

Predicting potential spatial distribution of Toothed *Leionema (Leionema Bilobum sub sp. Serrulatum)* using Weights-of-Evidence modelling with GIS

Xuan Zhu

*School of Earth, Atmosphere and Environment, Monash University
Clayton, VIC3800, Australia
Email: xuan.zhu@monash.edu*

Abstract: This paper presents a study on predicting the potential spatial distribution of Toothed *Leionema (Leionema Bilobum sub sp. Serrulatum)* using Weights-of-Evidence (WofE) predictive modelling with GIS. Toothed *Leionema* is one of four subspecies of *Leionema bilobum* from the Rutaceae family. It is a poorly investigated species which is considered rare in Victoria, Australia. WofE is a statistical method based on Bayesian theorem, which applies a log-linear form of Bayes' rules and uses the prior probability distribution, calculated based on the observed presence of the modelled species, to produce a posterior probability distribution of the species. This study aims to use WofE as a tool coupled with GIS to identify key environmental factors associated with the habitat of Toothed *Leionema*, based on field observations and to determine the potential spatial distribution of this species in the Strzelecki Ranges, part of the Great Dividing Range, located in the West Gippsland region of Victoria.

The integration of WofE with GIS allowed the distribution of the species to be predicted based on geographically referenced observation data. It involved combining a set of evidence themes representing environmental predictor variables with the known distribution of the species to generate a response theme representing the probabilities of occurrence of the species at every location. The evidence and response themes were represented as raster map layers. This research used Arc-WofE to implement WofE.

The modelling results indicate that elevation, aspect, distance to water and distance to plantation areas are significant factors associated with the spatial distribution of Toothed *Leionema*. The model shows that areas with elevations between 336m and 555m and a dominant south-western aspect that are close to plantation areas (within 710 m), and to water (between 1000-1250m), provide potentially suitable habitat for Toothed *Leionema* in the region.

The effectiveness and validity of the tool for predicting the likely distribution of the species for which there is little knowledge about their physiology and habitat requirements was also assessed. Both model validation with training and control sites and ground truthing suggest that the modelling results are valid and the predicted spatial distribution provides a useful indication of potentially suitable habitat for Toothed *Leionema*. The results not only offer some understanding of the habitat requirements of the species, but also provide a new set of data about the species for developing strategies for their conservation and management, particularly when forest plans are formulated or reviewed.

Keywords: *Species distribution modelling, Weights-of-Evidence, GIS, Toothed Leionema*

1. INTRODUCTION

Toothed Leionema (*Leionema Bilobum* sub sp. *Serrulatum*) is a “rare” plant species, which has been found in few populations in West Gippsland, Victoria, Australia. But its accurate distribution information is unavailable. An understanding and awareness of its distribution is important for reducing the risk of its conservation status deteriorating. This paper reports a study in which its potential spatial distribution in the Strzelecki Ranges in West Gippsland was predicted in order to support the development of forest management plans and implementation of a conservation strategy for the area.

Prediction of potential spatial distributions of species has been widely recognised as a significant component of conservation planning (Guisan and Zimmermann 2000; Scott et al. 2002). Most techniques for the predictive modelling of species distribution are statistical in nature and enable a probability of occurrence to be predicted in a location where no species information is available. The most commonly used statistical techniques include regression, classification techniques, environmental envelopes, Bayesian methods and neural networks. Among them, Bayesian methods have been extensively used to model the dynamics of single species and to forecast population structure changes in fragmented landscapes. Weights-of-Evidence (WofE) is a Bayesian method, which applies a log-linear form of Bayes’ rules and uses the prior probability distribution, calculated based on the observed presence of a species, to produce a posterior probability distribution of the species.

WofE was originally developed for medical diagnosis (Bonham-Carter et al. 1988). It was applied and tested for mapping habitat suitability for three-toed woodpecker with valid outcomes (Romero-Calcerrada and Luque 2006). However, very few examples exist of WofE being used to predict the spatial distribution of plant species based on environmental factors. The purpose of this research is to apply WofE modelling with GIS to identify key environmental factors associated with the Toothed Leionema habitat based on the field observed data and to use these environmental factors to determine its potential spatial distribution in the Strzelecki Ranges.

2. STUDY AREA

The Strzelecki Ranges is generally over 200m above sea level, with the highest point being Mount Tassie at 730m above sea level. It comprises of an uplift and consequent dissection of lower Cretaceous rocks. The average temperature is between 21 to 25 degrees Celsius. On average the region receives between 800 to 1400mm of rainfall annually, with the wettest periods between April and July (Master 1980). The vegetation in the area includes cool temperate rainforest, damp forest, herb rich foothills, lowland forest, warm temperate rainforest and wet forest (Boyle and Lowe 2004). The area is now predominantly owned and utilised by HVP Plantations Pty Ltd with 18,000 ha of softwood and 19,000 ha of hardwood plantation. Commercial plantations are comprised of forests dominated by Mountain Ash (*Eucalyptus regnans*), Messmate (*E. obliqua*), Bluegum (*E. globulus*), Shining Gum (*E. nitens*) and Radiata Pine (*Pinus radiata*), within which Toothed Leionema is found. Other understory species include numerous fern species, most commonly rough tree fern (*Cyathea australis*), occupying the higher and unsheltered regions with soft tree fern (*Dicksonia antarctica*) within the valleys, gullies and in other sheltered regions. The area has also been subject to the introduction of numerous invasive weed species since European settlement. Our study area was limited to a small portion of the Strzelecki ranges (Figure 1), covering an area of 25km × 30km where Toothed Leionema had been observed and accurate records of its location were available.

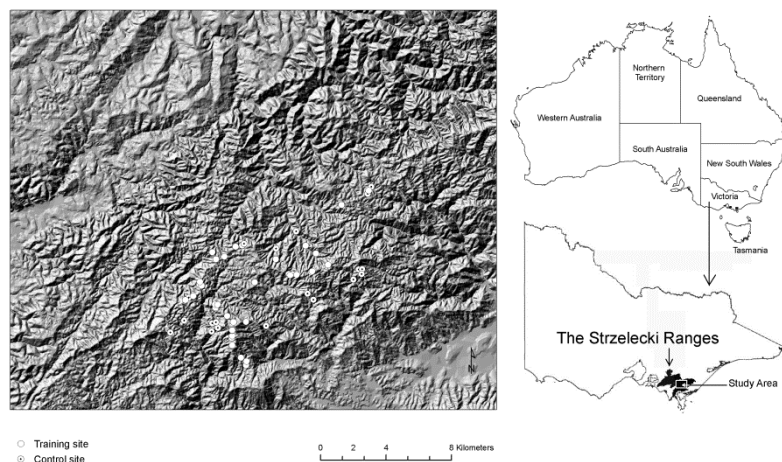


Figure 1. Study area

3. TOOTHED LEIONEMA

Toothed Leionema is one of four subspecies of *Leionema bilobum* from the Rutaceae family (Duretto et al. 2006). There is no detailed recorded knowledge of the physiology and distribution of this species. Toothed Leionema specimens were observed to resemble either a dense shrub or a small tree. The leaves are in the shape of oblong or narrow elliptic-oblong, about 16-52mm in length and 4-8mm in width. The average height of individual plants observed within the Strzelecki ranges was about 2.5m and a maximum height of approximately 3-3.5m. According to Duretto et al. (2006), Toothed Leionema actively reproduces via suckering. Essentially suckers are juvenile plants which sprout horizontally from the roots of a parent plant. There was some anecdotal evidence that suckering occurs after the roots of the parent plant have been damaged. While suckering was the primary source of recruitment observed in the field, there were many instances in the study area where seedlings were observed.

4. METHOD

4.1. Data collection

Presence data for Toothed Leionema in the Strzelecki Ranges were obtained from HVP and two field surveys. HVP provided 33 records of locations where Toothed Leionema was known to occur. These sites were verified during fieldwork and three additional sites were identified during the first round of fieldwork. The boundaries of Toothed Leionema thickets were mapped using Trimble GeoXH GPS with an accuracy of 10cm. These 36 sites were used as training sites in predictive modelling. The second field survey identified 14 new sites for the occurrence of Toothed Leionema in the modelled area. In addition, HVP provided a further 9 records. These 23 sites were used as control sites for model validation.

Due to the lack of adequate understanding of the general botany and the physiology of Toothed Leionema, indirect environmental variables, particularly topographic variables, were used in modelling. In this study, a 10m DEM provided by HVP was used to derive topographic indices. The Ecological Vegetation Classes (EVCs) data layer was acquired from the Victorian Department of Environment and Primary Industries (DEPI). Each EVC represents one or more plant communities that occur in similar types of environments and tend to show similar ecological responses to environmental factors (DSE 2004). EVCs can be used as a guide to the distribution of individual species and groups of species. The hydrology data layer from DEPI was used to analyse the proximity to streams and other water bodies. A forest data layer supplied by HVP provided information on native forests and plantation forests. This information was used as a proxy for the disturbance associated with plantation forestry.

4.2. Data analysis

WofE combines priori probabilities of the modelled species with their probabilities of occurrence conditional to the value of each predictor variable to calculate their posterior probabilities with known values of the predictor variables (Bonham-Carter et al. 1989). The integration of WofE with GIS allows the distribution of species to be predicted based on geographically referenced observation data. It involves combining a set of evidence themes representing environmental predictor variables (such as elevation and slope) with the known distribution of the modelled species to generate a response theme representing the probabilities of occurrence of the species at every location. The evidence and response themes are represented as raster map layers. This research used Arc-WofE to implement WofE (Raines et al. 2000).

In predicting the spatial distribution of a particular species, each evidence theme is classified according to its characteristics. A pair of weights is calculated for each class: W^+ measuring the influence of the presence of the class on the occurrence of the species, and W^- measuring the influence of the absence of the class on the occurrence of the species. Suppose there is a set of evidence themes and the number of classes in the i th evidence theme is m_i . The weights for class j of the i th evidence theme, W_{ij}^+ and W_{ij}^- , are calculated as log ratios of conditional probabilities (Bonham-Carter et al. 1988):

$$W_{ij}^+ = \ln \frac{P(E_{ij}|S)}{P(E_{ij}|\bar{S})} \quad (1)$$

$$W_{ij}^- = \ln \frac{P(\bar{E}_{ij}|S)}{P(\bar{E}_{ij}|\bar{S})} \quad (2)$$

where E_j is the presence of class j of evidence theme i ; \bar{E}_j is the absence of class j of evidence theme i ; S is the presence of the species; \bar{S} is the absence of the species; $P(E_j|S)$ is the conditional probability of being in class j of evidence theme i , given the presence of the species; $P(E_j|\bar{S})$ is the conditional probability of being in class j of evidence theme i , given the absence of the species; $P(\bar{E}_j|S)$ is the conditional probability of being not in class j of evidence theme i , given the presence of the species; and $P(\bar{E}_j|\bar{S})$ is the conditional probability of being not in class j of evidence theme i , given the absence of the species.

If more occurrences are found within class j of evidence theme i than would be expected by chance, W_{ij}^+ will be positive and W_{ij}^- will be negative. Conversely, W_{ij}^+ is negative and W_{ij}^- is positive when fewer occurrences are found within class j of evidence theme i than would be expected by chance. The difference between W_{ij}^+ and W_{ij}^- is called contrast, C_{ij} . It measures the strength of spatial association between the j th class and the known occurrences of the species. Larger the C_{ij} value, stronger the spatial association is. In general, the maximum value of C_{ij} ($j = 1, 2, \dots, m_i$) for the i th evidence theme gives the cut-off at which the predictive accuracy of the evidence theme is maximised. Therefore, the maximum value of C_{ij} is used to obtain the optimum cut-off (a class value) for reclassifying a continuous evidence theme into a binary map with a presence/absence pattern. The statistical significance of C_{ij} can be tested by its studentised value CS_{ij} (Bonham-Carter et al. 1988):

$$CS_{ij} = \frac{C_{ij}}{\sigma(C_{ij})} \quad (3)$$

where $\sigma(C_{ij})$ is the standard deviation of C_{ij} . When CS_{ij} is equal to or greater than 1.96, it is statistically significant at a significance level of 0.05. We call a class 'predictive' if its studentised contrast is equal to or greater than 1.96. If an evidence theme has no class that is predictive, it is considered 'insignificant' as it is not significantly associated with the spatial distribution of the species. Such an evidence theme has no prediction power, thus will be discarded from predictive modelling.

After weights and contrasts are calculated for all evidence themes, the significant evidence themes which are strongly associated with the spatial distribution of the species are identified according to studentised contrast values. These significant evidence themes are then combined to create a response theme or a probability map representing posterior probabilities of the occurrence of the species.

WofE assumes that the evidence themes are conditionally independent from one another with respect to the observed occurrences of the species. Lack of conditional independence between two evidence themes may result in an inflated posterior probability where predictive classes from the two themes are present (Agterberg and Cheng 2002). Theoretically, conditional independence of evidence themes implies that the sum of the posterior probabilities for all grid cells in the study area, T , is equal to the number of cells containing the species in the study area, NS (Agterberg and Cheng 2002). If $T > NS$, there is a lack of conditional independence among the evidence themes. According to Agterberg and Cheng (2002), the hypothesis of overall conditional independence could be accepted with a probability of 95% if $T - NS < 1.645 \sigma(T)$ where $\sigma(T)$ is the standard deviation of T .

5. RESULTS

5.1. Identification of key environmental variables associated with the presence of Toothed Leionema

The environmental variables used in predicting the spatial distribution of Toothed Leionema included elevation, slope, aspect, distance to water, EVC, and proximity to plantation areas. Weights, contrasts and studentised contrasts were calculated for each class in every evidence theme. The results showed that the studentised contrasts for all classes of the slope and EVC themes are less than 1.96, indicating that the contrasts for all the classes are not significantly greater than 0 at the significance level of 0.05. The two evidence themes have no prediction power, therefore were withdrawn from further analysis.

Toothed Leionema was found to be associated with south-westerly aspects in this region. As shown in Table 1, the south-west aspect has the highest contrast, and its studentised value is greater than 1.96. It has a high W^+ of 0.758, a low W^- of -0.163 and a low standard deviation of the contrast 0.385. Other classes in the aspect theme have a studentised contrast of less than 1.96. Therefore, the spatial association of the south-west

aspect with presence of the species is most significant statistically. Negative W^+ and positive W^- for the north-east, east and south-east aspects indicate that Toothed Leionema is less likely to be found in sites with these aspects than would be expected by chance. The north, north-west and west aspects have positive W^+ and negative W^- . They have some positive spatial association with the presence of the species, but this is not statistically significant.

Table 1. Weights, contrasts, standard deviations and studentised contrasts for aspect

| Class | W^+ | $\sigma(W^+)$ | W^- | $\sigma(W^-)$ | C | $\sigma(C)$ | CS |
|------------|---------|---------------|---------|---------------|---------|-------------|---------|
| Flat | 0.2762 | 0.5000 | -0.0297 | 0.1768 | 0.3059 | 0.5303 | 0.5768 |
| North | 0.4943 | 0.3780 | -0.0900 | 0.1857 | 0.5843 | 0.4211 | 1.3875 |
| North East | -1.4336 | 1.0000 | 0.0957 | 0.1690 | -1.5293 | 1.0142 | -1.5079 |
| East | -1.3340 | 1.0000 | 0.0833 | 0.1690 | -1.4173 | 1.0142 | -1.3975 |
| South East | -1.3885 | 1.0000 | 0.0899 | 0.1690 | -1.4783 | 1.0142 | -1.4577 |
| South | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| South West | 0.7580 | 0.3333 | -0.1631 | 0.1925 | 0.9211 | 0.3849 | 2.3932 |
| West | 0.5481 | 0.3780 | -0.0970 | 0.1857 | 0.6451 | 0.4211 | 1.5320 |
| North West | 0.3233 | 0.4083 | -0.0538 | 0.1826 | 0.3771 | 0.4472 | 0.8432 |

The elevation in the study area ranges from 50 to 690m. All the high contrasts with statistical significance are in the range between 350–550m indicating that this range of elevation is favourable for Toothed Leionema. The elevation group between 450 and 550m is particularly favourable, since W^+ is larger than 2 and the absolute value of W^- is less than 0.3. W^+ and contrast are negative or zero for the elevation of lower than 350m or higher than 550m, indicating that the distribution of Toothed Leionema is less likely associated with these ranges of elevation.

Toothed Leionema is not particularly dependent on close proximity to water. The highest contrast that is statistically significant is for a distance from water between 1,100 to 1,200m. When the distance to water is less than 1,100m, W^+ changes from negative values to positive ones at about 200m, and then becomes negative again in the range of 1,000 to 1,100m. It implies that this species is not significantly associated with access to water when it is located within 1,100m from water bodies. But it has a negative spatial association with the access to and availability of water when its distance to water is beyond 1,400m, as the distance of >1,400m has a statistically significant large negative W^+ (-2.21) and contrast (-2.48) with a studentised contrast of -2.437 (< -1.96). It means that at locations beyond 1,400m from water, fewer occurrences of Toothed Leionema are present than would be expected by chance.

Proximity to plantation shows significant values up to 700m, with higher W^+ value with increasing proximity to plantation. Therefore the species are more likely present in a closer proximity to plantation. However, it may be caused by sampling bias as most presence sites were provided by HVP, which would presumably be preferentially located near plantations.

It should be noted that CS is to test the significance of the environmental factors, not the predictive values.

5.2. Overall conditional independence

The observed number of cells containing the species in the study area $NS = 36$, and the expected number of cells containing the species $T = 38.959$. The difference $T - NS$ is 2.959 and the standard deviation of T is 2.677. As $T - NS < 1.645 \sigma(T) = 4.404$. Therefore, the conditional independence hypothesis can be accepted with a probability of 95%.

5.3. Potential spatial distribution of Toothed Leionema

The potential spatial distribution of Toothed Leionema within the study area was determined using the posterior probabilities of species presence, which is shown in Figure 2. Based on the modelling results, about 10% of the study area is predicted as high potential habitat for Toothed Leionema. These

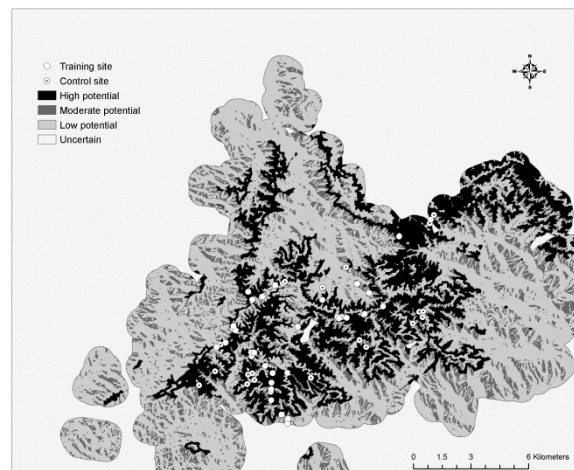


Figure 2. Predicted distribution of Toothed Leionema

areas have a mean elevation of 395m (ranging from 336-555m), a mean distance to water of 983m, a mean distance to plantation of 120m and a dominant south-western aspect. About 35% of the study area is predicted as moderate and low potential habitat for Toothed Leionema. Over 55% of the study area is considered uncertain. The uncertain areas are generally located more than 680m away from plantation areas with a north-western or northern aspect. The areas with moderate potential tend to have an aspect of south west, while those with low potential are likely to have an aspect of south or south east.

5.4. Model evaluation

The modelling results were validated against both the training sites and control sites. In this study, 26 of the 36 training sites (72%) were located inside the areas of high potential habitat predicted for Toothed Leionema; 5 fell in the areas of moderate potential, 4 in the areas of low potential and 1 in the uncertain areas. Of the 23 control sites, 15 (65%) fell within the areas of high potential, 4 in the areas of moderate potential, and 4 in the areas of low potential. None of the control sites is located in the uncertain areas. A predictive map is useful if it identifies at least 70% of the training sites and at least 50% of the control sites inside the areas of high potential. Therefore, our predictive results are valid.

In addition, ground truthing was carried out in the field after modelling. Over 40 locations across the four classes of potential areas were randomly selected and visited. 12 sites were visited in the areas predicted as high potential. Toothed Leionema was found at all sites except one. The environmental conditions observed at these sites remained fairly consistent in terms of those under which Toothed Leionema was previously found to be growing. These conditions include open understorey free from species such as ferns, open upper storey canopies predominantly consisting of species such as Mountain Ash (*Eucalyptus regnans*) and Messmate (*E. obliqua*) which allow for large amount of light penetration, and moist intact soil generally growing along a slope. A possible explanation to why Toothed Leionema was not found at one of the sites is that the site is situated directly opposite a private property. Numerous alien plant species that were introduced to the Strzelecki Ranges were identified in the immediate area, indicating that the area might have been modified in some way or Toothed Leionema had been out competed by the introduced alien species.

The areas with moderate potential are considerably smaller and far more scattered when compared to the areas with high potential. Due to their small size and scattered locations, it was difficult to deliberately visit these areas during ground truthing. As a result, many of the ground truthing sites selected for assessment contained in the moderate potential areas are intertwined with those falling in the other potential classes. GPS readings indicated that two of the new colonies discovered during ground truthing were located within the areas predicted as moderate potential.

The areas with low potential are predominantly located along steep slopes. Eight sites in these low potential areas were visited for ground truthing, and Toothed Leionema was found at only two of these.

The uncertain areas are where the presence of Toothed Leionema is not significantly correlated with the environmental factors. The presence of Toothed Leionema in these regions is largely random. No colonies of Toothed Leionema were found in areas of this category, although more than 10 sites were visited during ground truthing.

6. DISCUSSION AND CONCLUSIONS

This paper applied WofE coupled with GIS to identify key habitat elements for Toothed Leionema and to build a model to predict unknown locations. Four environmental factors, including aspect, elevation, distance to water and proximity to plantation areas, were found to be correlated with the observed occurrence of Toothed Leionema, and were shown to have significant predictive power to estimate the spatial distribution of the species. Approximately 10% of the study area was predicted to have high potential for their growth, which highlights the rarity of the species in the region. Both model validation with training and control sites and ground truthing undertaken in the field suggest that the modelling results are valid and the predicted spatial distribution provides a useful indication of potentially suitable habitat for Toothed Leionema.

WofE uses available biophysical data and existing observations to approximate the habitat requirements of the species of interest and to establish the relationships between the environmental factors and their spatial distributions. It is able to objectively choose the significant environmental variables as habitat requirements and determine their weights based on evidence rather than subjective judgements. With this ability, WofE can be used to objectively exclude insignificant environmental factors that have little prediction power from modelling.

However, like any other predictive modelling methods, WofE cannot provide a perfect representation of real and accurate distributions of the species. Ecosystems are highly complex and heterogeneous, and are very difficult to model accurately in every aspect of time and space (Guisan and Zimmermann 2000). The WofE model for the distribution of Toothed Leionema developed in this research is limited by its lack of sufficient environmental data and a limited number of known occurrences of the species. More understanding of the physiology and behaviour of the species is needed, and more data sets (such as soils, presence of ferns and upper storey species) are required to make the model more robust. It is also impossible to incorporate dynamic aspects of the species' succession into the model. In addition, the WofE model predicted a simulated distribution from field derived observations (i.e. based on empirical field data sets) rather than from theoretical physiological constraints. Therefore, the predicted potential spatial distribution is within the realised niche of Toothed Leionema, not a function of physiological performance and ecosystem constraints of the species and cannot be applied in changing environmental situations. Moreover, this study did not compare WofE with other species distribution modelling methods such as Maxent. It will be evaluated against other modeling techniques in the future.

Nevertheless, the WofE model has been useful in identifying key habitat conditions of Toothed Leionema and predicting areas where Toothed Leionema is likely to be present, in a way that has not been possible in the past. Although the predicted probabilities of occurrence resulted from WofE are not absolute and represent only a relative degree of potential distribution, they provide a measure of localities with likely occurrence of Toothed Leionema.

REFERENCES

- Agterberg, F.P. and Q. Cheng (2002). Conditional independence test for weights-of-evidence modeling. *Natural Resources Research*, 11, 249-255.
- Bonham-Carter, G.F., F.P. Agterberg and D.F. Wright (1988). Integration of geological data sets for gold exploration in Nova Scotia. *Photogrammetric Engineering and Remote Sensing*, 54, 1585-1592.
- Bonham-Carter, G.F., F.P. Agterberg and D.F. Wright (1989). Weights of evidence modelling: A new approach to mapping mineral potential. In: Agterberg, F.P. and G.F. Bonham-Carter (eds.) *Statistical Applications in the Earth Sciences*, pp.171-183. Geological Survey of Canada, Ottawa.
- Boyle, C. and K.W.Lowe (2004). *Biodiversity Action Planning – Strategic Overview for the Strzelecki Ranges Bioregion*. Department of Sustainability and Environment, Melbourne, Australia.
- Department of Sustainability and Environment (DSE) (2004). *EVC/Bioregion Benchmark for Vegetation Quality Assessment – Strzelecki Ranges Bioregion*. Department of Sustainability and Environment, Melbourne, Australia.
- Duretto, M.F., K.L. Durham, and E.A. James (2006). New subspecies of *Leionema bilobum* (Rutaceae). *Muelleria*, 23, 7-14.
- Guisan, A. and N.E. Zimmermann (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147-186.
- Master, I. (1980). *A Report on the Sunny Creek Catchment: A Proposal for Proclamation*. Land Conservation Council, Melbourne, Australia.
- Raines, G.L., G.F. Bonham-Carter and L. Kemp (2000). Predictive probabilistic modeling using ArcView GIS. *ArcUser*, April - June 2000, 45-48.
- Romero-Calcerrada, R. and S. Luqu (2006). Habitat quality assessment using Weights-of-Evidence based GIS modelling: the case of *Picooides tridactylus* as species indicator of the biodiversity value of the Finnish forest. *Ecological Modelling*, 196, 62-76.
- Scott, J.M., P.J. Heglund and M.L. Morrison (2002). *Predicting species occurrences: issues of accuracy and scale*. Island Press, Covelo, CA.