

PDEAR Modelling of Climate Change Prediction of Rare Protea Species

Danni Guo¹, Renkuan Guo², Guy F. Midgley¹, A.G. Rebelo¹, and Christien Thiart²

¹Kirstenbosch Research Center, South African National Biodiversity Institute, Private Bag X7, Claremont 7735, Cape Town, South Africa

²Department of Statistical Sciences, University of Cape Town, Private Bag, Rondebosch 7701, Cape Town, South Africa

Corresponding author: Dr Danni Guo

guo@sanbi.org

KEYWORDS: Partial Differential Equation Associated Regression, PDEAR model, climate change, Protea

1. Introduction

Global warming and climate changes are changing the environment and therefore changing the distribution and behaviour of the plant species. In this paper, we will model the Protea species in the population size of 1 to 10, in the Cape Floristic Region, in 2002, in South Africa. Since we have presence data only and these are incomplete sample data, so we will look at occurrence counts or frequency distributions of the Protea. We will use the *partial differential equation associated regression* (PDEAR) model, with consideration of precipitation and temperature, to fill in missing samples within the Cape Floristic Region, in 2002. Using future precipitation and temperature variables (Hewitson and Crane 2006, New et al. 2002, Tadross et al. 2005), we will use the PDEAR model to predict the future occurrence counts of rare Proteas population count.

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., 2007). The Protea Atlas Project collected samples of Fynbos's flowering Proteas in the Cape Floristic Region, South Africa, in 2002. In this case, we are focusing on the population category of Proteas that has an estimated population size from 1 to 10 per sample site.

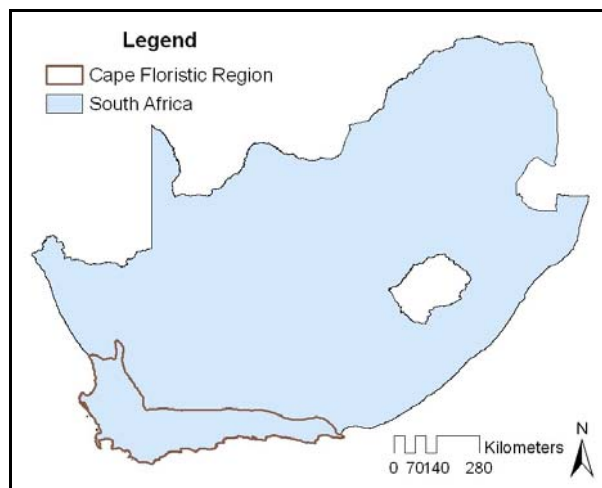


Figure 1. The Cape Floristic Region within South Africa

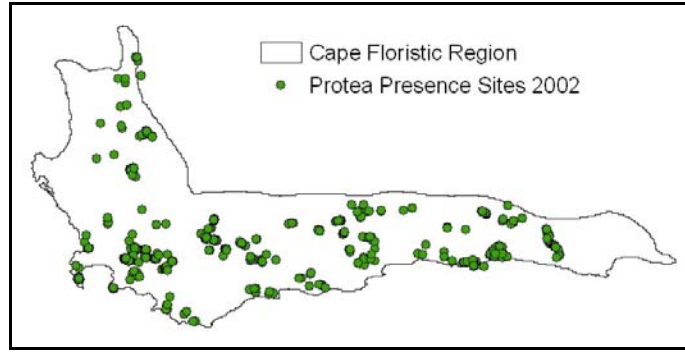


Figure 2. The Sample Locations of Proteas in the Population Size of 1-10, in the Cape Floristic Region, 2002

Figure 2 shows the locations of Proteas occurrence of the population size of 1 to 10, in the Cape Floristic Region. As one can see from the figure, the sample locations are not well spread. To solve this problem, we will use occurrence counts of the Proteas. The Cape Floristic Region is divided into 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. In Figure 3, the pink color are 0 in value, it shows the un-sampled locations. It is clear that a lot of the areas are un-sampled. The PDEAR model will be used here in order to predict the un-sampled cells.

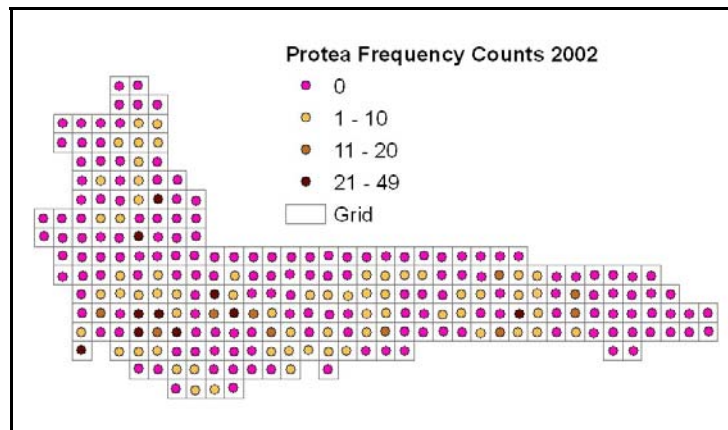


Figure 3. The Sampled Frequency Counts of Proteas in the Population Size of 1-10, in the Cape Floristic Region, 2002

3. The Theoretical Foundation for PDEAR modelling

PDEAR merges partial differential equation, multivariate regression and credibility measure theory into a new modelling family. For two-dimensional spatial modelling, the bivariate PDEAR model has following form:

$\begin{cases} PDES \\ CREG \end{cases}$	(1)
--	-----

where *PDES* denotes the first-order Pfaff linear partial differential equation system:

$\begin{cases} \frac{\partial \underline{z}}{\partial x_1} = B_1 \underline{z} + \underline{a}_1 \\ \frac{\partial \underline{z}}{\partial x_2} = B_2 \underline{z} + \underline{a}_2 \\ \underline{z}(\underline{x}^0) = \underline{z}^0 \end{cases}$	(2)
--	-----

with constant coefficients:

$B_i = \begin{bmatrix} b_{11}^{(i)} & b_{12}^{(i)} \\ b_{21}^{(i)} & b_{22}^{(i)} \end{bmatrix}, \underline{a}_i = \begin{bmatrix} a_1^{(i)} \\ a_2^{(i)} \end{bmatrix}, i = 1, 2$	(3)
--	-----

satisfying Frobenius consistency conditions $B_1 B_2 = B_2 B_1$, and *CREG* denotes coupled bivariate regression:

$\begin{cases} \underline{\Delta}_{x_{1i}}^{\partial \underline{z}} = B_1 \underline{z}(x_{1i}, x_{2i}) + \underline{a}_1 + \varepsilon_{1i} \\ \underline{\Delta}_{x_{2i}}^{\partial \underline{z}} = B_2 \underline{z}(x_{1i}, x_{2i}) + \underline{a}_2 + \varepsilon_{2i} \\ i = 1, 2, \dots, n \end{cases}$	(4)
--	-----

with partial divided difference:

$\left\{ \begin{array}{l} \underline{\Delta}_{x_{1i}}^{\partial \underline{z}} = \begin{bmatrix} \underline{\Delta}_{x_{1i}}^{\partial z_1} \\ \underline{\Delta}_{x_{1i}}^{\partial z_2} \end{bmatrix} = \begin{bmatrix} \frac{z_1(x_{1i}, x_{2i}) - z_1(x_{1(i-1)}, x_{2i})}{x_{1i} - x_{1(i-1)}} \\ \frac{z_2(x_{1i}, x_{2i}) - z_2(x_{1(i-1)}, x_{2i})}{x_{1i} - x_{1(i-1)}} \end{bmatrix} \\ \underline{\Delta}_{x_{2i}}^{\partial \underline{z}} = \begin{bmatrix} \underline{\Delta}_{x_{2i}}^{\partial z_1} \\ \underline{\Delta}_{x_{2i}}^{\partial z_2} \end{bmatrix} = \begin{bmatrix} \frac{z_1(x_{1i}, x_{2i}) - z_1(x_{1i}, x_{2(i-1)})}{x_{2i} - x_{2(i-1)}} \\ \frac{z_2(x_{1i}, x_{2i}) - z_2(x_{1i}, x_{2(i-1)})}{x_{2i} - x_{2(i-1)}} \end{bmatrix} \end{array} \right.$	(5)
---	-----

In Equation (1), the bivariate system is called the associated bivariate partial equation system, the bivariate regression system is called the coupled regression, and the error term is in nature of random fuzzy, in which the fuzziness comes from the replacements of the partial derivative vectors, say, $\partial \underline{z} / \partial x_1$, $\partial \underline{z} / \partial x_2$ in associated partial equation system by the divided partial differences, say, $\underline{\Delta}_{x_{1i}}^{\partial \underline{z}}$, $\underline{\Delta}_{x_{2i}}^{\partial \underline{z}}$ respectively, while the random error comes from sampling from the system under investigation. For more theoretical details, see Guo et al (2007^s, 2008).

For an intuitive explanation, we use *dear* (an abbreviation of *differential equation associated regression*) model, the univariate counterpart of PDEAR model to avoid the complexity in multi-dimensional spaces.

Without loss of generality, a simple linear differential equation, $dx/dt = a + bx$ is used for illustrations. Let $\hat{x}_i^{(0)}$ denote an approximation to the primitive function $x(t)$ at t_i , and let Dx_i/Dt_i be an approximation to the derivative function dx/dt at t_i , with $Dx_i = x(t_i) - x(t_{i-1})$, $Dt_i = t_i - t_{i-1}$.

Definition 2.1: If a dynamic system governed by $dx/dt = a + bx$ is sampled at its derivative level, denoted by $X^{(0)} = \{x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}\}$, the coupled equation system:

$$\begin{cases} \frac{dx}{dt} = a + bx \\ x_i^{(0)} = a + b\hat{x}_i^{(0)} + e_i, i = 1, 2, \dots, n \end{cases} \quad (6)$$

is called Type I DEAR model. If it is sampled at its primitive level, denoted by $X^{(0)} = \{x(t_1), x(t_2), \dots, x(t_n)\}$, the coupled equation system:

$$\begin{cases} \frac{dx}{dt} = a + bx \\ \frac{Dx_i}{Dt_i} = a + bx(t_i) + e_i, i = 1, 2, \dots, n \end{cases} \quad (7)$$

is called Type II DEAR model.

The second equations in equation system (6) and (7) are called coupled regressions, while the first one is called the associated differential equation.

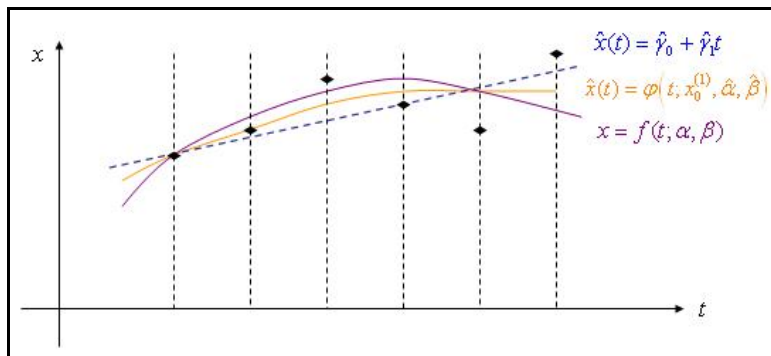


Figure 4. Two approximations to nonlinear curve $x(t) = f(t; a, b)$ in (t, x) space

In terms of Type I DEAR, we first fit the coupled regression in the (x, x') space (Figure 5), where x' denotes the derivative of x with respect to t .

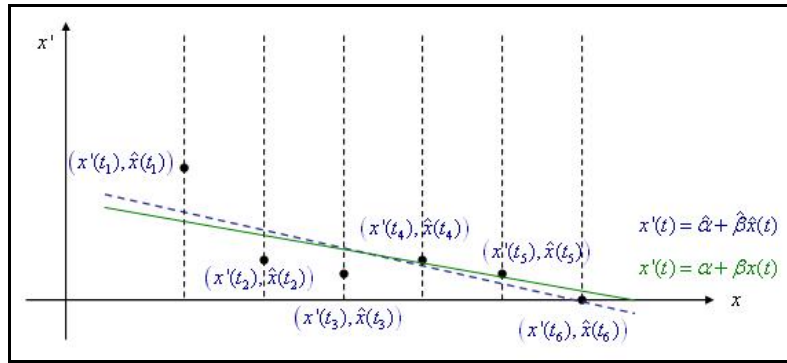


Figure 5. Type I approximation in (x, x') space

Then estimator of parameter $\underline{\theta} = (\alpha, \beta)$, denoted by $\hat{q} = (\hat{\alpha}, \hat{\beta})$ is obtained. Once the parameter $\underline{\theta} = (\alpha, \beta)$ is obtained, by solving the approximated linear differential equation $dx/dt = \hat{a} + \hat{b}x$, we will obtain an approximated nonlinear curve $x' = j(t; x_0^0, \hat{a}, \hat{b})$, (yellow-coloured curve in Figure 4), which is expected to approximate the primitive curve in relatively high accuracy.

If the sampling observations are collected at primitive function level, denoted as $X^{(0)} = \{x(t_1), x(t_2), \dots, x(t_n)\}$. Then in terms of Type II DEAR Type II, the derivatives could be approximated, for example, by the divided difference, i.e., Dx_i/Dt_i . Just as shown in Figure 6, fitting $\hat{x}' = Dx/Dt = \hat{a} + \hat{b}x$ for approximating line $x' = a + bx$. Then, the estimated parameter $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ will lead the nonlinear approximation $x' = j(t; x_0^0, \hat{a}, \hat{b})$ to the primitive function $x(t) = f(t; a, b)$ in (t, x) space (shown in Figure 4).

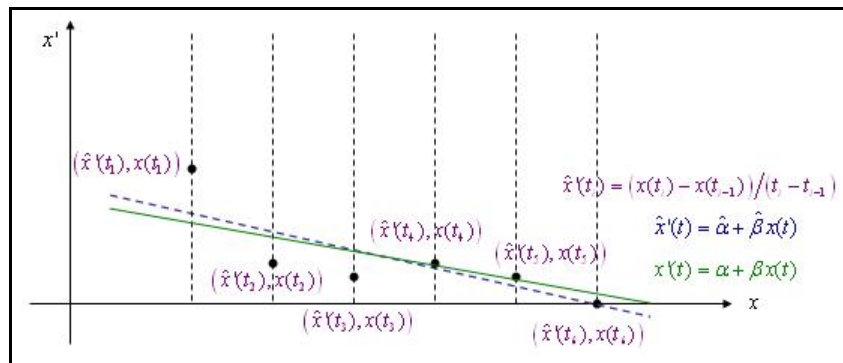


Figure 6. Type II approximation in (x, x') space

The error structure makes the coupled regression being no longer a regression in statistical linear model sense, rather in the random fuzzy variable theoretical sense, which engages Liu's credibility measure theory (2004, 2007) rather Zadeh's fuzzy mathematics (1965, 1978). More technical details can be found in 2007 publications of Guo et al listed in the reference.

4. PDEAR Predicted Protea Occurrence Counts

Using the predicted results from the PDEAR model, the un-sampled cells are predicted with frequency counts of the Protea. Figure 7 shows the PDEAR model predicted frequency counts of Proteas. The

figures show a general range of predicted values, but in fact the actual predicted values are numerical, and have predicted the sample values of where there were no samples at all.

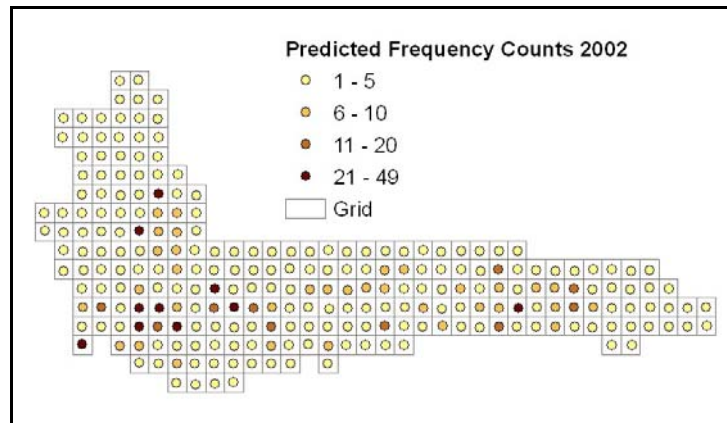


Figure 7. The PDEAR Model Predicted Frequency Counts of Proteas in the Population Size of 1-10, in the Cape Floristic Region, 2002

Now we can produce kriging prediction maps of the Proteas, using the predicted results from the PDEAR model. Figure 8 shows the distribution of occurrence counts of Proteas.

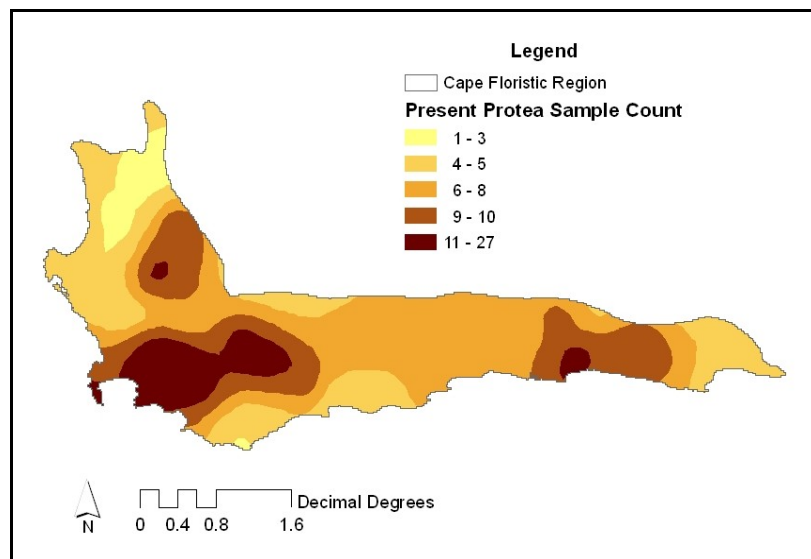


Figure 8. The Kriging Prediction Maps of Frequency Counts of Proteas in the Population Size of 1-10, in the Cape Floristic Region, 2002

Using the present precipitation and temperature data (1961-1990) and the projected future precipitation and temperature data (2070-2100) (Hewitson and Crane 2006, New et al. 2002, Tadross et al. 2005), rare Proteas counts are predicted for the future using the PDEAR model. See Figure 9.

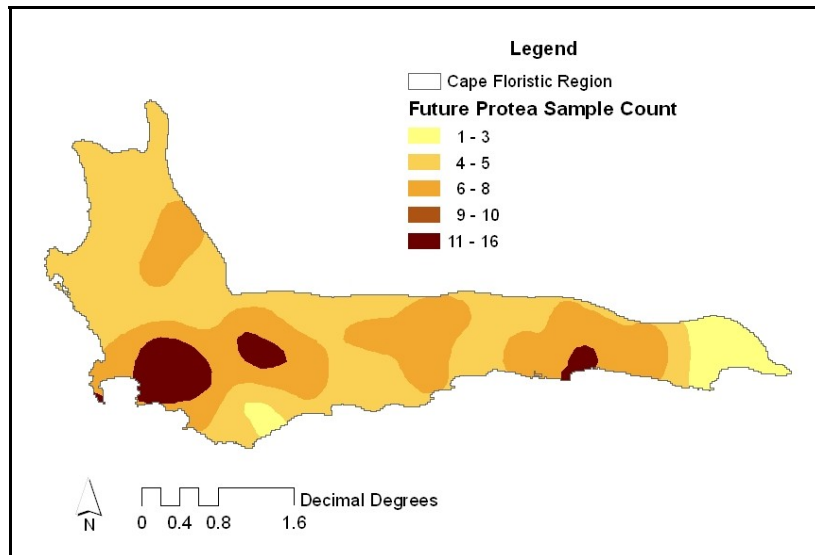


Figure 9. The Kriging Prediction Maps of Frequency Counts of Proteas in the Population Size of 1-10, in the Cape Floristic Region, in the future (2070-2100)

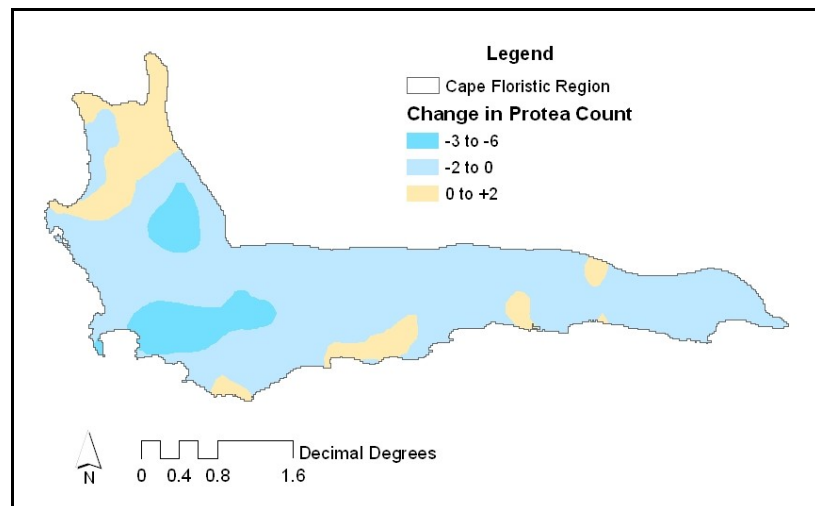


Figure 10. The Kriging Prediction Maps of Changes in Frequency Counts of Proteas between the Present and the future

Figure 10 shows the changes in Proteas counts between present and future. It is clear that in general, there is a decline in Protea counts, but in some isolated areas, a positive increase in Protea counts is shown. One must recognise that this population category of Proteas only has the population size of 1 to 10, so that it is very sensitive to environmental changes. Therefore, the Proteas between the years have changed dramatically in frequency count and spatial distribution patterns. Of particular interest is the dark blue areas, where there is a high decrease in Protea counts. It is important to recognise that under the changing climate conditions, Proteas are under threat and their distribution and numbers are changing in response to the climate, and we need to acknowledge the impacts of climate change.

5. Conclusion

The coupled regression component in a PDEAR model is in nature a special random fuzzy multivariate regression model. We developed a bivariate model for investigating the impacts from rainfall and temperature on the Protea species in the population size of 1 to 10, in the Cape Floristic

Region, in 2002, South Africa. Under same the biodiversity structure, we explore the future spatial change patterns of Protea species in the population size of 1 to 10 with future predicted rainfall and temperature. The spatial distribution and patterns are clearly will help us to explore global climate changing impacts on endangered species.

6. Acknowledgements

We would like to thank the Workshop on high-resolution climate change data for southern Africa held at the University of Witwatersrand, for providing the climate data used in this paper.

We would also 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

- Freeth R, Bomhard B and Midgley GF (2007). *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.
- Guo D, Guo R and Thiar C (2006). The Coupling of regression Modelling and Differential Equation Model in GM(1,1) Modelling and Extended GM(1,1) Models, *Journal of Grey System*, **9(2)** pp143-154.
- Guo D, Guo R, Midgley GF and Ebrahim I (2007). DEAR Modelling of Brunsvigia Litoralis Plant Distribution in Western Region District Council, South Africa , *Proceedings of the 6th International Conference on Information and Management Sciences*, Lhsa, Tibet, pp770-782.
- Guo, D., Guo, R., Midgley, G. F. and Rebelo, A. G. PDEMR Modelling of Protea Species in the Population Size of 1 to 10, in Cape Floristic Region from 1992 to 2002, South Africa, *Journal of Geographical Information Sciences*, ISSN 1082-4006, Vol. 12, No. 2, pp 67-78.
- Guo R, Guo D, Midgley GF and Rebelo AG (2007) Bivariate DEAR Modelling of Range Shifts in Aulax Umbellata Shrubs From 1992 to 2002 in South Africa, *Proceedings of the 6th International Conference on Information and Management Sciences*, Lhasa, Tibet, pp783-791.
- Guo, R. and Guo, D. (2007). Univariate DEMR Model – The Formation and Foundation. *Proceedings of IEEE International Conference on Grey Systems and Intelligent Services*, November 18-20, Nanjing, China, 2007, pp 13-18. IEEE Catalog Number: 07EX1882C, ISBN: 1-4244-1294-3, Library of Congress: 2007926987, Editor-in-Chief: Sifeng Liu
- Guo, R. and Guo, D. (2007). Multivariate DEMR Model – The Formation and Foundation. *Proceedings of IEEE International Conference on Grey Systems and Intelligent Services*, November 18-20, Nanjing, China, 2007, pp 19-25. IEEE Catalog Number: 07EX1882C, ISBN: 1-4244-1294-3, Library of Congress: 2007926987, Editor-in-Chief: Sifeng Liu
- Guo, R., Guo, D., Midgley, G. F. and Rebelo, A. G. (2007) PDEMR Model in Rare Protea Count Prediction. *Journal of Grey System*, Vol. 10, No. 2, September, pp 105-114, 2007. (ISSN: 1028-9488)
- Guo, R., Guo, D., Midgley, G. F. and Rebelo, A. G. (2008) PDEMR Modelling of Protea Rare Species Spatial Patterns. Accepted by *Journal of Uncertain Systems*.
- Hewitson BC and Crane RG (2006) Consensus between GCM Climate Change Projections with Empirical Downscaling: Precipitation Downscaling over South Africa, *International Journal of Climatology*, **26(10)** pp1315-1337.
- Liu BD (2004) *Uncertainty Theory: An Introduction to Its Axiomatic Foundations*. Springer-Verlag Heidelberg, Berlin.
- Liu BD (2007) *Uncertainty Theory*, 2nd Edition. Springer-Verlag Heidelberg, Berlin.
- Midgley GF, Hughes GO, Thuiller W and Rebelo AG (2006) Migration rate limitations on climate change-induced range shifts in Cape Proteaceae, *Diversity and Distributions*, **12** pp555-562.
- New M, Lister D, Hulme M and Makin I (2002). A high-resolution data set of surface climate over global land areas, *Climate Research*, **21** pp1-25.
- Rebelo AG (2001). *Proteas. A field guide to the proteas of southern Africa*. Fernwood Press, Vlaeberg, Cape Town, South Africa.

Tadross M, Jack C and Hewitson B (2005). On RCM-based projections of change in southern African summer climate, *Geophysical Research Letters*, **32(23)**.
Zadeh LA (1965). Fuzzy Sets, *Information and Control*, **8** pp338-353.
Zadeh LA (1978). Fuzzy Sets as a Basis for a Theory of Possibility, *Fuzzy Sets and Systems*, **1** pp3-28.

Biography

Dr Danni Guo is a specialist scientist in the Climate Change Research Group at the South African National Biodiversity Institute (SANBI). Her current research focus is on spatial modelling and GIS in biodiversity under sparse data uncertainty.

Dr Renkuan Guo is a Professor at the University of Cape Town, South Africa. His research interest includes: spatial modelling, uncertainty modeling in quality and reliability, particularly, repairable systems, small sample inferences, stochastic process modeling, particularly, financial markets analysis, Mathematical Foundation of uncertainty theory.

Dr Guy F. Midgley has been engaged in research on climate change impacts on ecosystems and biodiversity and development of adaptation strategies since the early 1990's, and now lead SANBI's Climate Change Research Group, and contribute to the work of the National Climate Change Committee (NCCC), the South African Scientific Committee on Global Change (SASCGC), and the Intergovernmental Panel on Climate Change (IPCC).

Dr Anthony G. Rebelo is based at the Threatened Species Programme of the South African National Biodiversity Institute. His recent work includes writing up the Fynbos component of the recently completed vegetation of South Africa.

Dr Christien Thiart is a Professor at the University of Cape Town, South Africa. Her research interests includes: GIS and Graphical computing: with special interest in analysis (spatial and statistical), the Linear model with special interest in Robust estimation techniques.