family of Partial Equation Model

The family of partial differential equation system under investigation takes its form

$$\begin{cases} \dfrac{\partial \underline{z}}{\partial x_i} = \underline{f}_i\left(\underline{z}, \underline{x}\right), \ i = 1, 2, \cdots, m \\ \underline{z}\left(\underline{x}^0\right) = \underline{z}^0 \end{cases} \tag{5}$$

where

$$\underline{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{bmatrix} \quad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \tag{6}$$

and

$$\underline{f}_i\left(\underline{z}, \underline{x}\right) = \left(f_{i1}\left(\underline{z}, \underline{x}\right), f_{i2}\left(\underline{z}, \underline{x}\right), \cdots, f_{im}\left(\underline{z}, \underline{x}\right)\right)^T . \tag{7}$$

5.2 A Linear Partial Equation System Model

A linear partial differential equation system takes its form

$$\begin{cases} \dfrac{\partial \underline{z}}{\partial x_i} = A_i\left(\underline{x}\right)\underline{z} + \underline{b}\left(\underline{x}\right), \ i = 1, 2, \cdots, n \\ \underline{z}\left(\underline{x}^0\right) = \underline{z}^0 \end{cases} \tag{8}$$

where

$$A_i\left(\underline{x}\right) = \left(a_{i,jk}\left(\underline{x}\right)\right)_{m \times m} . \tag{9}$$

The solution to a partial differential equation system is not necessary to exist. The following consistent theorem is a necessary condition for a partial equation system to have a solution.

Theorem 5.2.1. ([3]) *Assume that functions* $\underline{f}_i\left(\underline{z}, \underline{x}\right)$ *are continuously differentiable with respect to* $\underline{x}$ *and* $\underline{z}$ *respectively in a domain* $G \subset \mathbb{R}^m \times \mathbb{R}^n$. *Then the equation system has a solution for arbitrary initial data if and only if the following consistency conditions are satisfied*

$$\frac{\partial \underline{f}_i}{\partial x_j} + \frac{\partial \underline{f}_i}{\partial \underline{z}} \underline{f}_j = \frac{\partial \underline{f}_j}{\partial x_i} + \frac{\partial \underline{f}_j}{\partial \underline{z}} \underline{f}_i, \ \forall i \neq j = 1, 2, \cdots, n . \tag{10}$$

In addition, the solution is unique on the domain where it is defined. In the linear case, the solution is defined on the whole domain $D \subset \mathbb{R}^n$ *, where the coefficients and free terms are defined, provided the domain is surface-simply connected.*

Corollary 5.2.2. *For a linear partial differential equation system, the consistency conditions can be stated as*

$$\begin{cases} A_i A_j + \dfrac{\partial}{\partial x_j} A_i = A_j A_i + \dfrac{\partial}{\partial x_i} A_j \\ A_i \underline{b}_j + \dfrac{\partial}{\partial x_j} \underline{b}_i = A_j \underline{b}_i + \dfrac{\partial}{\partial x_i} \underline{b}_j \end{cases} \tag{11}$$

where $i \neq j = 1, 2, \cdots, n$.

5.3 The Consistency Conditions for A Bivariate PDE Model

Let a bivariate PDE takes the form

$$\begin{cases} \dfrac{\partial z}{\partial x} = \alpha_1(x,y)z + \beta_1(x,y) \\ \dfrac{\partial z}{\partial y} = \alpha_2(x,y)z + \beta_2(x,y). \end{cases} \tag{12}$$

Now let us investigate the formation of Eq. (12) satisfying the Corollary to Forbenius Theorem. Note that

$$\begin{aligned} A_i &= \alpha_1(x,y) \\ A_j &= \alpha_2(x,y) \\ b_i &= \beta_1(x,y) \\ b_j &= \beta_2(x,y) \\ x_i &= x \\ x_j &= y. \end{aligned} \tag{13}$$

For the first condition in (11)

$$\begin{aligned} & A_i A_j + \frac{\partial}{\partial x_j} A_i = A_j A_i + \frac{\partial}{\partial x_i} A_j \\ & \text{LHS}: \alpha_1(x,y)\alpha_2(x,y) + \frac{\partial}{\partial y}\alpha_1(x,y) \\ & \text{RHS}: \alpha_2(x,y)\alpha_1(x,y) + \frac{\partial}{\partial x}\alpha_2(x,y) \end{aligned} \tag{14}$$

which leads to the condition

$$\frac{\partial}{\partial y}\alpha_1(x,y) = \frac{\partial}{\partial x}\alpha_2(x,y). \tag{15}$$

As to the second condition in (11)

$$\begin{aligned} & A_i \underline{b}_j + \frac{\partial}{\partial x_j} \underline{b}_i = A_j \underline{b}_i + \frac{\partial}{\partial x_i} \underline{b}_j \\ & \text{LHS}: \alpha_1(x,y)\beta_2(x,y) + \frac{\partial}{\partial y}\beta_1(x,y) \\ & \text{RHS}: \alpha_2(x,y)\beta_1(x,y) + \frac{\partial}{\partial x}\beta_2(x,y) \end{aligned} \tag{16}$$

which leads to a fairly complicated condition

$$\alpha_1(x,y)\beta_2(x,y) + \frac{\partial}{\partial y}\beta_1(x,y) = \alpha_2(x,y)\beta_1(x,y) + \frac{\partial}{\partial x}\beta_2(x,y). \tag{17}$$

Combining Eq. (15) and Eq. (17), the consistency conditions can be expressed by

$$\begin{cases} \dfrac{\partial}{\partial y}\alpha_1(x,y) = \dfrac{\partial}{\partial x}\alpha_2(x,y) \\ \alpha_1(x,y)\beta_2(x,y) + \dfrac{\partial}{\partial y}\beta_1(x,y) = \alpha_2(x,y)\beta_1(x,y) + \dfrac{\partial}{\partial x}\beta_2(x,y). \end{cases} \tag{18}$$

6 The PDEMR Model Formation

Similar to DEMR model, PDEMR model is also constituted by two components: motivated partial differential equation (abbreviated as PDE) systems and coupled (multivariate) regression model. Let us use the linear PDE motivated regression for the basic definition.

Definition 6.1. *Coupled equation system*

$$\begin{cases} \dfrac{\partial \underline{z}}{\partial x_i} = A_i\left(\underline{x}\right)\underline{z} + \underline{b}\left(\underline{x}\right),\ i = 1, 2, \cdots, m \\ \underline{z}\left(\underline{x}^0\right) = \underline{z}^0 \\ \Delta^{\partial}_{x_i(k_i)}\underline{z} = A_i\left(\underline{x}\left(\underline{k}_i\right)\right)\underline{z}\left(\left(\underline{k}_i\right)\right) + \underline{b}\left(\underline{x}\left(\underline{k}_i\right)\right) \end{cases} \tag{19}$$

where

$$\begin{aligned} \Delta^{\partial}_{x_i(k_i)}\underline{z} &= \underline{z}\left(x_1\left(k_1\right), x_2\left(k_2\right), \cdots, x_i\left(k_i\right), \cdots, x_m\left(k_m\right)\right) \\ &\quad - \underline{z}\left(x_1\left(k_1\right), x_2\left(k_2\right), \cdots, x_i\left(k_i - 1\right), \cdots, x_m\left(k_m\right)\right) \end{aligned} \tag{20}$$

denotes the (first) partial difference of $\underline{z}\left(x_1, x_2, \cdots, x_m\right)$ *with respect to exploratory variable* x_i *at point* $\left(x_1\left(k_1\right), x_2\left(k_2\right), \cdots, x_i\left(k_i\right), \cdots, x_m\left(k_m\right)\right)$.

7 A PDEMR Modelling of Protea Frequency Count Spatial Distribution

7.1 A Bivariate Partial Differential Equation for Log-Count

Bear in mind that we intend to develop a counting model for filling those sites where the counts of a particular class was recorded as zero, typically is in the design note for observation and sampling data collection, however, was not attended for some technical reasons. The count z is a (integer) scalar function of coordinate (x, y) and thus it may be appropriate to us the log-transformation, i.e., $u(x, y) = \ln z(x, y)$. Note that

$$\begin{cases} \dfrac{\partial u\left(x, y\right)}{\partial x} = \dfrac{1}{z\left(x, y\right)} \dfrac{\partial z\left(x, y\right)}{\partial x} \\ \dfrac{\partial u\left(x, y\right)}{\partial y} = \dfrac{1}{z\left(x, y\right)} \dfrac{\partial z\left(x, y\right)}{\partial y}. \end{cases} \tag{21}$$

To obtain the insight of bivariate PDE model, we start with a bivariate partial differential equation system in the form of Eq. (22)

$$\begin{cases} \dfrac{\partial u}{\partial x} = \alpha_1 + 2\alpha_3 x + \alpha_4 y \\ \dfrac{\partial u}{\partial y} = \alpha_2 + \alpha_4 x + 2\alpha_5 y. \end{cases} \tag{22}$$

It is obvious that

$$\begin{aligned} &A_i\left(x, y\right) = \alpha_1 + 2\alpha_3 x + \alpha_4 y \\ &A_j\left(x, y\right) = \alpha_2 + \alpha_4 x + 2\alpha_5 y \\ &\beta_1\left(x, y\right) = 0 \\ &\beta_2\left(x, y\right) = 0. \end{aligned} \tag{23}$$

Accordingly, a homogeneous equation system is obtained and it is easy to check that the homogenous equation system Eq. (22) satisfies the consistency conditions set up in Corollary 5.2.2.

The matrix form of Eq. (22) can be written as

$$\begin{bmatrix} \dfrac{\partial u}{\partial x} \\ \dfrac{\partial u}{\partial y} \end{bmatrix} = \begin{bmatrix} \alpha_1 & 2\alpha_3 & \alpha_4 \\ \alpha_2 & \alpha_4 & 2\alpha_5 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \end{bmatrix} \tag{24}$$

or taking the transpose for both sides of Eq. (24),

$$\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & x & y \end{bmatrix} \begin{bmatrix} \alpha_1 & \alpha_2 \\ 2\alpha_3 & \alpha_4 \\ \alpha_4 & 2\alpha_5 \end{bmatrix}. \tag{25}$$

Let the parameter matrix be

$$\Lambda = \begin{bmatrix} \alpha_1 & \alpha_2 \\ 2\alpha_3 & \alpha_4 \\ \alpha_4 & 2\alpha_5 \end{bmatrix}. \tag{26}$$

The design matrix (a vector) is denoted as

$$\underline{X} = \begin{bmatrix} 1 \\ x \\ y \end{bmatrix}. \tag{27}$$

And the partial derivative vector is denoted as

$$\frac{\partial u}{\partial \underline{x}} = \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \end{bmatrix}. \tag{28}$$

Finally we have a matrix representation of the bivariate partial equation system Eq. (22)

$$\left(\frac{\partial u}{\partial \underline{x}} \right)^T = \underline{X}^T \Lambda . \tag{29}$$

7.2 The Divided Difference and Its Application in Approximating Partial Derivatives

The key step for PDEMR model setting is the translation from partial derivatives into partial differences. It is often the case that the observations are not equal-gap taken, but on the contrary. In bivariate circumstances, the way for defining difference for unequal-gapped data is even more complicated than that in one-dimensional case. Therefore, we intend to develop a scheme of the obtaining "best" partial difference for approximating the corresponding partial derivatives.

1. Divided difference.

Definition 7.2.1. *Given a function* $f(x)$ *on the interval [a,b]. Let the sequence* $\{x_1, x_2, \cdots, x_l\}$ *with* $\forall x_i \in [a,b]$ *and* $x_i < x_j$ *for any i<j. Then the quantity*

$$\Delta_{x_i}^{\partial} f \triangleq \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} \tag{30}$$

is called the (first) divided difference for function $f(\cdot)$ *at* x_i *.*

2. Partial divided difference.

Definition 7.2.2. *Given a bivariate function* $w(x,y)$ *on* $D \subset \mathbb{R}^2$*. Let* $(x_i, y_j) \in D$*. Then*

$$\Delta_{x_i}^{\partial} w = \frac{w(x_i, y_j) - w(x_{i-1}, y_j)}{x_i - x_{i-1}} \tag{31}$$

is defined as a (first) partial difference of $w(\cdot,\cdot)$ *with respect to exploratory variable x at* (x_i, y_j)*. Similarly,*

$$\Delta_{y_j}^{\partial} w = \frac{w(x_i, y_j) - w(x_i, y_{j-1})}{y_j - y_{j-1}} \tag{32}$$

is defined as the partial difference of $w(\cdot,\cdot)$ *with respect to exploratory variable y at* (x_i, y_j).

3. Using directional derivative for least-square estimated partial divided difference.

Let D be a sub-space of a 2-dimensional space, $\mathbb{R}\times\mathbb{R}$, any point of D, denoted as $M(x, y)$ corresponds to the value of a scalar function $s(x, y)$, if the position of M could be represented by a vector $\vec{r}$, then scalar function can be regarded as a function of variable vector $\vec{r}$, i.e., $s = s(\vec{r})$.

Definition 7.2.3. *Let*

$$\text{grad } s = \nabla s \triangleq \frac{\partial s}{\partial x}\vec{i} + \frac{\partial s}{\partial y}\vec{j} \tag{33}$$

be the gradient of scalar field $s(x, y)$ *at point* (x, y). *Let* $\vec{l}$ *be a unit directional vector with directional angular* θ_x *and* θ_y *such that*

$$\left|\vec{l}\right| = \cos^2(\theta_x) + \cos^2(\theta_y) = 1\,. \tag{34}$$

Then

$$\frac{\partial s}{\partial l} = \vec{l}\cdot \text{grad } s = \frac{\partial s}{\partial x}\cos(\theta_x) + \frac{\partial s}{\partial y}\cos(\theta_y) \tag{35}$$

is called the directional derivative with respect to directional vector $\vec{l}$ *at point* (x, y).

Let $\mathbb{k}_r(x, y) \subset D$ be a neighborhood of radius r, i.e., for any $(x_i, y_j) \in \mathbb{k}_r(x, y)$ the distances of (x_i, y_j) from point (x, y): $\sqrt{(x_i - x)^2 + (y_j - y)^2} < r$.

However, unless the functional form of the scalar field is available, then we can not obtain the accurate values of the directional derivatives. However, for each direction, $(x, y) \to (x_i, y_j)$, an approximate directional derivative can be calculated as

$$\frac{\widehat{\partial\varphi}}{\partial l} = \frac{s(x, y) - s(x_i, y_j)}{\sqrt{(x - x_i)^2 + (y - y_j)^2}}\,. \tag{36}$$

Furthermore, the cosines of the directional angular are also calculated as

$$\begin{cases} \cos(\theta_x) = \dfrac{x - x_i}{\sqrt{(x - x_i)^2 + (y - y_j)^2}} \\ \cos(\theta_y) = \dfrac{y - y_j}{\sqrt{(x - x_i)^2 + (y - y_j)^2}}. \end{cases} \tag{37}$$

Therefore, the $(x, y) \to (x_i, y_j)$ pair of point will generate an equation

$$\Delta_x^{\partial} s \cos(\theta_x) + \Delta_y^{\partial} s \cos(\theta_y) = \left(\frac{\widehat{\partial\varphi}}{\partial l}\right)_{(x,y)}^{(x_i, y_j)}. \tag{38}$$

In general, there will be $k(k-1)/2$ equations in total if there are k points in $\mathbb{k}_r(x, y) \subset D$, which overspecify the two unknown partial differences, Δ_x^{∂} and Δ_y^{∂} at (x, y) respectively. As a matter of fact, the partial differences will be least-square estimate.

7.3 The Coupled Bivariate Regression Model

Once the partial differences, either direct divided estimate or the least-square estimate defined in 7.2, Δ_x^{∂} and Δ_y^{∂} are ready for further analysis, let

$$\Delta_{n\times 2}=\begin{bmatrix}\Delta_{x_1}^{\partial}u & \Delta_{y_1}^{\partial}u\\ \Delta_{x_2}^{\partial}u & \Delta_{y_2}^{\partial}u\\ \vdots & \vdots\\ \Delta_{x_n}^{\partial}u & \Delta_{y_n}^{\partial}u\end{bmatrix},\quad X_{n\times 3}=\begin{bmatrix}1 & x_1 & y_1\\ 1 & x_2 & y_2\\ \vdots & \vdots & \vdots\\ 1 & x_n & y_n\end{bmatrix}$$

$$\Lambda_{3\times 2}=\begin{bmatrix}\alpha_1 & \alpha_2\\ 2\alpha_3 & \alpha_4\\ \alpha_4 & 2\alpha_5\end{bmatrix},\quad \mathrm{E}_{n\times 2}=\begin{bmatrix}e_{11} & e_{12}\\ e_{21} & e_{22}\\ \vdots & \vdots\\ e_{n1} & e_{n2}\end{bmatrix}. \tag{39}$$

Then the coupled regression model in matrix form will be

$$\Delta = X\Lambda + E. \tag{40}$$

Finally, the bivariate PDEMR model for the log-count will be

$$\begin{cases}\begin{bmatrix}\dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y}\end{bmatrix}=\begin{bmatrix}1 & x & y\end{bmatrix}\begin{bmatrix}\alpha_1 & \alpha_2\\ 2\alpha_3 & \alpha_4\\ \alpha_4 & 2\alpha_5\end{bmatrix}\\ \Delta = X\Lambda + E.\end{cases} \tag{41}$$

As to the error structure of the PDEMR formation in Eq. (41), we will investigate in the next subsection.

7.4 The Normal Random Fuzzy Error Structure and Its Estimation

Following multivariate regression modeling theory, it is assumed that the error matrix E in the bivariate regression model in Eq. (40) as follows

$$\mathrm{E}_{n\times 2}=\left[\underset{\sim}{e}_1,\underset{\sim}{e}_2\right]=\begin{bmatrix}e_{11} & e_{12}\\ e_{21} & e_{22}\\ \vdots & \vdots\\ e_{n1} & e_{n2}\end{bmatrix}. \tag{42}$$

Typically, the m observations on the j^{th} trial have correlation matrix $\Sigma=\left(\sigma_{ij}\right)$ and

$$\begin{aligned}&\mathrm{E}\left[\underset{\sim}{e}_i\right]=\underline{0},\ i=1,2\\ &\mathrm{Cov}\left[\underset{\sim}{e}_1,\underset{\sim}{e}_2\right]=\sigma_{12}I.\end{aligned} \tag{43}$$

In other words, the i^{th} "response" follows the linear regression model having error vector $\underset{\sim}{e}_i$ with $\mathrm{Cov}\left[\underset{\sim}{e}_i\right]=\sigma_{ii}I$. However, the errors for different "response" on the same "trial" may be correlated (Johnson and Wichern [9]).

Furthermore, for the coupled bivariate regression model, it is assumed that the error vector is normal random fuzzy vector, which is a sum of a fuzzy vector, denoted as $\left(m_1,m_2\right)^T$ and a normal random vector having mean zero and variance-covariance matrix $\Sigma=\left(\sigma_{ij}\right)$. Then for any given fixed value pair $\left(m_1,m_2\right)^T$, the error vector has joint distribution

$$f\left(x_1,x_2\right)=\frac{1}{2\pi\left|\Sigma\right|^{1/2}}\exp\left(-\begin{pmatrix}x_1-m_1\\ x_2-m_2\end{pmatrix}^T\Sigma^{-1}\begin{pmatrix}x_1-m_1\\ x_2-m_2\end{pmatrix}\right). \tag{44}$$

As to the fuzzy vector, $(m_1, m_2)^T$ we propose a bi-triangular joint membership function similar to univariate normal random fuzzy variable case

$$\mu_{(m_1,m_2)}(z_1,z_2) = \begin{cases} \dfrac{z_1-a_1}{b-a_1} + \dfrac{b-z_2}{b-a_2} & \text{if } (z_1,z_2)\in A_1 \\ \dfrac{c_1-z_1}{c_1-b} + \dfrac{b-z_2}{c_2-b} & \text{if } (z_1,z_2)\in A_2 \\ 0, & \text{otherwise} \end{cases} \tag{45}$$

where the domain areas in Eq. (45) are defined by

$$\begin{aligned} A_1 &= \left\{(z_1,z_2): a_1 \le z_1 \le b, a_2 \le z_2 \le b, \frac{z_1}{a_1} + \frac{z_2}{a_2} \ge 1\right\} \\ A_2 &= \left\{(z_1,z_2): b \le z_1 \le c_1, b \le z_2 \le c_2, \frac{z_1}{c_1} + \frac{z_2}{c_2} \le 1\right\} \\ A_3 &= \left\{(z_1,z_2): z_1 \ge c_1, z_2 \ge c_2, \frac{z_1}{c_1} + \frac{z_2}{c_2} \ge 1\right\}. \end{aligned} \tag{46}$$

The following figure illustrates the shape of the bi-triangular membership function proposed.

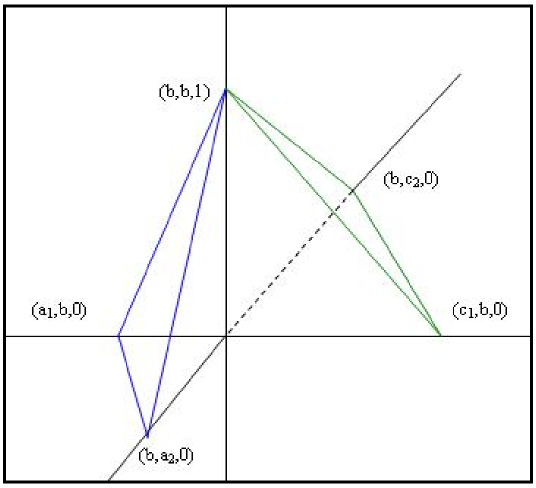


Figure 5. Bi-triangular joint membership function

The chance measure for the bi-normal random fuzzy vector can be derived as an extension to Model I (see Liu [11]). Let event $B = \{\xi_1 \le x_1, \xi_2 \le x_2\} \in \mathfrak{B}(\mathbb{R}^2)$, then the chance measure is

$$\mathrm{Ch}\left\{\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix} \in B\right\} = \begin{cases} \sup\limits_{\underline{x}}\left(\dfrac{\mu_{\tilde{\underline{m}}}(\underline{x})}{2} \wedge \displaystyle\int_{\tilde{\underline{m}}+\underline{\eta}\in B} f_{\underline{\eta}}(\underline{x}+\underline{y})d\underline{y}\right), & \text{if } \sup\limits_{\underline{x}}\left(\dfrac{\mu_{\tilde{\underline{m}}}(\underline{x})}{2} \wedge \displaystyle\int_{\tilde{\underline{m}}+\underline{\eta}\in B} f_{\underline{\eta}}(\underline{x}+\underline{y})d\underline{y}\right) < 0.5 \\ 1-\sup\limits_{\underline{x}}\left(\dfrac{\mu_{\tilde{\underline{m}}}(\underline{x})}{2} \wedge \displaystyle\int_{\tilde{\underline{m}}+\underline{\eta}\in B^c} f_{\underline{\eta}}(\underline{x}+\underline{y})d\underline{y}\right), & \text{if } \sup\limits_{\underline{x}}\left(\dfrac{\mu_{\tilde{\underline{m}}}(\underline{x})}{2} \wedge \displaystyle\int_{\tilde{\underline{m}}+\underline{\eta}\in B} f_{\underline{\eta}}(\underline{x}+\underline{y})d\underline{y}\right) \ge 0.5 \end{cases} \tag{47}$$

where bi-normal random fuzzy vector $\underline{\xi} = \tilde{\underline{m}} + \underline{\eta}$. Without any doubts, the Model I type of chance measure looks very neat, but the estimation procedure would be difficult to handle.

For parameter estimation purpose, we will still intend to address the relevant average chance measure. At our current mathematical manipulation level, it is impossible to derive a bivariate average chance measure for the bi-triangular membership fuzzy mean vector and bivariate normal distribution with mean zero and variance-covariance matrix Σ. Nevertheless, for the error vector fuzzy component,

$$\mu_{(m_1,m_2)}(z_1,z_2)=\begin{cases}\frac{z_1+h_1}{h_1}-\frac{z_2}{h_2}, & \text{if } (z_1,z_2)\in A_1\\ \frac{h_1-z_1}{h_1}-\frac{z_2}{h_2}, & \text{if } (z_1,z_2)\in A_2\\ 0, & \text{otherwise}\end{cases} \tag{48}$$

the bivariate normal random component is distributed as

$$f(x_1,x_2)=\frac{1}{2\pi\left\|\begin{bmatrix}\sigma_1 & \sigma_{12}\\ \sigma_{12} & \sigma_2\end{bmatrix}\right\|^{1/2}}\exp\left(-\begin{pmatrix}x_1\\ x_2\end{pmatrix}^T\begin{bmatrix}\sigma_1 & \sigma_{12}\\ \sigma_{12} & \sigma_2\end{bmatrix}^{-1}\begin{pmatrix}x_1\\ x_2\end{pmatrix}\right). \tag{49}$$

Note that he parameters to be estimated are (h_1,h_2) and $(\sigma_1,\sigma_{12},\sigma_2)$. Therefore, the estimation procedure for the bivariate estimation problem can be converted into a three-step univariate average chance estimations:

Step 1. Perform the Maximum Average Chance estimation on parameter (h_1,σ_1) utilizing the bivariate regression error vector $\underset{\sim}{\hat{e}}_1$, which in component level is $e_{1j}\sim N\left(\tilde{\zeta}_1,\sigma_1\right), j=1,2,\cdots,n$ being assumed to be independent normal random fuzzy variables. The form of the average chance density is given in Appendix Eq. (A12).

Step 2. Perform the Maximum Average Chance estimation on parameter (h_2,σ_2) utilizing the bivariate regression error vector $\underset{\sim}{\hat{e}}_2$, which in component level is $e_{2j}\sim N\left(\tilde{\zeta}_2,\sigma_2\right), j=1,2,\cdots,n$ being assumed to be independent normal random fuzzy variables. The form of the average chance density is given in Appendix Eq. (A12) too.

Step 3. Perform a Maximum Average Chance estimation on parameter σ_{12} utilizing the "data" of the sum of two error vectors, i.e., $\underset{\sim}{\hat{e}}_1+\underset{\sim}{\hat{e}}_2$ conditional on the estimated parameters $\left(\hat{h}_1,\hat{h}_2,\hat{\sigma}_1,\hat{\sigma}_2\right)$ obtained from Step 1 and Step 2. The reason behind it is the fuzzy mean of the component error sum vector is $\zeta_{1j}+\zeta_{2j}, j=1,2,\cdots,n$ and the variance $VAR\left(\hat{e}_{1j}+\hat{e}_{2j}\right)=\sigma_1+2\sigma_{12}+\sigma_2$. The fuzzy membership for $\zeta_{1j}+\zeta_{2j}, j=1,2,\cdots,n$ are defined by $(-(h_1+h_2),0,h_1+h_2)$, while the normal component is defined as $N(0,\sigma_1+2\sigma_{12}+\sigma_2)$. Thus the average chance density takes a form

$$\begin{aligned}\psi(x)=&\frac{1}{2(h_1+h_2)}\left(\Phi\left(\frac{x+(h_1+h_2)}{\sqrt{\sigma_1+2\sigma_{12}+\sigma_2}}\right)-\Phi\left(\frac{x-(h_1+h_2)}{\sqrt{\sigma_1+2\sigma_{12}+\sigma_2}}\right)\right)+\frac{1}{\sqrt{\sigma_1+2\sigma_{12}+\sigma_2}}\phi\left(\frac{x-(h_1+h_2)}{\sqrt{\sigma_1+2\sigma_{12}+\sigma_2}}\right)\\ &-\frac{1}{2(h_1+h_2)}\left(\left(\frac{x+(h_1+h_2)}{\sqrt{\sigma_1+2\sigma_{12}+\sigma_2}}\right)\phi\left(\frac{x+(h_1+h_2)}{\sigma}\right)-\left(\frac{x-(h_1+h_2)}{\sqrt{\sigma_1+2\sigma_{12}+\sigma_2}}\right)\phi\left(\frac{x-(h_1+h_2)}{\sqrt{\sigma_1+2\sigma_{12}+\sigma_2}}\right)\right)\\ &+\frac{x+(h_1+h_2)}{2(h_1+h_2)\sqrt{\sigma_1+2\sigma_{12}+\sigma_2}}\left(\phi\left(\frac{x+(h_1+h_2)}{\sqrt{\sigma_1+2\sigma_{12}+\sigma_2}}\right)-\phi\left(\frac{x-(h_1+h_2)}{\sqrt{\sigma_1+2\sigma_{12}+\sigma_2}}\right)\right).\end{aligned} \tag{50}$$

As to the Maximum Average Chance Estimation procedure, we will state it in the subsection 7.5.

7.5 The Maximum Average Chance Estimation

In linear model theory, it is often assumed that the model error structure follows multivariate normal distribution and thus the likelihood function can be obtained. For random fuzzy variable theory, an average chance function can be defined (for theoretical details, see Appendix). We will propose a data-assimilating algorithm for determining the unknown parameters of the average chance distribution underlying the PDEMR model. We use data-assimilation to contrast with statistical estimation because while both methods determine unknown parameters in terms of sampling data, however, statistical estimation is performed under the hypothesized (random) population probability distribution, but the data-assimilation will be performed according to a chance distribution, particularly, the average chance distribution, which is not population probability distribution at all. In statistics, the commonly used principle is maximum likelihood estimation, where the estimated parameter(s) maximize the likelihood function. Parallel to maximum likelihood estimation, we will define average chance function according to a hypothesized random fuzzy population average chance distribution, and then for the data-assimilated parameter(s) we maximize the average chance function, which may be regarded as a counterpart of likelihood function.

Definition 7.1. *(Average chance function). Let* $\{x_1, x_2, \cdots, x_n\}$ *be a simple random sample drawing from a given population with assumed probability distribution* $F(x;\underline{\theta})$ *, where parameter-vector* $\underline{\theta} = (\theta_1, \theta_2, , \theta_\gamma)^T$ *,* $\gamma \geq 1$ *and parameter component* θ_0 *is a fuzzy variable with credibility distribution* $\Lambda_{\theta_0}(y)$ *defined by parameter-vector* $\underline{\rho}$ *, and whose average chance distribution is derivable and denoted as* $\Psi(x)$*. Then the joint average chance density, denoted as*

$$C(\underline{\rho}^T, \theta_2, \cdots, \theta_\gamma \mid \{x_1, x_2, \cdots, x_n\} = \prod_{i=1}^{n} \psi\left(x_i \mid \underline{\rho}^T, \theta_2, \cdots, \theta_\gamma\right) \tag{51}$$

is called the (average) chance function. Similarly to log-likelihood function, the function

$$\begin{aligned} & l_C\left(\underline{\rho}^T, \theta_2, \cdots, \theta_\gamma \mid \{x_1, x_2, \cdots, x_n\}\right) \\ &= \ln C\left(\underline{\rho}^T, \theta_2, \cdots, \theta_\gamma \mid \{x_1, x_2, \cdots, x_n\}\right) \\ &= \sum_{i=1}^{n} \ln\left(\psi\left(x_i \mid \underline{\rho}^T, \theta_2, \cdots, \theta_\gamma\right)\right) \end{aligned} \tag{52}$$

is called the log-chance function given the simple random sample $\{x_1, x_2, \cdots, x_n\}$.

Maximum Average Chance Principle: *For a given simple random sample* $\{x_1, x_2, \cdots, x_n\}$ *, the optimal data-assimilated parameter-vector* $\left(\underline{\rho}^T, \theta_2, \cdots, \theta_\gamma\right)^T$ *maximizes the average chance function or equivalently, maximizes the log-chance function.*

Let us investigate the maximum average chance estimate for a normal random fuzzy variable, under a triangular credibility fuzzy mean with parameters (a, b, c) and fixed variance parameter σ^2.

In a full data-assimilated parameter estimation of the coupled regression model specified by the univariate model, we have that

$$\hat{e}_i = y_i - \underline{x_i}\underline{\alpha} \sim N\left(\tilde{e}, \sigma^2\right) \tag{53}$$

where

$$\underline{x_i} = \left(1, \hat{x}(i), \Delta x(i), \cdots, \Delta^{p-1} x(i)\right) \tag{54}$$

and $y_i = \Delta^p x(i)$, therefore the contribution of i^{th} sample element to the average chance function is

$$\begin{aligned} \psi(x_i) = & \frac{1}{2(b-a)}\left(\Phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - a}{\sigma}\right) - \Phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - b}{\sigma}\right)\right) \\ & + \frac{(y_i - \underline{x_i}\underline{\alpha}) - a}{2(b-a)\sigma}\left(\phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - a}{\sigma}\right) - \phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - b}{\sigma}\right)\right) \\ & + \frac{1}{2(c-b)}\left(\Phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - b}{\sigma}\right) - \Phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - c}{\sigma}\right)\right) \\ & + \frac{(y_i - \underline{x_i}\underline{\alpha}) + c - 2b}{2(c-b)\sigma}\left(\phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - b}{\sigma}\right) - \phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - c}{\sigma}\right)\right) + \frac{1}{\sigma}\phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - c}{\sigma}\right) \\ & - \frac{1}{2(b-a)}\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - a}{\sigma}\phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - a}{\sigma}\right) - \frac{(y_i - \underline{x_i}\underline{\alpha}) - b}{\sigma}\phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - b}{\sigma}\right)\right) \\ & - \frac{1}{2(c-b)}\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - b}{\sigma}\phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - b}{\sigma}\right) - \frac{(y_i - \underline{x_i}\underline{\alpha}) - c}{\sigma}\phi\left(\frac{(y_i - \underline{x_i}\underline{\alpha}) - c}{\sigma}\right)\right). \end{aligned} \tag{55}$$

Let (a, b, c) specify fuzzy mean $\tilde{e}$, and σ^2 the variance. Then the full log-chance function is

$$l_C\left(\left(a,b,c\right),\sigma \mid \left\{x_1,x_2,\cdots,x_n\right\}\right)=\sum_{i=1}^{n}\ln\left(\psi\left(x_i \mid \left(a,b,c\right),\sigma\right)\right). \tag{56}$$

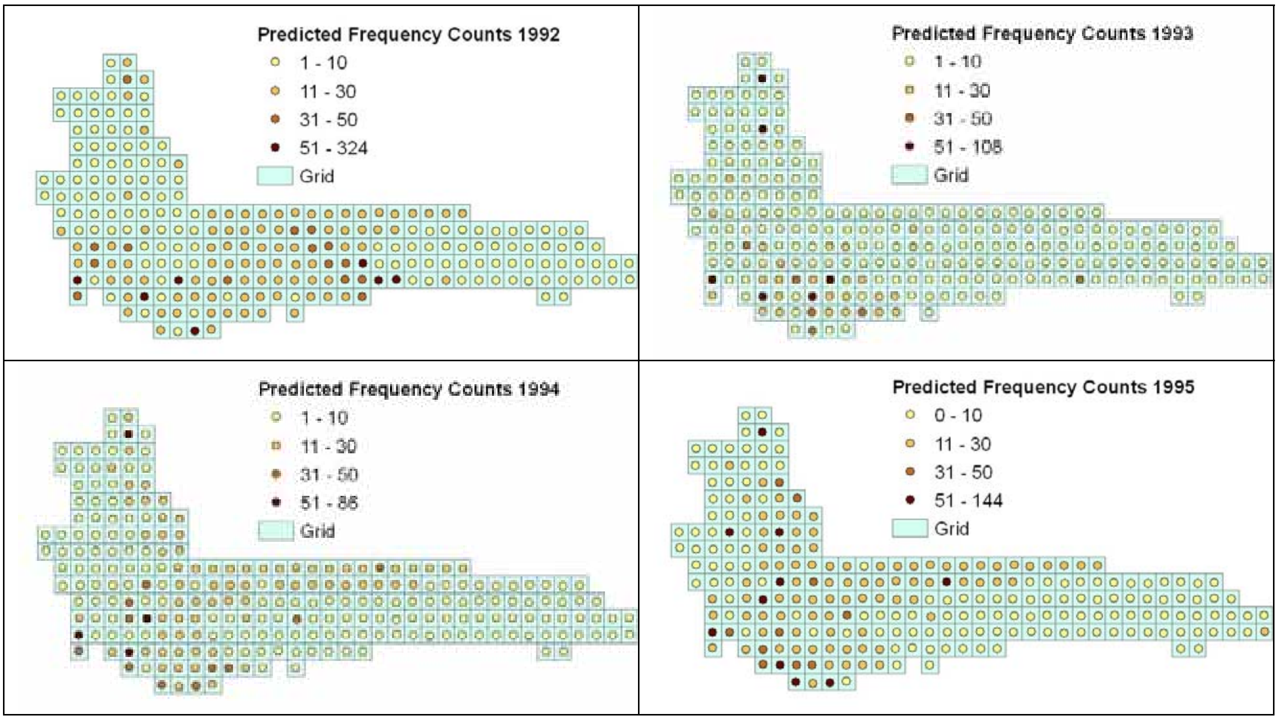
Thus the search the unknown parameters as an optimization problem may be converted into the problem of solving 4 nonlinear equation system as follows

$$\begin{cases}\dfrac{\partial l_C\left(\underline{\alpha},\ \left(a,b,c\right),\sigma \mid \left\{x_1,x_2,\cdots,x_n\right\}\right)}{\partial a}=0\\ \dfrac{\partial l_C\left(\underline{\alpha},\ \left(a,b,c\right),\sigma \mid \left\{x_1,x_2,\cdots,x_n\right\}\right)}{\partial b}=0\\ \dfrac{\partial l_C\left(\underline{\alpha},\ \left(a,b,c\right),\sigma \mid \left\{x_1,x_2,\cdots,x_n\right\}\right)}{\partial c}=0\\ \dfrac{\partial l_C\left(\underline{\alpha},\ \left(a,b,c\right),\sigma \mid \left\{x_1,x_2,\cdots,x_n\right\}\right)}{\partial \sigma}=0.\end{cases} \tag{57}$$

For the case of an isosceles triangular membership function (shown in Appendix Eq. (A11)) the number of parameters to be estimated reduced to 2, i.e., $\left(h_1,\sigma_1\right)$ for **Step 1** and $\left(h_2,\sigma_2\right)$ **Step 2,** respectively, and in Subsection **7.4** and accordingly the average chance density is given in Appendix Eq. (A12). As to **Step 3**, the parameter left to be estimated are σ_{12} because $\left(\hat{h}_1,\hat{h}_2,\hat{\sigma}_1,\hat{\sigma}_2\right)$ are obtained in **Step 1** and **Step 2**.

8 PDEMR Predicted Protea Frequency Counts

Using the predicted results from the PDEMR model, the un-sampled cells are predicted with frequency counts of the Protea. Figure 6 shows the predicted frequency counts of Proteas in the population size of 10 to 100, in the Cape Floristic Region, from 1992 to 2002.



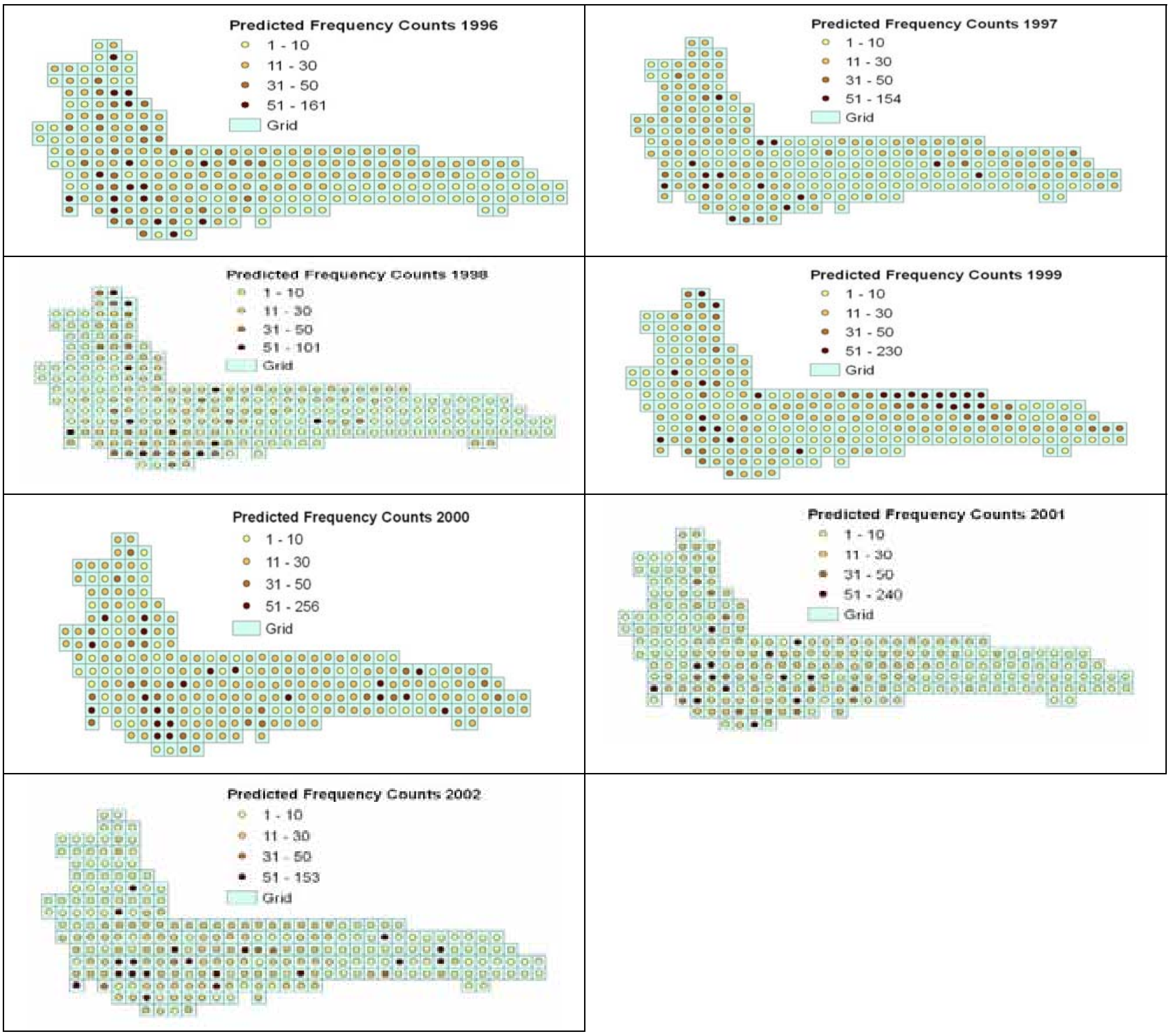
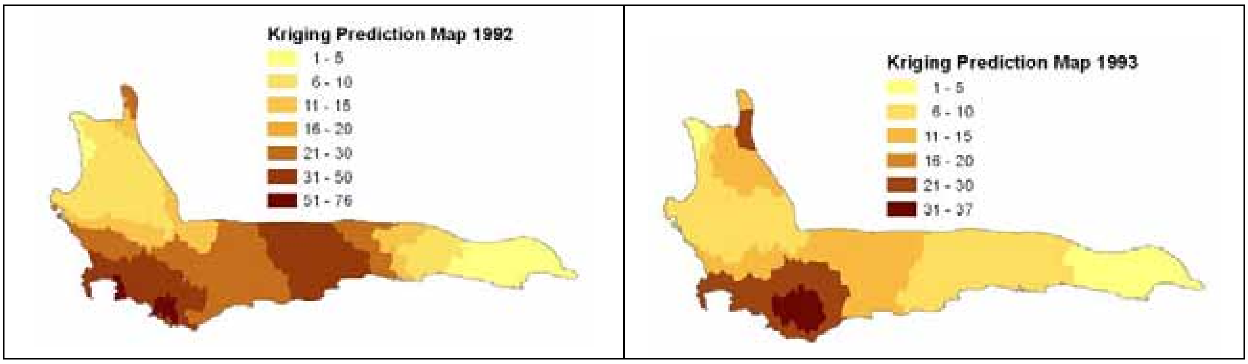
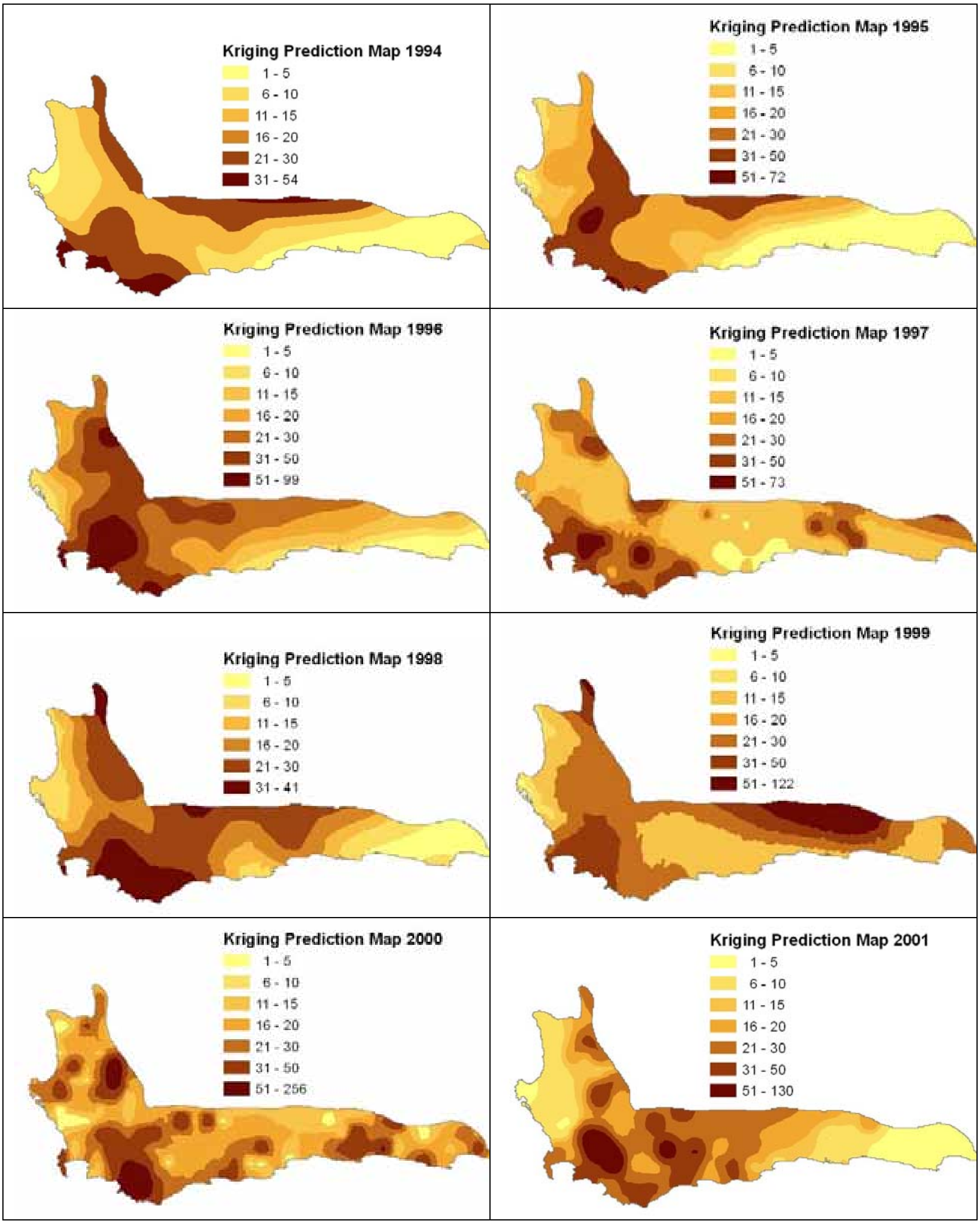


Figure 6. The PDEMR model predicted frequency counts of proteas in the population Size of 10-100, in the Cape Floristic Region, 1992-2002

Finally, we can produce kriging prediction maps of the Protea species, using the predicted results from the PDEMR model. Figure 7 shows the distribution and patterns of frequency counts of Proteas. One can see the changes in the density of occurrence of the Proteas in the Cape Floristic Region over the 11 years.




Kriging Prediction Map 1994
1 - 5
6 - 10
11 - 15
16 - 20
21 - 30
31 - 54
Kriging Prediction Map 1995
1 - 5
6 - 10
11 - 15
16 - 20
21 - 30
31 - 50
51 - 72
Kriging Prediction Map 1996
1 - 5
6 - 10
11 - 15
16 - 20
21 - 30
31 - 50
51 - 99
Kriging Prediction Map 1997
1 - 5
6 - 10
11 - 15
16 - 20
21 - 30
31 - 50
51 - 73
Kriging Prediction Map 1998
1 - 5
6 - 10
11 - 15
16 - 20
21 - 30
31 - 41
Kriging Prediction Map 1999
1 - 5
6 - 10
11 - 15
16 - 20
21 - 30
31 - 50
51 - 122
Kriging Prediction Map 2000
1 - 5
6 - 10
11 - 15
16 - 20
21 - 30
31 - 50
51 - 256
Kriging Prediction Map 2001
1 - 5
6 - 10
11 - 15
16 - 20
21 - 30
31 - 50
51 - 130

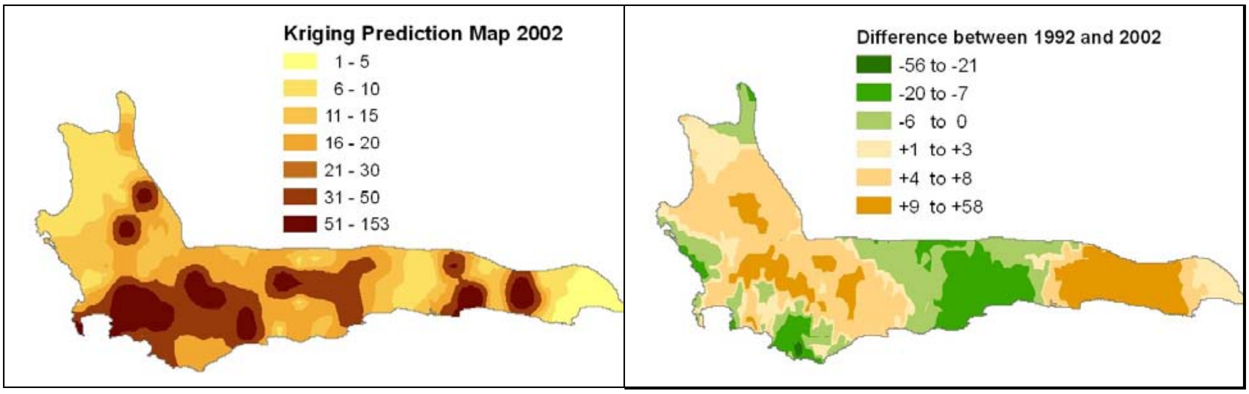


Figure 7. The kriging prediction maps of frequency counts of proteas in the population size of 10-100, in the Cape Floristic Region, 1992-2002

The light colours in Figure 6 referred to low counts and the dark colours referred to high counts of the Proteas. It is interesting to note that the kriging prediction maps of 1992 and 2002 are very different, and the frequency counts of the Proteas has increased in 2002. The pattern of distribution are also very different, the high counts of Proteas in 2002 seems to occur in small fragmented areas. The final map in Figure 6 shows the difference between 1992 and 2002, one can that there are areas of positive changes and areas of negative changes, over the 11 years.

9 Conclusion

In this paper, we solved two crucial problems with regard to the ecological dataset, presence data only and incomplete sample data. We used the *partial differential equation motivated regression* (PDEMR) model, which merges the partial differential equation theory, (statistical) linear model theory and credibility measure theory together. The coupled regression component in a PDEMR model is in nature a special random fuzzy multivariate regression model. We developed a bivariate model for prediction of the Protea species in the population size of 10 to 100, in the Cape Floristic Region, 1992 to 2002, in South Africa. The model has produced very good results, which helped to produce kriging prediction maps. The spatial distribution and pattern are clear to see and understand in the kriging maps.

Finally, it is necessary to pointed out that conceptually the motivated partial equation is common one as in partial differential equation literature. The parameters in the motivated differential equation are real-valued numbers. However, after the coupled multivariate regression, and even further the maximum average chance estimation for the error structure, the estimated parameters are random fuzzy in nature. However, in this paper we will not facilitate further details because this will depend on our research on the asymptotic analysis of the Maximum Average Chance estimators in the near future.

Acknowledgements

We would like to thank the Protea Atlas Project for providing the species data for our research.
This research is partially supported by South African National Research Foundation Grant FA2006042800024.

References

[1] Deng, J. L., *Grey Systems (Social Economical)*, The Publishing House of Defense Industry, Beijing, China, 1985. (in Chinese)

[2] Freeth, R., B. Bomhard and Midgley, G., *Adapting to Climate Change in the Cape Floristic Region: Building Resilience of People and Plants in Protected Areas*, South African National Biodiversity Institute (SANBI), Climate Change Research Group, Kirstenbosch Research Center, 2007.

[3] Frobenius, G., Ueber das Pfaffsche problem, *J. Reine Angew. Math.*, vol.82, pp.230-315, 1877.

[4] Guo, D., R. Guo *et al*., DEMR modelling of brunsvigia litoralis plant distribution in western region district council, South Africa, *Proc. of the Sixth International Conference on Information and Management Sciences*, Lhsa, Tibet, China, July 1-6, 2007.

[5] Guo, D., R. Guo and C. Thiart, The coupling of regression modelling and differential equation model in GM(1,1) modelling and extended GM(1,1) models, *Journal of Grey System*, vol.9, no.2, pp.143-154, 2006.

[6] Guo, R., Repairable system modelling via grey differential equations, *Journal of Grey System*, vol.8, no.1, pp.69-92, 2005.

[7] Guo, R., D. Guo *et al*., Bivariate DEMR modelling of range shifts in aulax umbellata shrubs from 1992 to 2002 in South Africa, in *Proc. 6th International Conference on Information and Management Sciences*, Lhasa, Tibet, pp.783-791, 2007.

[8] Hakopian, H. A. and M.G. Tonoyan, Partial differential analogs of ordinary differential equations and systems, *New York Journal of Mathematics*, vol.10, pp.89-116, 2004.

[9] Johnson, R. A. and D. W. Wichern, *Applied Multivariate Statistical Analysis*, Prentice-Hall Inc, 1982.

[10] Liu, B. D., *Uncertainty Theory: An Introduction to Its Axiomatic Foundations*, Springer-Verlag Heidelberg, Berlin, 2004.

[11] Liu, B. D., *Uncertainty Theory*, 2nd Edition, Springer-Verlag, Berlin, 2007.

[12] Liu, Y. K. and B. D. Liu, Random fuzzy programming with chance measures defined by fuzzy integrals, *Mathematical and Computer Modelling*, vol.2, no.3, pp.249-256, 2002.

[13] Midgley, G. F., G. O. Hughes *et al.*, Migration rate limitations on climate change-induced range shifts in Cape Proteaceae, *Diversity and Distributions*, vol.12, pp.555-562, 2006.

[14] Rebelo, A. G., *Proteas. A field guide to the proteas of southern Africa*. Fernwood Press, Vlaeberg, Cape Town, South Africa, 2001.

[15] Zadeh, L. A., Fuzzy Sets, *Information and Control*, vol. 8, pp. 338-353, 1965.

[16] Zadeh, L. A., Fuzzy Sets as a Basis for a Theory of Possibility, *Fuzzy Sets and Systems*, vol.1, pp.3-28, 1978.

Appendix: Theory of Random Fuzzy Variable

First we need to review the fuzzy credibility measure theory foundation proposed by Liu [10], and then state the concept of random fuzzy variable. The theory of Liu [10, 11] is different from that initiated by Zadeh [15, 16].

Let Θ be a nonempty set, and 2^{Θ} the power set on Θ. Each element, let us say, $A \subset \Theta$, $A \in 2^{\Theta}$ is called an event. A number denoted as $\mathrm{Cr}\{A\}$, $0 \le \mathrm{Cr}\{A\} \le 1$, is assigned to event $A \in 2^{\Theta}$, which indicates the credibility grade with which event A occurs. $\mathrm{Cr}\{A\}$ satisfies the following axioms [10]:

Axiom 1: $\mathrm{Cr}\{\Theta\} = 1$.

Axiom 2: $\mathrm{Cr}\{\cdot\}$ *is non-decreasing, i.e., whenever* $A \subset B$, $\mathrm{Cr}\{A\} \le \mathrm{Cr}\{B\}$.

Axiom 3: $\mathrm{Cr}\{\cdot\}$ *is self-dual, i.e., for any* $A \in 2^{\Theta}$, $\mathrm{Cr}\{A\} + \mathrm{Cr}\{A^c\} = 1$.

Axiom 4: $\mathrm{Cr}\{\bigcup_i A_i\} \wedge 0.5 = \sup_i \left[\mathrm{Cr}\{A_i\}\right]$ *for any* $\{A_i\}$ *with* $\mathrm{Cr}\{A_i\} \le 0.5$.

Definition A.1. ([10]) *Any set function* $\mathrm{Cr}: 2^{\Theta} \to [0,1]$ *satisfies Axioms 1-4 is called a* $(\vee, \wedge)$*-credibility measure (or classical credibility measure). The triple* $(\Theta, 2^{\Theta}, \mathrm{Cr})$ *is called the* $(\vee, \wedge)$*-credibility measure space.*

Definition A.2. ([10]) *A fuzzy variable* ξ *is a mapping from credibility space* $(\Theta, 2^{\Theta}, \mathrm{Cr})$ *to the set of real numbers, i.e.,* $\xi : (\Theta, 2^{\Theta}, \mathrm{Cr}) \to \mathbb{R}$.

Definition A.3. ([10]) *The credibility distribution* $\Phi : \mathbb{R} \to [0,1]$ *of a fuzzy variable* ξ *on* $(\Theta, 2^{\Theta}, \mathrm{Cr})$ *is*

$$\Phi(x) = \mathrm{Cr}\left\{\theta \in \Theta \,\middle|\, \xi(\theta) \le x\right\}. \tag{A1}$$

Now we are ready to state the random fuzzy variable concept.

Definition A.4. *A random fuzzy variable, denoted as* $\xi = \left\{ X_{\beta(\theta)}, \theta \in \Theta \right\}$ *, is a collection of random variables* X_β *defined on the common probability space* $(\Omega, \mathfrak{A}, \Pr)$ *and indexed by a fuzzy variable* $\beta(\theta)$ *defined on the credibility space* $\left(\Theta, 2^\Theta, \mathrm{Cr}\right)$.

Definition A.5. ([10]) *Let* ξ *be a random fuzzy variable, then the average chance measure denoted by* $\mathrm{ch}\{\cdot\}$*, of a random fuzzy event* $\{\xi \le x\}$*, is*

$$\mathrm{ch}\{\xi \le x\} = \int_0^1 \mathrm{Cr}\left\{\theta \in \Theta | \Pr\{\xi(\theta) \le x\} \ge \alpha\right\} \mathrm{d}\alpha \cdot \tag{A2}$$

Then function $\Psi(\cdot)$ *is called as average chance distribution if and only if:*

$$\Psi(x) = \mathrm{Ch}\{\xi \le x\}. \tag{A3}$$

Now, let us to derive the average chance distribution for a normal random fuzzy variable $\xi \overset{d}{\sim} N(\zeta, \sigma^2)$, where the mean ζ is a triangular fuzzy variable and standard deviation σ is a given positive real number. Note that fuzzy event

$$\begin{aligned} &\left\{\theta \in \Theta : \Pr\{\xi(\theta) \le x\} \ge \alpha\right\} \Leftrightarrow \left\{\theta \in \Theta : \Phi\left(\frac{x - \zeta(\theta)}{\sigma}\right) \ge \alpha\right\} \\ &\Leftrightarrow \left\{\theta \in \Theta : x \ge \zeta(\theta) + \sigma\Phi^{-1}(\alpha)\right\} \Leftrightarrow \left\{\theta \in \Theta : \zeta(\theta) \le x - \sigma\Phi^{-1}(\alpha)\right\}. \end{aligned} \tag{A4}$$

The fuzzy mean is assumed to have a triangular membership function

$$\mu_\zeta(w) = \begin{cases} \dfrac{w - a_\zeta}{b_\zeta - a_\zeta}, & a_\zeta \le w < b_\zeta \\ \dfrac{c_\nu - w}{c_\nu - b_\nu}, & b_\zeta \le w < a_\zeta \\ 0, & \text{otherwise} \end{cases} \tag{A5}$$

and

$$\Lambda(w) = \mathrm{Cr}\{\zeta \le \mathrm{w}\} = \begin{cases} 0, & w < a_\nu \\ \dfrac{w - a_\zeta}{2(b_\zeta - a_\zeta)}, & a_\zeta \le w < b_\zeta \\ \dfrac{w + c_\zeta - 2b_\zeta}{2(c_\zeta - b_\zeta)}, & b_\zeta \le w < c_\zeta \\ 1, & w \ge c_\zeta \end{cases} \tag{A6}$$

which gives the credibility distribution for the fuzzy mean, ζ .

Then the critical step is to derive the expression of $\mathrm{Cr}\left\{\zeta(\theta) \in \Theta | \Pr\{\xi(\omega, \theta) \le x\} \ge \alpha\right\}$. For normal random fuzzy variable with a triangular fuzzy mean,

$$\left\{\zeta(\theta) : \Pr\{\xi(\omega, \theta) \le x\} \ge \alpha\right\} \Leftrightarrow \left\{\theta \in \Theta : \zeta(\theta) \le x - \sigma\Phi^{-1}(\alpha)\right\} \tag{A7}$$

where $\Phi(s) = \int_{-\infty}^{s} e^{-u^2/2} \mathrm{d}u$ denotes the standard normal cumulative distribution function.

Then the range for the integration of the integrand $\mathrm{Cr}\left\{\theta\in\Theta:\zeta(\theta)\le x-\sigma\Phi^{-1}(\alpha)\right\}$ with respect to α is listed in Table 2.

Table 2. Integration range with respect to α

$g(\alpha)$	Range for α	$\mathrm{Cr}\left\{\theta\in\Theta:\zeta(\theta)\le x-\sigma\Phi^{-1}(\alpha)\right\}$
$-\infty<g(\alpha)<a_\zeta$	$\Phi\left(\frac{x-a_\zeta}{\sigma}\right)<\alpha<1$	0
$a_\zeta\le g(\alpha)<b_\zeta$	$\Phi\left(\frac{x-b_\zeta}{\sigma}\right)<\alpha<\Phi\left(\frac{x-a_\zeta}{\sigma}\right)$	$\frac{x-\sigma\Phi^{-1}(\alpha)-a_\zeta}{2(b_\zeta-a_\zeta)}$
$b_\zeta\le g(\alpha)<c_\zeta$	$\Phi\left(\frac{x-c_\zeta}{\sigma}\right)<\alpha<\Phi\left(\frac{x-b_\zeta}{\sigma}\right)$	$\frac{x-\sigma\Phi^{-1}(\alpha)+c_\zeta-2b_\zeta}{2(c_\zeta-b_\zeta)}$
$g(\alpha)\ge c_\zeta$	$0<\alpha<\Phi\left(\frac{x-c_\zeta}{\sigma}\right)$	1

where $\zeta=g(\alpha)=x-\sigma\Phi^{-1}(\alpha)$.

Then we obtain the average chance measure for the event $\{\xi(\omega,\theta)\le x\}$:

$$\mathrm{ch}\{\xi(\omega,\theta)\le x\}=\int_{\Phi\left(\frac{x-b_\zeta}{\sigma}\right)}^{\Phi\left(\frac{x-a_\zeta}{\sigma}\right)}\frac{x-\sigma\Phi^{-1}(\alpha)-a_\zeta}{2(b_\zeta-a_\zeta)}\mathrm{d}\alpha+\int_{\Phi\left(\frac{x-c_\zeta}{\sigma}\right)}^{\Phi\left(\frac{x-b_\zeta}{\sigma}\right)}\frac{x-\sigma\Phi^{-1}(\alpha)+c_\zeta-2b_\zeta}{2(c_\zeta-b_\zeta)}\mathrm{d}\alpha+\int_0^{\Phi\left(\frac{x-c_\zeta}{\sigma}\right)}1\times\mathrm{d}\alpha \tag{A8}$$

which leads to the average chance distribution:

$$\begin{aligned}\Psi(x)=&\frac{x-a_\zeta}{2(b_\zeta-a_\zeta)}\left(\Phi\left(\frac{x-a_\zeta}{\sigma}\right)-\Phi\left(\frac{x-b_\zeta}{\sigma}\right)\right)\\&+\frac{x+c_\zeta-2b_\zeta}{2(c_\zeta-b_\zeta)}\left(\Phi\left(\frac{x-b_\zeta}{\sigma}\right)-\Phi\left(\frac{x-c_\zeta}{\sigma}\right)\right)\\&+\Phi\left(\frac{x-c_\zeta}{\sigma}\right)-\frac{\sigma}{2(b_\zeta-a_\zeta)}\int_{\frac{x-b_\zeta}{\sigma}}^{\frac{x-a_\zeta}{\sigma}}u\phi(u)\mathrm{d}u-\frac{\sigma}{2(c_\zeta-b_\zeta)}\int_{\frac{x-c_\zeta}{\sigma}}^{\frac{x-b_\zeta}{\sigma}}u\phi(u)\mathrm{d}u.\end{aligned} \tag{A9}$$

Take the derivative with respect to x, the average chance density is obtained

$$\begin{aligned}\psi(x)=&\frac{1}{2(b_\zeta-a_\zeta)}\left(\Phi\left(\frac{x-a_\zeta}{\sigma}\right)-\Phi\left(\frac{x-b_\zeta}{\sigma}\right)\right)+\frac{x-a_\zeta}{2(b_\zeta-a_\zeta)\sigma}\left(\phi\left(\frac{x-a_\zeta}{\sigma}\right)-\phi\left(\frac{x-b_\zeta}{\sigma}\right)\right)\\&+\frac{1}{2(c_\zeta-b_\zeta)}\left(\Phi\left(\frac{x-b_\zeta}{\sigma}\right)-\Phi\left(\frac{x-c_\zeta}{\sigma}\right)\right)+\frac{x+c_\zeta-2b_\zeta}{2(c_\zeta-b_\zeta)\sigma}\left(\phi\left(\frac{x-b_\zeta}{\sigma}\right)-\phi\left(\frac{x-c_\zeta}{\sigma}\right)\right)+\frac{1}{\sigma}\phi\left(\frac{x-c_\zeta}{\sigma}\right)\\&-\frac{1}{2(b_\zeta-a_\zeta)}\left(\left(\frac{x-a_\zeta}{\sigma}\right)\phi\left(\frac{x-a_\zeta}{\sigma}\right)-\left(\frac{x-b_\zeta}{\sigma}\right)\phi\left(\frac{x-b_\zeta}{\sigma}\right)\right)\\&-\frac{1}{2(c_\zeta-b_\zeta)}\left(\left(\frac{x-b_\zeta}{\sigma}\right)\phi\left(\frac{x-b_\zeta}{\sigma}\right)-\left(\frac{x-c_\zeta}{\sigma}\right)\phi\left(\frac{x-c_\zeta}{\sigma}\right)\right).\end{aligned} \tag{A10}$$

Finally, if the fuzzy mean is assumed to have a isosceles triangular membership function

$$\mu_\zeta(w)=\begin{cases}\dfrac{w+h}{h}, & -h\le w<0\\ \dfrac{h-w}{h}, & 0\le w<h\\ 0, & \text{otherwise}\end{cases}, \tag{A11}$$

the average chance density takes the form

$$\begin{aligned}\psi(x)=&\frac{1}{2h}\left(\Phi\left(\frac{x+h}{\sigma}\right)-\Phi\left(\frac{x-h}{\sigma}\right)\right)+\frac{1}{\sigma}\phi\left(\frac{x-h}{\sigma}\right)\\ &-\frac{1}{2h}\left(\left(\frac{x+h}{\sigma}\right)\phi\left(\frac{x+h}{\sigma}\right)-\left(\frac{x-h}{\sigma}\right)\phi\left(\frac{x-h}{\sigma}\right)\right)\\ &+\frac{x+h}{2h\sigma}\left(\phi\left(\frac{x+h}{\sigma}\right)-\phi\left(\frac{x-h}{\sigma}\right)\right).\end{aligned} \tag{A12}$$