Can a plant's cultural status and weed history provide a generalised weed risk score?

Roderick P. Randall
Invasive Species, Biosecurity and Regulation, Department of Agriculture and Food, South Perth, Western Australia 6151, Australia
(rod.randall@agric.wa.gov.au)

Summary Previous editions of *A Global Compendium of Weeds* (Randall 2012), whilst providing references relat ing to a plant's weed history, have not provided any real guidelines about the level of risk these species pose.

Using the extensive dataset behind the *Compendium,* including data indicating a plant's means of entry to new environments, dispersal pathways and impacts this paper considers the possibility of a generalised weed risk score.

Of the 40,874 referenced weeds listed in the database 9855 have a score indicating their level of weed risk. With further data collection more species can easily attain a score, and for species with no score the system is simple to use even without the underlying dataset.

Keywords *Global Compendium of Weeds*, database, risk, introduction, naturalised, environmental, agricultural, score, ranking.

INTRODUCTION

For several decades researchers have looked at various groupings of plant biological characteristics to determine those most useful in assessing a species' weed potential (Richardson *et al*. 1990, Rejmánek 1995, Rejmánek and Richardson 1996, Reichard and Hamilton 1997). Later publications looked at various other combinations of criteria to try and predict weediness, such as taxonomy (Daehler 1998, Diez *et al*. 2009), impacts and environment (Pyšek *et al*. 2012), weed groupings (Morin *et al.* 2013), comparing naturalisation rates (Duncan *et al.* 2004) and distribution traits (Pyšek *et al*. 2009), among others.

This paper considers the question: Could a database containing a relatively small range of dispersal, cultural (use) and weed status attributes for plants produce a generalised weed risk score?

Phases For a plant species to become a problematic weed in a new region it must progress through three phases:

1. **Enter** and establish a population in the new environment
2. **Disperse** throughout that environment normally from the initial population/s
3. **Impact** on the environment or human activities within its new range

The entry of a species into a new environment is considered an incursion. Once a species has become established and starts to spread without human assistance it can be considered naturalised and a cycle of dispersal, establishment, naturalisation and dispersal continues throughout the new environment.

After naturalisation the impacts may be immediate and costly. Species of *Striga* or *Orobanche* can have huge financial ramifications, restricting access to some markets just by the mere knowledge of their introduction into a previously clean cropping region. Most weeds, however, do not begin to accumulate noticeable impacts until years after their incursion, normally when public awareness starts to occur (Figure 1). With effective surveillance systems in place detection should hopefully occur in the first incursion phase. Detection in later phases usually means eradication is no longer an option and management systems need to be developed to cope with the new species.

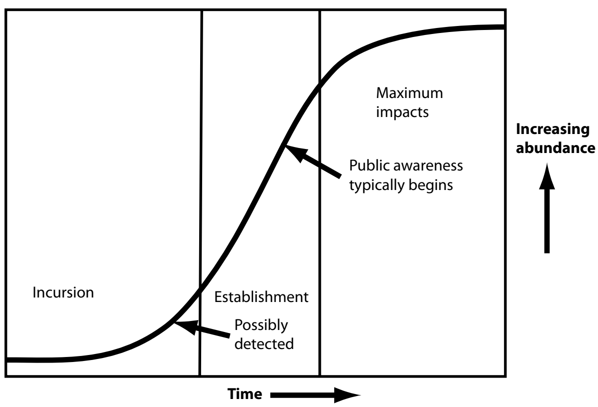

Figure 1. A typical 'Invasion Curve' with the phases of an organism's Incursion (Entry), Establishment (Dispersal) and Maximum impacts.

Rather than relying on a number of biological or ecological traits this proposed system considers the outcomes of a plant species' interaction with humans and its observed behaviour around the world. This approach is supported in part by a number of studies that concluded that the single most useful characteristic of a plant in predicting weediness is its behaviour as a 'Weed Elsewhere' (Gordon *et al*. 2008, Daehler and Carino 2000, Daehler *et al*. 2004, Mack 1996, Parker *et al.* 2007, Rejmánek *et al.* 2005). The 'Weed Elsewhere' question is not a function of a plant's biology or ecology but rather how a species' behavior is perceived by humans.

Consequently a number of the database characteristics used in this study are drawn from attributes that are closely allied to human behaviors, observations and values. This includes categories like 'intentionally grown' or 'considered to have medicinal values' and other categories of human values where these plants interact and impact, their uses and how they are dispersed by people.

MATERIALS AND METHODS

The Department of Agriculture and Food, Western Australia currently maintains a database that contains over 1.2 million plant names referentially linked by their taxonomy to 4.62 million bibliographically referenced records (Randall 2016). These referenced records contain data on hundreds of thousands of plant species, including their cultivated and naturalised status, dispersal mechanisms and their weed status from various regions around the world.

For the purposes of this system a species incursion risk is determined by (1) its ability to enter new environments, (2) to spread and establish over wider areas by its dispersal options and lastly (3) by its impacts in these environments. By matching selected database categories to these three phases and placing a value on each category a relatively simple scoring system was devised with the intention of producing a generalised weed risk score.

Not all categories were given the same value. Within the 'entry' phase the score value allocated was based on the percentage of known weeds as a proportion of the total species number within each category. Essentially the weediest categories were weighted with a higher score (Table 1). Within the 'dispersal' phase all categories were given the same score as there is a much higher percentage of weeds in each of these categories. In the 'impact' phase the 'invasive' category was valued the same as 'environmental' and 'agricultural' for reasons discussed later.

Fifteen categories, five in each phase (Table 1), were selected from those available in the plants database (Randall 2016) with the end result that 189,996 plant species were represented in at least one of the fifteen categories.

The database categories were selected to try and separate the more significant end of the weed spectrum rather than try to generate an all inclusive ranking of the many thousands of 'minor weeds', many of which may never generate the level of risk to warrant an eventual score.

Within each phase the scoring is additive, with the subsequent scores of the three phases multiplied. The maximum score any phase can add up to is four, with the maximum score therefore being 64:

<u>WEED RISK SCORE</u> = <u>ENTRY</u> (A + B + C + D + E) x <u>DISPERSAL</u> (F + G + H + I + J) × <u>IMPACT</u> (K + L + M + N)

For a species to generate a weed risk outcome it needs at least one score in each of the three phases, otherwise its outcome will be zero. Essentially the system assumes that if a species, has little ability to enter new regions, does not manage to easily disperse or has no definable impacts then it would not be considered a significant weed.

Scores were then generated for 9855 eligible species, which were then exported and graphed to try and determine the most effective allocation of the scores for a range of four risk outcomes.

A number of different cut offs between each category were originally tested in a graph of the risk ranking versus species distribution (Figure 2). The final distribution cut offs selected were those shown in Figure 3 (the blue lines) which closely follows the logarithmically reducing relationship between the categories, as illustrated in Figure 2.

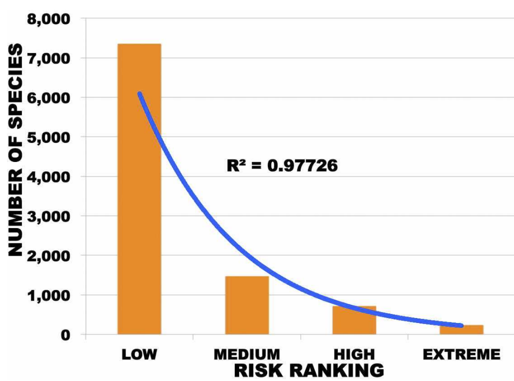

Figure 2. Risk ranking versus number of species per rank.

Table 1. The system categories, definitions, total species numbers from database, weed percentages and allocated values for each.

PHASE	CATEGORY		No. of species	% weeds	Value
ENTRY	Introduction via the most important pathways (human mediated)				
A	Pastures	includes fodder, forage, silage, hay, straw	3,170	67.0	1.3
B	Forestry	timber, pulp, firewood, shade, utility	1,422	61.8	1.2
C	Herbal	plants considered of medicinal use	32,842	37.5	0.6
D	Crops	edible plants	26,270	31.1	0.6
E	Ornamentals	grown by people for all sorts of reasons	149,636	13.1	0.3
DISPERSAL	Significant dispersal mechanisms				
F	Human dispersed	Sown as crops, garden plants, sold, bartered Intentional dispersal via human mediated pathways including machinery and vehicles	53,614	30.8	0.8
G	Animal dispersed	Includes all records of dispersal by animals	4,953	80.6	0.8
H	Wind or Water	When recorded as dispersed by either	4,004	89.1	0.8
I	Contaminants	Contaminants in seed for sowing, hay, fodder, other agricultural products	5,777	100	0.8
J	Escapees	Plants that escape from rehab sites, crops, orchards, gardens, collections etc	5,902	100	0.8
IMPACT	The most significant or potentially significant impacts				
K	Agricultural weed	Known weed of an agricultural crop	15,927	100	1
L	Environmental weed	Impacts are reported on the natural environment	5,471	100	1
M	Noxious weed	Quarantine species or Prohibited species as well as Declared, Noxious or Pest plants	4,238	100	1
N	Invasive	Species that can impact significantly on the environment or agricultural (primary production) operations, are seen to be actively spreading, often with impacts increasing	6,076	100	1

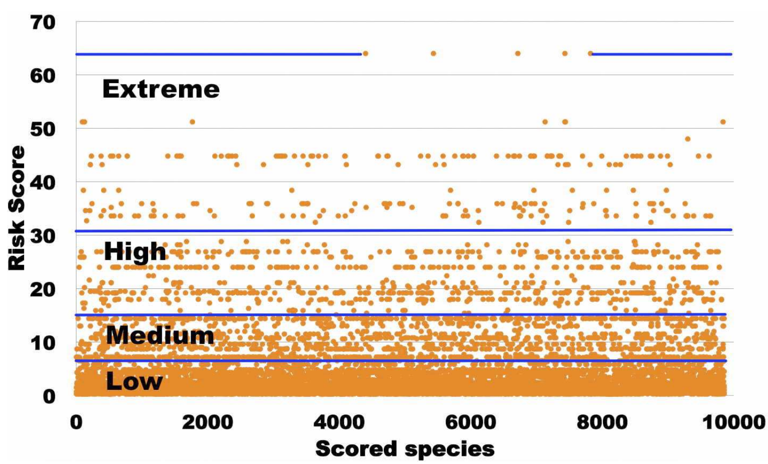

Figure 3. Risk score versus number of scored species (n = 9855) showing allocated risk ranks (blue lines).

Table 2 lists these cut off scores and includes the numbers and percentages of species in each rank category and as a percentage of all known weeds.

There is also an effective fifth rank outcome, that of 'Unscored', which includes all those species for which insufficient data exist to allow a score to be calculated. This outcome contains 180,141 species with at least one category captured in at least one phase.

RESULTS

The number of species in each of the four risk categories is outlined in Table 2. The 235 'Extreme' risk species identified by this process are all well known global weeds with the top 13 of this 'extreme' group containing very well known invasive species (Appendix 1).

Some of the 180,141 'Unscored' species could potentially generate a risk ranking with further research and therefore could be considered Data Deficient. However some species may never generate a risk score, these species could potentially be considered Non-Weedy species.

The only available measure of the likelihood of a species being Data Deficient, or a Non-Weed, is to consider the total number of records for that species. The more records a species has then potentially the lower is its likelihood of being data deficient. The implication of this is that species with no calculated risk rank, no weed references and numerous non-weed related records could effectively be considered non-weedy.

To determine just how many species may meet this potential Non-Weedy outcome the number of database records for all species with at least one phase score in the system, and no weed references (140,298), were exported and graphed to determine their distribution in a similar manner to the risk scored species (Figure 4).

While a more definitive analysis is needed, preliminary work implies that those species referenced 30 or more times with no actual weed references in the *Plant Database*, could be considered as candidate 'non-weeds'.

In terms of numbers of species, 30 or more references encompasses 1622 candidates, shown as points above the blue line in Figure 4. This also implies that the remaining 138,676 species from Table 2 could be considered data deficient.

Whilst interesting, this aspect of the *Plants Database* requires more statistical analysis and verification and clarification of the species names, including the determination of any synonymy and correction of spelling errors.

Table 2. Allocated Risk Ranks and species statistics for each.

Risk Score	Risk Rank	No. of species	% of species scored (n = 9855)	% of all known weeds (n = 40,874)
Zero	Unscored	180,141	–	–
>0–6	Low	7,437	75.4	18.19
7–15	Medium	1,472	14.9	3.60
16–31	High	712	7.2	1.74
32–64	Extreme	235	2.3	0.57

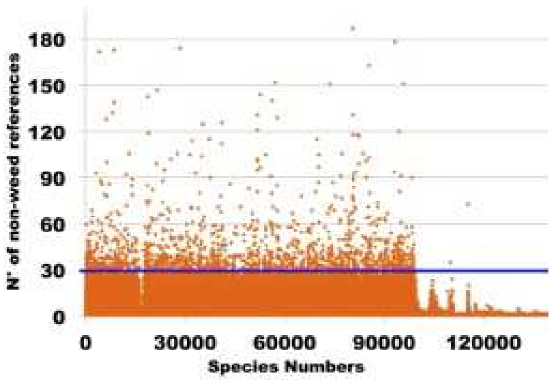

Figure 4. Number of references (for no weed reference species) versus species numbers (n = 140,298). Blue line = 30 records (vertical axis truncated at 200 records).

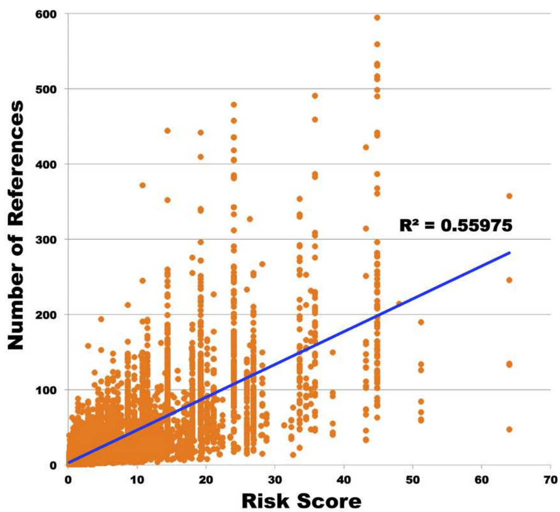

Figure 5. Species risk score versus the number of Weed References in the *Plants Database* (n = 9855).

That so many species are ranked in the database is an indication of the amount of relevant data held. Currently the database contains over 396,000 weed related records and another 4.26 million records relating to various other aspects of the species it holds.

With such a large number of weed related records it was considered that a potential relationship may exist between the number of weedy records for any species in the database and its ranked outcome (Figure 5). While there is a reasonably strong relationship between the risk score and the number of weed references for the dataset, there is still a significant degree of variation. The effect of cascading references (reference not a primary source) would have an impact on this approach.

The numerous non-weedy data sources also play a part in providing viable outcomes but a similarly plotted figure comparing instead the total number of references versus risk score showed no clear relationship.

DISCUSSION

Definitions and Categories Determining the correct category for a species from the literature can be difficult, sometimes impossible, so the descriptive term used by the authors is generally left unchanged. Therefore the score used in this system will more likely reflect the utility of the descriptive term used rather more than its formal definition might suggest. This is an issue that currently affects much of invasive species science and while a widely cited definitions paper has been written, those definitions are not well applied (Richardson *et al*. 2000).

'Invasive' is a term that is widely used and often misrepresented in the literature. Quite often the terms agricultural or environmental weed would be more informative and appropriate. Currently 'Invasive' scores the same as these more defined impacts as the majority of sources in the literature do not apply the term correctly.

Many plant species reported as environmental weeds could often be more correctly considered reports of naturalisation. Environmental impacts are harder to ascertain than agricultural and a great deal of work needs to be done to describe and quantify the full extent and costs of the environmental impacts of weeds. For the most part however species reported as environmental and agricultural weeds have been accepted as such unless the criteria as explained in the source indicates a clear mismatch.

Dispersal The vast majority of plant propagules are dispersed over relatively small distances, normally in metres to tens of metres (Cain *et al.* 2000, Horn *et al.* 2001, Nathan and Muller-Landau 2000, Tackenberg *et al.* 2003). This makes the human mediated dispersal of plants and plant propagules a significant class of global dispersal and is the reason why most natural means of dispersal were ignored.

Unsurprisingly a large proportion of species in each of the dispersal categories are also recorded as weeds. The same score was allocated to each as no single dispersal means is particularly important but multiple means of dispersal do mean an increased level of propagule pressure (Lockwood *et al*. 2005, Simberloff 2009). Furthermore demand for ornamental species is often driven by fashion, advertising and pricing, human foibles that have led to many successful plant incursions (Dehnen-Schmutz *et al*. 2007).

Unscored Species Many species in the database are yet to have a risk score calculated but may well do so as further data are incorporated. While the *Plants Database* (Randall 2016) is certainly not complete, it does have weed risk scores determined for 9855 weed species which represents a significant saving in time and resources for potential users of this system. It should also be remembered that many of these species may never meet enough criteria to enable them to attain a risk score, these being non-weedy plants.

For any species not yet scored, and no *Plants Database* available, Appendix 2 provides an example of three species scored with an example/template table that could be used to collate the information and tally up the scores. Any species currently unscored could have its risk rank determined using this table; the real effort of course lies in searching through references for the relevant data to complete the category scores within each phase.

The issue of non-weedy species being potentially identifiable by this system will be further investigated.

Benefits One of the biggest issues of the majority of 'ranking' or 'prioritisation' systems is that most have no species banked (assessed) in the system for others to use right away. With many thousands of species already assessed this proposed ranking system is likely to meet the requirements of most users without their having to research any, or many, species. As data collection in the database continues the numbers of species with ranked risks will continue to increase over time.

Consider the Climate While the generalised weed risk score provides a good indication of a species' weed potential globally, any potential risk should be considered in respect of a species' ability to naturalise in the region of concern. If a suitable climate exists and there are no cultural, agronomic or management restraints against the establishment or spread of the suspect species then the systems risk rank outcome could be considered appropriate and acted on accordingly.

However if there are no suitable climates available then the likelihood of the assessed species establishment in the region is likely to be low. There are very few plant species that are globally invasive and those are all well documented and hence already captured in this system.

A simple system that could be employed to modify the risk score based on an analysis of its preferred climate would be to apply a discount based on the degree of the climate match. For example a good climate match would mean no discounting of the risk score, whereas a moderate climate match could reduce the score by 25%, a low match by 75% and a poor climate match could reduce the score by 95%.

The degree to which a species' climate requirements are met in a new environment should therefore provide a level of discount that could be applied to its generalised risk score.

For a state like Western Australia, which covers over 30° of latitude across a range of climates, it is unlikely many species scores would be discounted; there is generally always somewhere a plant could survive and thrive. However in a region like Tasmania many tropical species may find their risk ranks heavily discounted.

Microclimates It is also important not to overlook potentially valuable microclimates. Many plants can become weeds in production glasshouses in regions where they would not ordinarily establish or survive in the prevailing weather conditions outside. Under such circumstances the costs of such weeds to industry and consumers can be high.

A system like this will not identify a potential weed where no records of weediness exist. Even known weeds that have no incorporated references for agricultural or environmental impacts will remain unscored in this system. The intent is to highlight and rank the worst known weeds rather than try and rank all possible weedy species, known and unknown. That this system can determine risk rankings for 33% of the world's known weeds is testament to the many and varied impacts that so many plant species have globally.

Overall this system could provide a great many users with a ready made, data populated, tool to assist in the ranking of their current weed species, representing significant savings in time and resources.

The potential savings for individuals and government, and the reduced downstream impacts to the environment and agriculture, by removing just a single 'extreme' risk species from a rehab program, cropping or landscaping project could be incalculable.

Any future editions of the *Global Compendium of Weeds* will also contain this system's risk ranking to provide readers with more data on species weed potential.

ACKNOWLEDGMENTS

The full results of this project are available as an Excel spread-sheet which contains all the category scores, phase scores, final risk scores and rank outcomes for 9855 plant species. This supplementary data spreadsheet is available on the conference USB drive along with the rest of the conference proceedings.

Many thanks to the thousands of authors whose work has gone into and made entirely possible this potential tool and lastly to the very useful suggestions from the reviewer.

REFERENCES

Cain, M.L., Milligan, B.G. and Strand, A.E. (2000). Invited special paper: long-distance seed dispersal in plant populations. *American Journal of Botany* 87 (9), 1217-27.

Daehler, C.C. (1998). The taxonomic distribution of invasive angiosperm plants: ecological insights and comparison to agricultural weeds. *Biological Conservation* 84, 167-80.

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 93-102.

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

Dehnen-Schmutz, K., Touza, J., Perrings, C. and Williamson, M. (2007). A century of the ornamental plant trade and its impact on invasion success. *Diversity and Distributions* 13, 527-34.

Diez, J.M., Williams, P.A., Randall, R.P., Sullivan, J.J., Hulme, P.E. and Duncan, R.P. (2009). Learning from failures: testing broad taxonomic hypotheses about plant naturalization. *Ecology Letters* 12 (11), 1174-83.

Duncan, R.P., Randall, R.P. and Williams, P.A. (2004). Comparative naturalization rates in the exotic floras of Australia and New Zealand. Paper presented at the fourth Southern Connections Conference, 19–23 January, University of Cape Town, Cape Town, South Africa.

Gordon, D.R., Onderdonk, D.A., Fox, A.M., Stocker, R.K. and Gantz, C. (2008) Predicting invasive plants in Florida using the Australian Weed Risk Assessment. *Invasive Plant Science and Management* 1 (2), 178-95.

Horn, H.S., Nathan, R. and Kaplan, S.R. (2001). Long-distance dispersal of tree seeds by wind. *Ecological Research* 16, 877-85.

Lockwood, J.L., Cassey, P. and Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *TRENDS in Ecology and Evolution* 20 (5), 223-8.

Mack, R.N. (1996). Predicting the identity and fate of plant invaders: emergent and emerging approaches. *Biological Conservation* 78, 107-21.

Morin, L., Paini, D.R. and Randall, R.P. (2013). Can global weed assemblages be used to predict future weeds? *PLoS ONE* 8 (2): e55547. doi:10.1371/journal.pone.0055547

Nathan, R. and Muller-Landau, H.C. (2000). Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *TREE* 15 (7), 278-85.

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.

Pyšek, P., Jarošík, V., Hulme, P.E., Pergl, J., Hejda, M., Schaffner, U. and Vilà, M. (2012). A global assessment of invasive plant impacts on resident species, communities and ecosystems: the interaction of impact measures, invading species' traits and environment. *Global Change Biology* 18, 1725-37. doi:10.1111/j.1365-2486.2011.02636.x

Pyšek, P., Jarošík, V., Pergl, J., Randall, R., Chytrý, M., Kühn, I., Tichý, L., Danihelka, J., Chrtekjun, J. and Sádlo, J. (2009). The global invasion success of central European plants is related to distribution characteristics in their native range and species traits. *Diversity and Distributions* 15 (5), 891-903.

Reichard, S.H. and Hamilton, C.W. (1997) Predicting invasions of woody plants introduced into North America. *Conservation Biology* 11 (1), 193-203.

Rejmánek, M. (1995). What makes a species invasive. *In* Pyšek, P., Prach, K., Rejmánek, M. and Wade, P.M. (eds), 'Plant invasions', pp. 3-13. (SPB Academic Publishing, The Hague, Netherlands).

Rejmánek, M. and Richardson, D.M. (1996). What attributes make some plant species more invasive? *Ecology* 77, 1655-60.

Rejmánek, M., Richardson, D.M., Higgins, S.I., Pitcairn, M. and Grotkopp, E. (2005). Ecology of invasive plants: state of the art. *In* Mooney, H.A., McNeely, J.A., Neville, L., Schei, P.J. and Waage, J. (eds), 'Invasive alien species: a new synthesis', pp. 104-61. (Island Press, Washington, DC).

Randall, R.P. (2016). 'Plants Database'. Department of Agriculture and Food, Western Australia. Unpublished database (4.62 million records).

Randall, R.P. (2012). 'A Global Compendium of Weeds, Second Edition', 1124 pp. (Department of Agriculture and Food, Western Australia).

Richardson, D.M., Cowling, R.M. and Le Maitre, D.C. (1990). Assessing the risk of invasive success in *Pinus* and *Banksia* in South African mountain fynbos. *Journal of Vegetation Science*, 1 629-42. doi: 10.2307/3235569

Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D. and West, C.J. (2000). Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions* 3 (6), 14-93.

Simberloff, D. (2009). The role of propagule pressure in biological invasions. *Annual Review of Ecology, Evolution and Systematics* 40, 81-102.

Tackenberg, O., Poschlod, P, and Bonn, S. (2003). Assessment of wind dispersal potential in plant species. *Ecological Monographs* 73 (2), 191-205.

Appendix 1. Top 13 Ranked Weeds (the next 75 species all score the same at 44.8).

Species	Global Risk Score
Gleditsia triacanthos L.	64
Leucaena leucocephala (Lam.) De Wit.	64
Parkinsonia aculeata L.	64
Prosopis glandulosa Torr.	64
Robinia pseudoacacia L.	64
Acacia farnesiana (L.) Willd.	51.2
Acacia nilotica (L.) Willd. ex Delile	51.2
Calotropis procera (Aiton) W.T.Aiton	51.2
Pithecellobium dulce (Roxb.) Benth.	51.2
Prosopis juliflora (Sw.) DC.	51.2
Prosopis pallida (Humb. & Bonpl. ex Willd.) Kunth	51.2
Ziziphus mauritiana Lam.	51.2
Trifolium pratense L.	48

Appendix 2. Generic Weed Risk Ranking System as applied to three species.

Phase Categories (No = Zero)	***Malva sylvestris* L.** **tall mallow**	***Abies pinsapo* Boiss.** **Spanish fir**	***Emex australis* Steinh.** **doublegee**
ENTRY			
Pastures (Yes = 1.3)	1.3	0	0
Forestry (Yes = 1.2)	0	1.2	0
Crop or Edible plant (Yes = 0.6)	0.6	0	0.6
Herbal/Medicinal (Yes = 0.6)	0.6	0.6	0
Ornamental (Yes = 0.3)	0.3	0.3	0.3
Sum	2.8	2.1	0.9
DISPERSAL	×	×	×
Human Dispersed (Yes = 0.8)	0.8	0.8	0.8
Animal Dispersed (Yes = 0.8)	0.8	0	0.8
Contaminant (Yes = 0.8)	0.8	0	0.8
Escapee (Yes = 0.8)	0.8	0	0
Wind and/or Water (Yes = 0.8)	0	0	0.8
Sum	3.2	0.8	3.2
IMPACT	×	×	×
Agricultural Weed (Yes = 1.0)	1	0	1
Environmental Weed (Yes = 1.0)	1	1	1
Noxious Weed (Yes = 1.0)	1	0	1
Invasive (Yes =1.0)	1	1	1
Sum	4.0	2.0	4.0
Weed Rank Score = Entry × Dispersal × Impact	**35.84**	**3.36**	**11.52**
Generic Risk Ranking	**Extreme**	**Low**	**Medium**