# Likelihood and consequences: reframing the Australian weed risk assessment to reflect a standard model of risk

**Curtis C. Daehler**[A] and **John G. Virtue**[B]
[A] Department of Botany, University of Hawai'i, Honolulu, HI 96822, USA. Email: daehler@hawaii.edu
[B] Department of Water Land and Biodiversity Conservation, Adelaide, South Australia 5001, Australia.

## Summary

**Whereas standard risk assessment models express risk as a function of likelihood and consequences, the Australian weed risk assessment (WRA) system expresses risk as a single score based on the answers to 49 questions, with no explicit differentiation of likelihood and consequences. We identified WRA questions that most closely reflect likelihood of spread and those that reflect consequences of spread (impacts) to determine whether explicit consideration of likelihood and consequences could reduce WRA mis-classifications or provide insights into why species may be misclassified. Data from a previously published test of the Australian WRA system in Hawai'i were reanalysed. As expected, most major weeds had high scores for both likelihood and consequences components of risk, while both scores tended to be low for non-weeds. Major and minor weeds did not differ in terms of consequence scores, but major weeds averaged higher likelihood scores. A composite score obtained by multiplying likelihood and consequences was better at identifying weeds than the original WRA score while correctly identifying non-weeds at the same rate as the original WRA score. Explicit consideration of likelihood and consequences is a promising approach to improve WRA, but questions remain about non-independence of these components.**

**Keywords: Impacts, likelihood, risk assessment, spread.**

## Introduction

A standard approach to risk assessment is to estimate the likelihood (L) that an event will occur and the consequences (C) of that event if it does occur (Bowden *et al.* 2001). Likelihood is a probability whereas consequences are expressed as a magnitude. Each can be estimated quantitatively or semi-quantitatively. A risk assessment score (R) is then obtained by multiplying the likelihood by the consequences (R = L × C) (Bowden *et al.* 2001). R is a quantity or index on which policy decisions may be enacted. For example, in weed risk assessment (WRA), plants that have R-values above a specific threshold may be denied entry at the border because they are deemed as having high potential to become weeds.

Richardson *et al.* (2000) sought to standardize definitions in plant invasions and defined weeds as 'plants (not necessarily non-native) that grow in sites where they are not wanted and which usually have detectable, negative economic, environmental and/or social effects'. These undesirable effects are synonymous with the term 'impacts', which in turn equates to 'consequences' in the language of risk. In this paper we focus specifically on introduced (non-native) weeds, as most WRA systems are designed and implemented to prevent weed introductions. To frame WRA within a standard risk model, we define consequences (C) as the magnitude of impacts that may occur if an introduced plant successfully spreads beyond deliberate plantings or beyond the point of accidental introduction (if not deliberately planted).

Likelihood (L) of the consequences is then defined as the likelihood that the plant will be introduced and spread. For proposed or deliberate introductions, the plant is certain to be introduced, so we consider only the likelihood of spread or naturalization (Pysek *et al.* 2004) after introduction multiplied by the consequences, as an index for assessing weed risk. Spread alone does not confer weed status.

The Australian WRA (Pheloung *et al.* 1999) is used to assess proposed plant introductions to Australia, and it is increasingly being tested for use in other parts of the world including the Bonin Islands, Japan (Kato *et al.* 2006), Hawai'i and other Pacific Islands (Daehler and Carino 2000, Daehler *et al.* 2004), the Czech Republic (Křivánek and Pyšek 2006), and Florida, USA (Gordon *et al.* 2008). This system consists of 49 questions in sections covering domestication/cultivation, climate and distribution, weed history, undesirable traits, plant type, reproduction, dispersal mechanisms and persistence, each of which contributes to the WRA score. Question scores are summed, with a score >6 indicating a species has high risk of becoming a weed. In Australia, such plants are denied entry. Although separate WRA scores can be determined for risk of becoming an agricultural weed, environmental weed, or nuisance weed (Pheloung *et al.* 1999), the scoring system does not explicitly distinguish or separate likelihood and consequence components of risk.

There are at least two reasons to assess risk explicitly as a function of likelihood and consequences. First, R = L × C is a widely accepted standard for assessing risk among all sorts of regulatory and safety programs ranging from occupational health (Sadhra and Rampal 1999) to crime prevention (Fennelly 2004). Furthermore, the World Trade Organization also specifies that risk assessment for biological commodities should include likelihood and consequences components (World Trade Organization 2007). Our second interest in assessing risk in terms of L and C is that it could yield new insights into differences between weeds and non-weeds, or between major weeds and minor weeds. For example, we hypothesized that separation of L and C components would allow clear differentiation of major and minor weeds (Figure 1), and that expressing risk as L × C could help improve WRA accuracy.

## Materials and methods

### *Classifying WRA questions*

The 49 questions of the Australian WRA were classified as either relating primarily to likelihood of spread or consequences of spread (impacts) (Table 1). The likelihood questions could have yielded a score ranging from −26 to 36, whereas the questions relating to consequences could have yielded a score ranging from −1 to 21. Based on these upper and lower bounds, likelihood and consequences scores were standardized so that each could range from 0 to 10.

### *Comparison of original and modified WRA scoring*

Previous results based on the Australian WRA (Daehler *et al.* 2004) were reanalysed after separating the likelihood and consequences components of the WRA. The data consisted of WRAs for 172 species that had been present in Hawai'i for at least 50 years. Each species was independently rated by experts based on the species' actual behaviour in Hawai'i (major pest, minor pest or not a pest). WRA results were compared with expert ratings to judge accuracy of the WRA. We consider the term 'pest', as used in the original analysis (Daehler *et al.* 2004) to be synonymous with our use of 'weed' in the present paper, as both are defined by undesirable impacts.

### *Discriminant analysis*

To compare the overall ability of the original WRA scoring and our modified scoring

**Table 1. Australian WRA questions and their designation as primarily contributing to likelihood spread (L) or consequences (C).**

| | Question | Category |
|---|---|---|
| 1.01 | Is the species highly domesticated? If answer is 'no' go to question 2.01 | L |
| 1.02 | Has the species become naturalized where grown? | L |
| 1.03 | Does the species have weedy races? | L |
| 2.01 | Species suited to Australian climates? (0 – low; 1 – intermediate; 2 – high) | L |
| 2.02 | Quality of climate match data? (0 – low; 1 – intermediate; 2 – high) | |
| 2.03 | Broad climate suitability (environmental versatility)? | L |
| 2.04 | Native or naturalized in regions with extended dry periods? | L |
| 2.05 | Does the species have a history of repeated introductions outside its natural range? | L |
| 3.01 | Naturalized beyond native range? | L |
| 3.02 | Garden/amenity/disturbance weed? | C |
| 3.03 | Weed of agriculture/horticulture/forestry? | C |
| 3.04 | Environmental weed? | C |
| 3.05 | Congeneric weed? | C |
| 4.01 | Produces spines, thorns or burrs? | C |
| 4.02 | Allelopathic? | C |
| 4.03 | Parasitic? | C |
| 4.04 | Unpalatable to grazing animals? | C |
| 4.05 | Toxic to animals? | C |
| 4.06 | Host for recognized pests and pathogens? | C |
| 4.07 | Causes allergies or is otherwise toxic to humans? | C |
| 4.08 | Creates a fire hazard in natural ecosystems? | C |
| 4.09 | Is a shade tolerant plant at some stage of its life cycle? | L |
| 4.10 | Grows on infertile soils? | L |
| 4.11 | Climbing or smothering growth habit? | C |
| 4.12 | Forms dense thickets? | C |
| 5.01 | Aquatic? | L |
| 5.02 | Grass? | L |
| 5.03 | Nitrogen fixing woody plant? | L |
| 5.04 | Geophyte? | L |
| 6.01 | Evidence of substantial reproductive failure in native habitat? | L |
| 6.02 | Produces viable seed? | L |
| 6.03 | Hybridizes naturally? | L |
| 6.04 | Self-fertilization? | L |
| 6.05 | Requires specialist pollinators? | L |
| 6.06 | Reproduction by vegetative propagation? | L |
| 6.07 | Minimum generative time (years)? | L |
| 7.01 | Propagules likely to be dispersed unintentionally? | L |
| 7.02 | Propagules dispersed intentionally by people? | L |
| 7.03 | Propagules likely to disperse as a produce contaminant? | L |
| 7.04 | Propagules adapted to wind dispersal? | L |
| 7.05 | Propagules buoyant? | L |
| 7.06 | Propagules bird dispersed? | L |
| 7.07 | Propagules dispersed by other animals (externally)? | L |
| 7.08 | Propagules dispersed by other animals (internally)? | L |
| 8.01 | Prolific seed production? | L |
| 8.02 | Evidence that a persistent propagule bank is formed (>1 y)? | L |
| 8.03 | Well controlled by herbicides? | L |
| 8.04 | Tolerates or benefits from mutilation, cultivation or fire? | L |
| 8.05 | Effective natural enemies present in Australia? | L |

(L × C) to separate weeds from non-weeds, discriminant analysis was conducted (SYSTAT version 10, Systat Software, Inc., San Jose, CA). A random sample consisting of 65% of the 172 expert-rated species was used to determine the best discriminant function differentiating weeds from non-weeds considering the following variables: the original WRA score, the likelihood score (L), the consequences score (C), or the composite L × C score. The remaining 35% of the data were then used to test the discriminant function. This procedure was repeated five times to generate mean rates of correct classification and ranges of correct classification for each scoring system. For these analyses, the original expert ratings of major and minor pests were pooled into a single category (weeds).

## Results

Our predicted areas of concentration for major weeds and non-weeds (Figure 1) in relation to likelihood and consequences scores were generally confirmed by the empirical findings, although no species had very low likelihood of spread (Figure 2). Contrary to predictions, species with high likelihood of spread and low consequences had proportionately more major weeds than expected, while species with low likelihood of spread and high consequences were largely absent in the data set (Figure 2).

A closer examination of major weeds and minor weeds revealed that the consequences score was not a significant discriminator ($P > 0.15$). However, major weeds were characterized by higher average scores for likelihood of spread than minor weeds (Figure 3). The discriminant function for major and minor weeds (plotted on Figure 3) included only likelihood of spread and correctly identified 68% of major weeds and 69% of minor weeds, based on 35% of the data that were reserved for testing.

### *Original WRA scores versus L × C*

For practical purposes, the original WRA strives for binary classification (weeds should be rejected from entry and non-weeds should be admitted). Major and minor weeds were therefore grouped into a single 'weed' category and the predictive ability of L × C was tested and compared to the predictive ability of the original WRA scores. Compared to the original WRA scores, the rate of correct classification of weeds was significantly greater when scores were expressed in terms of L × C (averaging 91% versus 82% for the original WRA scores), although rates of non-weed identification were not improved (Table 2). Consideration of L and C independently in the discriminant function did not improve rates of correct classification over the composite L × C score (Table 2).

## Discussion

Our findings of improved correct classifications rates when considering WRA scores formulated as L × C suggest that this is a promising approach that merits further consideration. Nevertheless, some caveats should also be considered. In particular, when assessing risk, likelihood and consequences should ideally be independent variables. In our attempt to translate Australian WRA questions into likelihood of spread versus consequences of spread, some questions clearly related to one and not the other. For example, 'toxic to animals' relates to consequences rather than potential for spread, while 'high seed production' relates to likelihood of spread but does not inform us of consequences. However, some WRA questions may relate to both likelihood of spread and consequences. For example, nitrogen fixation in a woody plant (question 5.03) may relate to an increased likelihood of naturalization in nutrient-poor soil, but when nitrogen fixation leads to unwanted soil nutrient enrichment, it becomes an impact.

We considered the 'weed elsewhere' questions to be primarily indicators of consequences, since being a weed elsewhere is one of the best single predictors of becoming a weed (having unwanted impacts) upon introduction to a new region (Reichard and Hamilton 1997, Mack *et al.* 2002). But being a 'weed elsewhere' might also reflect a species' unusually strong ability to spread. We did not assign the weed elsewhere questions to the 'likelihood' category because we considered the 'naturalized elsewhere' question to better reflect a history of successful spread and thus likelihood of spread independent of impacts. Naturalization indicates spread, but it does not imply impacts (Richardson *et al.* 2000).

Perhaps of more fundamental concern is the possibility that our perception of whether something is a weed is determined by its ability to spread, just as much as by its impacts. In this case, literature references may have classified a plant as a 'weed elsewhere' based on observations of its spread, rather than its actual impacts. This would lead to confounding of likelihood and spread for the 'weed elsewhere' questions. Some evidence that experts might weigh perceptions of spread heavily in their classification of species as weeds can be gleaned from Figure 2, which shows that questions relating to likelihood of spread were more important in differentiation major and minor weeds than were questions relating to consequences (impacts). The best strategy to minimize cross-correlation of likelihood and consequences in the 'weed elsewhere' questions is to ensure that positive answers are based on concrete impacts, rather than on life history traits. For example, ruderals are often perceived as 'weedy species' and called 'weeds' although they may not have concrete negative impacts. Positive answers to 'weed elsewhere' questions from literature sources should be firmly based on impact.

An alternative explanation for the separation of major and minor weeds based only on likelihood of spread is that weed spread and impacts are not independent. Dispersal, establishment, reproduction and persistence attributes that enable a high rate of spread at local and distant scales may also allow a high population density. Density or abundance is a major component of impact (Parker *et al.* 1999).

Rates of correct classification for the original WRA score shown in Table 2 are different from those reported in Daehler *et al.* (2004) for several reasons. First, Table 2 combines major and minor weeds into a single category. Second, the rates of

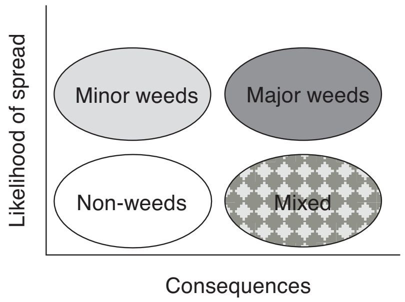


**Figure 1. Hypothesized separation of different expert designations based on likelihood and consequences scores derived from the Australian WRA. Minor weeds were expected to separate from major weeds due to lower impact (consequences). Non-weeds (most species) were expected to have both low consequences and likelihood of spread. Species with high scores for consequences and low scores for likelihood of spread were expected to be most variable, including a few major weeds (those that were successful in spreading despite low likelihood).**

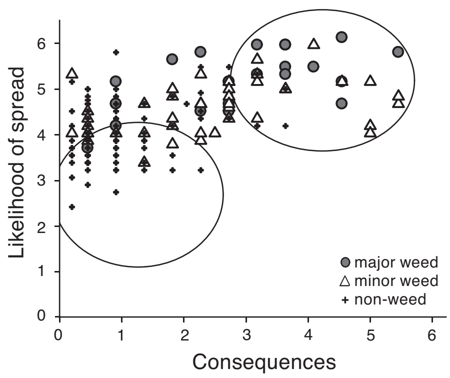


**Figure 2. Actual separation of different expert designations based on likelihood and consequences scores. Outlining circles indicate areas of general agreement with predictions (Figure 1).**

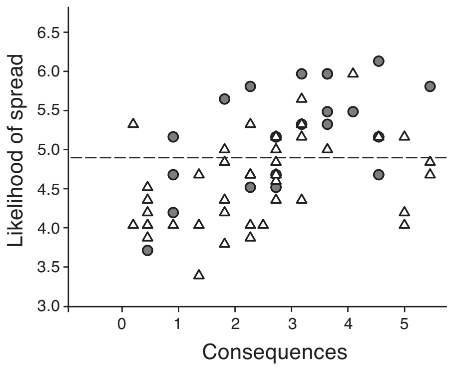


**Figure 3. Comparison of major (solid circles) and minor weeds (triangles). The dashed line is the discriminant function derived from 65% of the data. Consequences was dropped from the function due to its non-significant contribution.**

**Table 2. Mean rates of correct classification for reserve data based on discriminant function analysis, comparing original WRA score, likelihood × consequences (L × C), and likelihood and/or consequences (L, C). Bracketed values indicate range of correct classification rates observed among five analyses based on different reserve data.**

| | WRA Score | L × C | L, C |
|---|---|---|---|
| Weeds[A] | 81.8 [78–89] | 91** [87–93] | 87** [83–93] |
| Non-weeds | 78.2 [68–91] | 78.2 [69–86] | 80.2 [69–91] |
| Overall | 80.2 | 85.8 | 83.8 |

[A] Major and minor weeds combined.
** Significant difference from original WRA score, P <0.01 (chi-square test for difference in proportions).

correct classification in Table 2 are for a random subset of 35% of the species, not for the full set of 172 species. This procedure of reserving data to test the discriminant function provides a better estimate of the predictive ability of the scores. Finally, we do not allow for an 'evaluate further' category in the discriminant analysis. Instead, the discriminant function assigns a hard cut-off between weeds and non-weeds. While this procedure does not reflect the common practice of allowing for classification as 'evaluate further', the intent of the discriminant function analysis was to compare the reformulated WRA score with the original in an objective manner. An optimal range or scores for 'evaluate further' could be defined for the L × C WRA score, but this would require the defining and balancing of acceptable risks to determine the resulting size of the 'evaluate further' category. Daehler *et al.* (2004) proposed a decision tree that greatly reduced the pool of species rated as 'evaluate further' with no apparent risk of admitting major weeds, but some additional minor weeds may be admitted by this procedure.

Some other WRA systems explicitly consider likelihood and consequences. One example is the United States Department of Agriculture (USDA) system for listing of noxious weeds (Lehtonen 2001). In this system, likelihood and consequences components are summed rather than multiplied. Although summation of likelihood and consequences is not a standard practice in risk assessment (Bowden *et al.* 2001), the result could be effectively the same as multiplying in this case because the USDA system only generates coarse classifications (high, medium and low) for likelihood and consequences components.

Parker *et al.* (2007) developed a rapid WRA system for the United States that involves multiplying components of likelihood and consequences. Results were compared to those of the additive Australian model used by Daehler *et al.* (2004) and found to be comparable. But Parker *et al.* (2007) pointed out that a multiplicative model is preferable to an additive model because with an additive model (R = L + C), it is possible for a species to receive a high risk rating even if L = 0. In contrast, with a multiplicative system (R = L × C) it is unlikely for a species to receive a high risk rating if either L or C is very low.

All weed risk assessment systems to date are either qualitative or semi-quantitative, because of difficulties in measuring likelihood and consequences in absolute terms. However, Suter (2006) points out that methods are readily available for incorporating uncertainties into quantitative components of risk assessments, and quantitative risk assessment could be used more frequently. More explicit recognition of weed risk as a function of likelihood and consequences in border WRA predictive systems would be a first step in this direction; such an approach is already being adopted for post-border WRA prioritization systems (Anonymous 2006).

## Acknowledgments

This paper was improved by discussions with peers at the 2nd International Workshop on Weed Risk Assessment, as well as by comments from an anonymous reviewer.

## References

Anonymous (2006). HB 294-2006 'National post-border weed risk management protocol'. (Standards Australia International Ltd., Sydney, Standards New Zealand, Auckland and CRC Australian Weed Management, Adelaide).

Bowden, A.R., Lane, M.R. and Martin, J.H. (2001). 'Triple bottom line risk management: enhancing profit, environmental performance, and community benefits'. (J. Wiley and Sons, New York).

Daehler, C.C. and Carino, D.A. (2000). Predicting invasive plants: prospects for a general screening system based on current regional models. *Biological Invasions* 2, 92-103.

Daehler, C.C., Denslow, J.S., Ansari, S. and Kuo, H. (2004). A risk assessment system for screening out invasive pest plants from Hawai'i and other Pacific Islands. *Conservation Biology* 18, 360-8.

Fennelly, L.J. (2004). 'Handbook of loss prevention and crime prevention'. (Butterworth-Heinemann, Oxford).

Gordon, D.R., Onderdonk, D.A., Fox, A.M. and Stocker, R.K. (2008). Consistent accuracy of the Australian weed risk assessment system across varied geographies. *Diversity and Distributions* 14, 234-42.

Kato, H., Hata, K., Yamamoto, H. and Yoshioka, T. (2006). Effectiveness of the weed risk assessment system for the Bonin Islands *In* 'Assessment and control of biological invasion risk', eds F. Koike, M.N. Clout, M. Kawamichi, M.D. Poorter and K. Iwatsuki, pp. 65-72. (Shoukadoh Book Sellers, Kyoto, Japan).

Křivánek, M. and Pyšek, P. (2006). Predicting invasions by woody species in a temperate zone: a test of three risk assessment schemes in the Czech Republic (Central Europe). *Diversity and Distributions* 12, 319-27.

Lehtonen, P.P. (2001). Pest risk assessment in the United States: guidelines for qualitative assessments of weeds. *In* 'Weed risk assessment', eds R.H. Groves, F.D. Panetta, and J.G. Virtue, pp. 117-23. (CSIRO Publishing, Collingwood, Victoria).

Mack, R.N., Barrett, S.C.H., Defur, P.L., Macdonald, W.I., Madden, L.V., Marshall, D.S., Mccullough, D.G., Mcevoy, P.B., Nyrop, J.P., Reichard, S.E.H., Rice, K.J. and Tolin, S.A. (2002). 'Predicting invasions of nonindigenous plants and plant pests' (National Academy Press, Washington, DC).

Parker, I.M., Simberloff, D., Lonsdale, W.M., Goodell, K., Wonham, M., Kareiva, P.M., Williamson, M.H., Von Holle, B., Moyle, P.B., Byers, J.E. and Goldwasser, L. (1999). Impact: toward a framework for understanding the ecological effects of invaders. *Biological Invasions* 1, 3-19.

Parker, C., Caton, B.P. and Fowler, L. (2007). Ranking nonindigenous weed species by their potential to invade the United States. *Weed Science* 55, 386-97.

Pysek, P., Richardson, D.M., Rejmanek, M., Webster, G.L., Williamson, M. and Kirschner, J. (2004). Alien plants in checklists and floras: towards better communication between taxonomists and ecologists. *Taxon* 53(1), 141-3.

Richardson, D.M., Pysek, P., Rejmanek, M., Barbour, M.G., Panetta, F.D. and West, C.J. (2000). Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions* 6, 93-107.

Reichard, S.H. and Hamilton, C.W. (1997). Predicting invasions of woody plants introduced into North America. *Conservation Biology* 11, 193-203.

Sadhra, S.S. and Rampal, K.G. (1999). 'Occupational health: risk assessment and management'. (Blackwell Publishing, Oxford).

Suter, G.W. II (2006). 'Ecological risk assessment' 2nd edition. (CRC Press, Boca Raton, FL).

World Trade Organization (2007). 'WTO analytical index: guide to WTO law and practice' (Cambridge University Press, Cambridge).