Report Cover Page

ACERA Project		
0904		
Title		
Quantifying effects of uncertainty in responses to key weed risk assessment scoring questions		
Author(s) / Address(es)		
Peter Caley, ABARES, Department of Agriculture Fisheries and Forestry, Canberra, ACT		
Material Type and Status (Internal draft, Final Technical or Project report, Manuscript, Manual, Software)		
Final report		
Summary		
This report addresses several of the priorities identified during the first phase of this project relating to uncertainty in weed risk management systems by John Virtue (South Australian Dept. of Land, Water & Conservation). In particular, it assesses the effect of uncertain knowledge on our ability to identify invasive plants using classification systems. The Australian Weed Risk Assessment (WRA) system uses information on a species' current weed status in other parts of the world, climate and environmental preferences, and biological attributes to predict whether it would be a serious weed (i.e. invasive) if introduced to Australia. Documented evidence of invasiveness elsewhere is considered the best predictor of being a weed in Australia, though such evidence is not always available or is uncertain. This report assesses how critical it is to know whether a species is invasive elsewhere when undertaking a weed risk assessment, and the implications when such information is unknown or uncertain. The results show that when knowledge of invasiveness elsewhere is unknown, the probability of being invasive is generally reduced. This suggests that there is indeed some information to assist with weed classification when a species' invasiveness elsewhere is unknown. The report also helps to quantify how much opportunity in terms of residence time would be considered sufficient for a species to demonstrate invasive potential. The report estimates that based on past introductions in temperate climes (e.g. woody perennials introduced to Brandenburg [Germany] and South Australia), at least a century of observation would be required before a woody perennial species could be declared incapable of becoming a major weed with any certainty. This time could probably be reduced substantially by experimental trials that fast-track the naturalisation process. Using the original data used to "train" the WRA system, though excluding information on documented invasiveness for a species introduced outside their native range, a classification model was built that performed well in identifying weeds, though poorly at identifying non-weeds. This contrasts somewhat with the previous belief that much of the classification ability of the current WRA system arises from information pertaining to a species' introduction record and invasive status elsewhere. The report identifies some challenges in assembling the data needed to reassess the current and proposed WRA systems.		
ACERA Use only	Received By:	Date:
	ACERA / AMSI SAC Approval:	Date:
	DAFF Endorsement: ( ) Yes ( ) No	Date:

Quantifying effects of uncertainty in responses to key weed risk assessment scoring questions: ACERA 0904

Peter Caley

Australian Bureau of Agricultural and Resource Economics and Sciences

Final report

December 2010

FINAL VERSION

Australian Government
Australian Bureau of Agricultural and Resource Economics and Sciences

Acknowledgements

This report is a product of the Australian Centre of Excellence for Risk Analysis (ACERA). Jeanine Baker, Ian Gordon, Quentin Hart, Greg Hood, Mick McCarthy and an anonymous reviewer provided comment on draft versions. In preparing this report, the authors acknowledge the financial and other support provided by the Department of Agriculture, Fisheries and Forestry (DAFF).

Table of contents

Tables

Figures

1. Executive Summary

The Australian Weed Risk Assessment (WRA) system uses information on a species current weed status in other parts of the world, climate and environmental preferences, and biological attributes to predict whether it would be a serious weed (i.e. invasive) if introduced to Australia. Documented evidence of invasive behaviour elsewhere is considered the best predictor of being a weed in Australia. This reports shows that:

- When invasive behaviour elsewhere is unknown or uncertain, the risk of a species being invasive when introduced to Australia is intermediate between species that are known to be invasive elsewhere (higher risk) and species that are known not to be invasive (lower risk).
- The WRA system's scoring takes a conservative approach when assessing a species whose invasive behaviour elsewhere is unknown or uncertain, in that it minimises the chances of mistakenly introducing a weed.
- At least a century of observing no invasiveness post-introduction would be required before a plant species could be declared incapable of becoming a major weed with any certainty.
- An alternative classification model built using data that excludes questions pertaining to invasiveness behaviour elsewhere performed well in identifying weeds, though it was poor at identifying non-weeds.
- There is a need to reassess the current WRA system using independent data, rather than via cross-validation with the data used to build the current system. This may not be straightforward, as the pool of plants introduced to Australia from which known invasive and non-invasive plants may be identified has been subject to selection bias from ongoing screening.

2. Introduction

Invasions of alien plants incur substantial economic and environmental costs in many countries; hence avoiding the intentional introduction of additional invasive plants is a high priority. The Australian Weed Risk Assessment (WRA) system (Pheloung *et al.* 1999) and a growing number of modified implementations (e.g. Krivanek and Pysek 2006, Gordon *et al.* 2008, Nishida *et al.* 2009) use information on a species' current weed status outside of its native range, climate and environmental preferences, and its biological attributes to predict whether it would be a serious weed (i.e., invasive) if introduced to a defined new environment. In a review of screening models for invasive species, Hayes and Barry (2008) identified climate match, evidence for invasiveness elsewhere (meaning outside a species' native range) and propagule pressure as the three consistent predictors of invasiveness. Data for these three predictors of invasiveness for introduced species are sometimes unavailable or uncertain. In particular, the strength of evidence for invasiveness elsewhere depends on the location, extent and timing of introductions—which will interact with factors such as the fecundity and generation interval of the plant—and the intensity of surveillance. For example, the observed time-lag between introduction and evidence of naturalisation (a precursor to becoming invasive) differs between annual and perennial plants (Caley *et al.* 2008), and such time-lags are shorter in tropical climates (Daehler 2009).

Current guidelines for using the WRA system (Gordon *et al.* in press) recognise that a potential weed must have had opportunities to demonstrate invasiveness. A lack of evidence results in a "no" or "unknown" response, depending on the amount of information available on the species. Known residence time (outside of native range) in the absence of demonstrated invasiveness is the major piece of information underpinning the inference that a species has not previously been invasive when introduced outside its native range. There is a need to define how much opportunity in terms of residence time is enough. An early study by Kowarik (1995) documented a mean time in excess of 100 years between introduction and naturalisation for the Brandenburg (Germany) introduced woody perennial flora. Caley *et al.* (2008) undertook a Bayesian analysis of the estimated time to, and probability of, recorded naturalisation for woody perennials introduced to South Australia through the nursery trade. A prior distribution on the time to naturalisation was taken from the Brandenburg records, and a prior distribution on the probability of naturalisation from the "Ten's Rule" of Williamson and Fitter (1996). The result was a mean lag time of 149 years and naturalisation probability of around 19%. Note that most if not all estimates of residence times recorded prior to naturalisation arise from introductions with very different introduction histories—different species are introduced in different numbers (propagule pressure), in different numbers of locations, and over different time periods. In contrast, experimental trials involving deliberate plantings (e.g. approach VIII of Mack 1996) should enable faster evidence of naturalisation to be recorded, by ensuring sufficient propagules are given the chance to demonstrate invasiveness in a more widely representative set of conditions.

'Invasiveness elsewhere' is typically treated as a dichotomous variable (either ‘Evidence of invasiveness elsewhere’ or ‘No evidence of invasiveness elsewhere’). In reality, a continuum of evidence will accumulate based on the timing, knowledge and extent of introductions and their fate (Figure 1). There are several possible scenarios leading to a species invasiveness elsewhere being considered unknown. First, it may be known unequivocally that there have been no introductions outside its native range, thus preventing any demonstration of invasiveness. Second, despite known introductions, there is no record of the fate of these introductions due to lack of monitoring. Third, introductions of a species might have been documented outside of its native range, but their timing and extent could be insufficient to demonstrate invasiveness. Lastly, we might not know whether or not there have been introductions outside of the native range.

For most species, a lack of observed range expansions when introduced outside their native range is used to infer that they are not invasive. These are essentially observational experiments with little control over influential variables such as the number and location of propagules released. Data are obtained either by actively monitoring known introductions outside of the natural range (‘watch and wait’), or as is more often the case, by ‘passively’ waiting for invasiveness to be observed and published in the accessible literature. The latter approach would lend itself to reporting bias. A designed alternative is active experimentation as described by Pattison and Mack (2009).

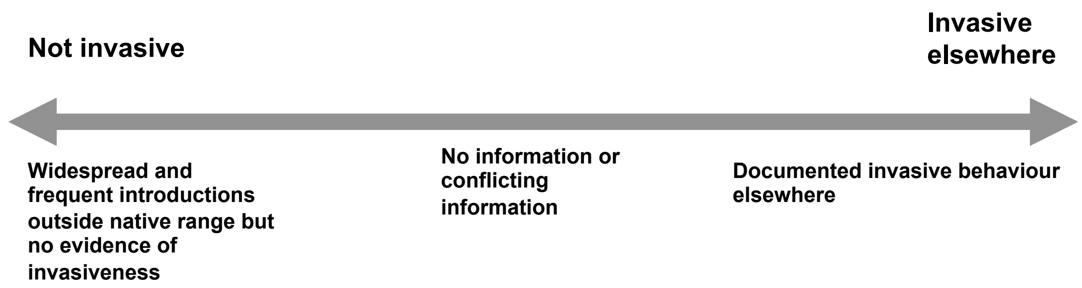

Figure 1. Diagrammatic representation of the continuum of evidence of invasiveness of a species when introduced outside of its native range.

The objectives of the current project were derived from those contained within the original (though uncompleted) ACERA Project 0904 undertaken by John Virtue (South Australian Dept. of Land, Water & Conservation). The objectives of the original project and their degree of completion prior to the current project (in square parentheses) were to:

- survey current government practices for addressing uncertainty in qualitative risk assessments [completed]
- review international best practice in uncertainty analysis of qualitative risk assessments [partially completed]
- develop a robust methodology for uncertainty analysis of qualitative risk assessments that could be broadly adopted by government [incomplete]
- incorporate social and perceptual inputs in the development, dissemination and uptake of uncertainty analysis [not done]
- develop a training module for dissemination amongst government [not done].

As part of the first objective, a questionnaire was sent to members of the National Weed Risk Management Forum (NWRMF) to gauge their understanding of uncertainty, which was further investigated in a national workshop of the NWRMF. The workshop mapped out sources of uncertainty in the weed risk management process and four key areas were identified as priorities for further development: (1) model uncertainty in Weed Risk Management (WRM) systems, including how this is influenced by the desired goal, context and endpoint; (2) addressing species knowledge gaps; (3) intervals and categorisation [of responses to weed attributed questions]; and (4) communicating uncertainty.

With the remaining project funds, and with the agreement of ACERA this re-scoped project sought to contribute to (1) and (4). In particular, because 'Invasiveness elsewhere' is considered such a key determinant of the likelihood of invasion in a new location, this study explores ways that evidence for invasiveness elsewhere can be described and quantified, and how it is handled in the current WRA system. It also helps quantify how much knowledge (as measured by residence time) is required before a species can be safely considered unlikely to become invasive. The revised objectives of this project were to:

- Quantitatively explore the variability and uncertainty in deterministic versus probabilistic weed risk assessments arising from the 'invasiveness elsewhere' question.

- Determine how estimated invasiveness may change over time arising from the accumulation of additional data with respect to evidence of invasiveness elsewhere.

3. Key terminology and concepts

3.1 Classification of alien plants

The Australian Weeds Strategy (Natural Resource Management Ministerial Council 2006) considers a weed "as a plant that requires some form of action to reduce its harmful effects on the economy, the environment, human health and amenity". In an attempt to avoid ambiguity, this definition needs to be placed within an invasive species context. Within the hierarchical scheme for classification of alien plants proposed by Pyšek *et al.* (2004), weeds can be considered a subset of invasive plants which themselves are a subset of naturalised plants. The definitions (abbreviated) from Pyšek *et al.* (2004) for these categories are as follows:

Naturalised—alien plants that sustain self-replacing [sustaining] populations.

Invasive—a subset of naturalised plants that produce reproductive offspring, often in very large numbers, at considerable distances from the parent plants, and thus have the potential to spread over a large area. [The ability to have deleterious effects on pre-existing biota is usually implicit in the definition of invasiveness].

Weeds—plants that grow in sites where they are not wanted and which have detectable economic or environmental impact or both. [Implicitly invasive].

Other commonly used terms are:

Residence time—time since first successful introduction into a new range (Wilson *et al.* 2007).

Propagule pressure—also termed “introduction effort”, is a composite measure of the number of individuals released into a non-native region (Lockwood *et al.* 2005).

For the purposes of this study, a weed is considered to be synonymous with invasive, as implicit in the definition of a weed is the ability to establish (i.e. invade) in at least some habitats (where they are considered unwanted).

4. Methods

4.1 Making inference on invasiveness elsewhere

Uncertainty in knowledge of a species' invasiveness elsewhere interacts with the introduction record in informing prospective classification of invasiveness (Table 1). Species with known invasiveness elsewhere are of little interest, as they are rarely accepted for importation under the current system unless they have a very poor climate match within the recipient region. This would be unusual, as Australia has most climates represented.

Table 1. Matrix showing the degree to which knowledge of a species' invasive record elsewhere interacting with its introduction record elsewhere provides information for predicting invasiveness.

Invasive record elsewhere	Introduction record elsewhere		
	Known introduced	No introduction record	Known not introduced
Known invasive	Informative	Informative[1]	-
Unknown (no data)	Informative?	Uninformative?	Uninformative
Not invasive*	Informative	-	-

* In the presence of adequate monitoring.
- Mutually exclusive events

For species with a known record of introductions elsewhere but no evidence of invasiveness (when monitoring is assumed adequate to detect invasiveness), assessors must decide whether to classify the species as 'not invasive' elsewhere or retain its invasive status as unknown. This should occur after assessors have weighed up whether opportunity to demonstrate invasiveness has occurred. 'Opportunity' encompasses the extent of introductions in terms of the range of environments and number of propagules introduced, along with the time since introduction (residence time). The measure 'propagules introduced × spatial extent of introductions × environmental suitability × residence time' should capture most of the opportunity to demonstrate invasiveness. Formally this could be represented as an integral of introduction rate over space, environment and time. Assessors reckon this process mentally. The classification of 'not invasive elsewhere' is based on the failure to demonstrate invasiveness despite what is considered ample opportunity (and monitoring)—but how much opportunity constitutes enough?

For species with unknown introduction records and hence unknown invasive records elsewhere, "absence of evidence [in this case of invasiveness elsewhere] does not equate with evidence of absence" is the oft-quoted maxim. Whilst theoretically true, this may be overly restrictive in practice. For example, there is presumably a non-zero prior probability that a species with an unknown introduction record has in fact been the subject of unrecorded introductions. These introductions may have failed to establish, hence there is no record of invasiveness one way or the other. This is especially true of regions that do not routinely collect introduction records. Here an assessment if made of how the current WRA system handles such uncertainty and or unknown information based on the wording of the questions and the guidelines for how they should be answered.

4.2 Accounting for delay in evidence of invasiveness

Two key quantities underlying the number of invasive plants are the probability that an introduced plant will naturalise (and possibly become invasive), and what is the time until this occurs following introduction. It is now well recognised that there may be a considerable delay between a species introduction, and subsequent demonstration of invasiveness (Crooks 2005). For plants, the delay is most pronounced for woody perennials as they have a longer generation interval. The distribution of the delay between introduction and naturalisation (a

[1] Some species are introduced accidentally (e.g. as contaminants)–there is no introduction record for these species though they may be recorded as invasive.

necessary precursor to becoming 'invasive') of past introductions is used as a means of estimating how the chances of future naturalisation decrease as residence time increases. The South Australian dataset collated by Caley *et al.* (2008) is used to estimate the time to, and probability of, naturalisation of introduced woody perennial plants. These data contain the estimated dates of the first introduction of 2230 woody plant species to South Australia up until 1983, based on their first appearance in nursery catalogues dating back as far as 1843 (Mulvaney 1991). Data on the time of first recorded naturalisation for these plants were obtained from Kloot (1986) and specimens lodged subsequently in the South Australian herbarium until early 2007 (Robyn Barker pers. comm.). Up until 2007, 188 of 2230 (8.4%) of woody perennials listed in nursery catalogues were recorded as having naturalised.

Whereas Caley *et al.* (2008) had only two categories of plants – those that can naturalise and those that cannot, here plants that can naturalise are further broken down into two types, “invasive” naturalisers and “non-invasive” naturalisers. The distinction is somewhat arbitrary as invasiveness is a continuum, but it serves a useful purpose as it is "major" weeds that we are most concerned with identifying. Kloot (1986) classified naturalised plants as either "major", "minor", or "inconsequential". Here, plants considered "minor" or "inconsequential" weeds are combined with those having no classification and categorised as naturalisers not considered invasive, and "major" weeds are categorised as naturalisers considered invasive. Where necessary, the species classification of Kloot (1986) was changed to reflect current declared weed species in South Australia.

The model parameterisation of the naturalisation process is:

p_0 = probability that plant will never naturalise;

p_1 = probability that plant will naturalise, but be considered non-invasive;

p_2 = probability that plant will naturalise, and be considered invasive;

T_1 = the time to naturalisation for non-invasive naturalisers with probability density function $f_1(t)$ and cumulative density function $F_1(t)$; and

T_2 = the time to naturalisation for invasive naturalisers with probability density function $f_2(t)$ and cumulative density function $F_2(t)$.

Let there be n recorded naturalisations of introduced plants considered non-invasive (minor or insignificant weeds), m recorded naturalisations of introduced plants considered invasive (major weeds), and q known introductions for plants that have not naturalised at the end of the observation period (right-censored observations). For the South Australian dataset, n =171, m =17, and q = 2,042. The likelihood function where θ represents the parameters to be estimated is:

$$L(\theta \mid data) = \prod_{i=1}^{m} p_1 f_2(t_i) \prod_{j=1}^{n} p_2 f_1(t_i) \prod_{k=1}^{q} \left(1 - p_1 F_1(t_i) - p_2 F_2(t_i)\right) \qquad \text{Eqn 1}$$

There is a positive relationship between the estimated probability of naturalisation and the estimated time to it occurring. In the current context, it is the time to naturalisation that is of interest. Without informed priors on the probability of naturalisation, the estimated time to naturalisation is poorly defined (Caley *et al.* 2008). Moderately informative priors were used: $p_1 \sim$ Uniform(0.01, 0.03) and $p_2 \sim$ Uniform(0.1, 0.3). That is, between 10 and 30 percent of introduced woody perennials will naturalise (though not become invasive), and between 1 and 3 percent of introduced woody perennials will naturalise and become invasive. The lower bound for these priors seem reasonable, as already nearly 10% of introduced woody perennials introduced to South Australia have naturalised, and nearly 1% of introductions are classified as major weeds, with more transitions expected. Both $f_1(t)$ and $f_2(t)$ were assumed to be Gamma distributed with uninformative priors for their rate and shape parameters (both set to 0.001).

The model was fitted to the observed data via Markov Chain Monte Carlo (MCMC) methods utilising the Metropolis–Hastings algorithm with informative priors on p_1 and p_2 as detailed above. Shape and rate parameter values for T_1 and T_2 were updated jointly within each step of the Markov chain. Normal proposal densities were used for all parameters. The acceptance rate for each step of the Markov chain was tuned for each parameter to be in the order of 70% by adjusting each proposal distribution. For the purpose of describing the posterior distributions for each of the parameters, the chain was run for 500,000 steps following a burn-in period of 50,000 steps, and then subsampled at every 100^{th} step. The degree of thinning was necessary as the chain 'mixed' quite slowly.

The code used to implement the analysis was run on simulated datasets with known parameters to confirm it was estimating parameters correctly.

4.3 Estimating effects of not knowing invasiveness elsewhere

4.3.1 Empirically from WRA system training dataset

The simplest approach to assessing the effect of not knowing invasiveness elsewhere is to estimate the probability (or likelihood) of an introduced plant being invasive in Australia based only on their invasive status elsewhere. This approach is somewhat rudimentary, as it ignores other plant attributes that may confer invasiveness. There are, however, arguments that invasiveness elsewhere is the single most important predictor of invasiveness (see Hayes and Barry (2008) and the Results section). In an operational setting, of greater interest is the posterior probability of a plant being invasive (Caley *et al.* 2006), which incorporates the prior probability of invasiveness and the likelihood. A standard application of Baye's rule gives the posterior probability of weediness given the weed elsewhere ('WE') question was not answered ('NA') as:

$$P(\text{Weed} \mid \text{NA to WE}) = \frac{P(\text{NA to WE} \mid \text{Weed})P(\text{Weed})}{P(\text{NA to WE} \mid \text{Weed})P(\text{Weed}) + P(\text{NA to WE} \mid \overline{\text{Weed}})P(\overline{\text{Weed}})} \quad \text{Eqn 2}$$

Here, $\overline{\text{Weed}}$ denotes "not a weed". Posterior probabilities for P(Weed | Yes to WE) and P(Weed | No to WE) were calculated similarly. Equation 3 was applied to the expert survey dataset (hereafter the 'training dataset') that was used to manually train the current WRA system (Pheloung *et al.* 1999). The dataset was generated by asking a small number of experts to answer questions (49 in total) pertaining to a species' current weed status in other parts of the world, its climatic and environmental preferences, and biological attributes. It contains assessment of 370 taxa, of which 286 (77%) were considered weeds, and 84 (23%) were not weeds. An extract of the questions is shown in Table 2. The lack of an answer is assumed to be because weediness elsewhere is either unknown (no data), or the data are conflicting. The data set contains multiple appraisals for some taxa by different experts. Where multiple appraisals occurred, these data were collapsed into a single record as follows. For logical questions with responses 'Yes', 'No', or 'NA', the majority logical response was chosen, regardless of the number of NAs. Where the logical responses of the experts were equally divided, again regardless of the number of NAs, then the answer to the question was deemed 'NA' (if the experts can't reach a majority agreement, then the answer is unknown).The final training dataset consisted of 370 taxa, of which 286 (77%) were classified as weeds, and 84 (23%) as non-weeds.

The approach used is similar to Caley *et al.* (2006) and involved:

1. Taking a bootstrap sample of weeds and non-weeds from the WRA training dataset. This provides a means of estimating uncertainty arising from natural variation.

2. Calculating the probabilities of responses to the 'Weed elsewhere' question conditional on weed status (either "Yes", "NA", or "No").

3. Calculating the posterior probabilities of weediness conditional on the response to the 'Weed elsewhere' question and prior probability of weediness using Equation 1.

Results were calculated for all possible prior probabilities of weediness for plants being assessed.

4.3.2 Via a statistical model classifier

The effect of uncertain knowledge in relation to weed elsewhere status was examined by comparing the performance of a classification model built using all variables in the training dataset, with a model fitted to the training dataset with the predictor variables pertaining to weediness elsewhere removed. Caley and Kuhnert (2006) modelled the WRA training data using a single classification and regression tree (CART) model, but much better predictions are possible using ensemble methods such as random forests (Breiman 2001). Random forest classifiers were chosen as implemented by the `randomForest` library in the software package R (R Development Core Team 2008). Briefly, a random forest model is fitted by fitting modified CART models to bootstrapped versions of the dataset, thus generating a 'forest' of trees. The modification to the CART model involves choosing a random subset of the explanatory variables to be considered for each "split". For classification, the 'forest' of trees each cast a vote for the predicted class, with the average across all trees used. Two random forest classifiers were built. The first was built using all the training data, while the second was built using the training data with questions relating to weediness elsewhere removed. The questions removed were Q3.01–Q3.05 which relate to evidence of naturalisation outside of native habitat, demonstrated weediness, and evidence of weedy congeners. Also removed were Q1.02, Q1.03, and Q2.05 which require knowledge of introductions outside of a species' native range (see Table 2).

There are many missing values within the WRA training dataset, which is unsurprising given the number of questions. The `randomForest()` function within the software package R uses imputation for missing values. The imputation method `na.roughfix()` imputes missing factors by replacing them with the most frequent level within each classification. This is unlikely to introduce bias into the results. The alternative method for handling missing values, `rfImpute()`, wasn't considered as it appears to lead to overly optimistic estimation of classification errors (Breiman 2003). Each forest consisted of 10,000 trees, which would be more than ample based on the recommendation of Hastie *et al.* (2009). Models were compared using measures of sensitivity and specificity. Sensitivity in this context is the probability that a plant species is invasive given the model predicts it to be invasive. Specificity is the probability that a plant species is not invasive given the model predicts it to be not invasive.

5. Results

5.1 Use of invasiveness elsewhere data in current WRA system

Information on climate match, evidence for invasiveness elsewhere, and propagule pressure are in Questions 2.01–2.05 ('Climate and distribution') and 3.01–3.05 ('Weed elsewhere') of the WRA (Table 2; Walton *et al.* 1998). Logical reckoning in the scoring system, based on interactions between the responses to these questions, is evident in the WRA scoring process arrived at via a ‘lookup’ table (Table 3) (Pheloung *et al.* 1999). For example, the degree of uncertainty in the quality of climate matching data (Question 2.02) in combination with the predicted climate suitability, determines the score resulting from being a weed elsewhere (either "Yes" or "No") (Table 3). For example, a predicted low climate suitability based on good data lessens the score assigned to a species known to be a weed elsewhere. Species for which their weed status elsewhere is 'Unknown' incur a maximum possible score by default. This is conservative, since it is equivalent to that of a species having evidence of being invasive elsewhere and having a high climate suitability rating based on good data. The WRA recognises the information content for the combination of a 'Yes' for Question 2.05 ('History of repeated introductions outside of native range ') and 'No' for Question 3.01 ('Naturalised beyond native range'): the result is 2 points being deducted from the score (Table 3). An uncertain ('?') response to Question 2.05 ('History of repeated introductions outside of native range') in combination with a 'No' to Question 3.01 ('Naturalised beyond native range') results in a single point being deducted (Table 3) – this indicates the current system considers this combination to be informative (in contrast to Table 1 which considers this combination of answers to not be possible).

Table 2. Extract from the Australian Weed Risk Assessment system (Pheloung *et al.* 1999) showing guidelines for interpreting the questions pertaining to 'Climate and distribution' and 'Weed elsewhere'.

Question	WRA Guidelines
1	**Domestication / cultivation**
1.01	**Is the species highly domesticated? If answer is “no” go to Question 2.01** The taxon must have been cultivated and subjected to substantial human selection for at least 20 generations. Domestication generally reduces the weediness of a species by breeding out noxious characteristics.
1.02	**Has the species become naturalised where grown?** Is a domesticated plant, which has introduced from another region, growing, reproducing and maintaining itself in the area in which it is growing. A “yes” answer to question 1.01 will be modified by the response to this question.
1.03	**Does the species have weedy races?** Only answer this question if the species you are assessing is a sub-species, cultivar or registered variety of a domesticated species. If the taxon is a less weedy subspecies, variety or cultivar, then there must be good evidence that it does not retain the capacity to revert to a weedy form. A “yes” answer to question 1.01 will be modified by the response to this question.

2	**Climate and distribution**
2.01	**Species suited to Australian climates (0-low; 1-intermediate; 2-high)** This question applies to any one Australian climate type, or more than one. Ideally, base the climate matching on an approved computer prediction system such as CLIMEX, BIOCLIM or Climate. If no computer analysis is carried out then assign the maximum score (2).
2.02	**Quality of climate match data (0-low; 1-intermediate; 2-high)** The score for this question is an indication of the quality of the data used to generate the climate analysis. Reliable specific data scores 2, general climate references scores 1, broad climate or distribution data scores 0. If a computer analysis was not carried out assign the maximum score of 2.
2.03	**Broad climate suitability (environmental versatility)** Score “yes” for this question if the species is found to grow in a broad range of climate types. Output from the climate matching program may be used for this question. Otherwise base the response on the natural occurrence of the species in 3 or more distinct climate categories. Use the map of climatic regions provided or one available in a comprehensive atlas.
2.04	**Native or naturalised in regions with extended dry periods** The species is able to grow in areas with rainfall in the driest quarter less than 25 mm. Plants from this group may potentially grow and survive in arid Australian conditions.
2.05	**Does the species have a history of repeated introductions outside its natural range?** This history should be well documented. A potential weed must have opportunities to show its potential. A score for Question 2.05 will modify the score for a “no” answer to Question 3.01. Species with repeated introductions that have not established are a lower risk.
3	**Weed elsewhere**
3.01	**Naturalised beyond native range** A naturalised species will be cited in floras of localities which are clearly outside of the native range. If the native range is uncertain and the known extent of the naturally growing plants is within the area of uncertainty then the answer is “don't know”.
3.02	**Garden/amenity/disturbance weed** The plant is generally an intrusive weed of gardens, parklands, roadsides, quarries, etc. This question carries less weight than 3.03 or 3.04. If a plant is listed as a weed in relevant references but the type of weed is uncertain or it is a minor weed — score “yes” for 3.02.
3.03	**Weed of agriculture/horticulture/forestry** The plant is generally a weed of agriculture/horticulture/forestry and causes productivity losses and/or costs due to control. This question carries more weight than 3.02. If a plant is listed as a weed in relevant references but the type of weed is uncertain or it is a minor weed — score “yes” for 3.02.
3.04	**Environmental weed** The plant is documented to alter the structure or normal activity of a natural ecosystem. This question carries more weight than 3.02. If a plant is listed as a weed in relevant references but the type of weed is uncertain or it is a minor weed — score “yes” for 3.02.
3.05	**Congeneric weed** Documented evidence that one or more species, with similar biology, within the genus of the species being evaluated are weeds.

Table 3. The 'Lookup' table from the Australian Weed Risk Assessment system (Pheloung *et al.* 1999) detailing how the score for a plant arising from its weed elsewhere status is determined based on knowledge of introduction history and predicted climate match with Australian conditions. Where the answer to Questions 3.01–3.05 is unknown the Default score is used. Question names in inverted commas are abbreviated (see Table 2 for more details). A higher score increased the likelihood that a species will be considered invasive.

Questions 3.01–3.05 answered "Yes"		Rating to question (0–low, 1–intermediate, 2–high)								Default
Inputs	Question 2.01 ("Climate suitability")	0	0	0	1	1	1	2	2	2
	Question 2.02 ("Data quality")	0	1	2	0	1	2	0	1	2
		Score								
Weed elsewhere questions to be scored	Question 3.01 ("Naturalised")	2	1	1	2	2	1	2	2	2
	Question 3.02 ("Garden weed")	2	1	1	2	2	1	2	2	2
	Question 3.03 ("Production weed")	3	2	1	4	3	2	4	4	4
	Question 3.04 ("Environment weed")	3	2	1	4	3	2	4	4	4
	Question 3.05 ("Congeneric weed")	2	1	1	2	2	1	2	2	2

Questions 3.01–3.05 answered "No"		Rating to question		
Inputs	Question 2.05 ("Repeated introductions outside of natural range")	?	N	Y
		Score		
Weed elsewhere question to be scored	Question 3.01 ("Naturalised")	-1	0	-2
	Questions 3.02–3.05 ("Weed")	0	0	0

5.2 Accounting for delay in evidence of invasiveness

As mentioned in the Methods, the ability of the analysis to converge on parameter estimates required the use of moderately informative priors on the probability of naturalisation, so the results should be interpreted with the validity of these priors in mind. That said, the choice of priors for the probability of naturalisation or becoming invasive are reasonable based on the literature. Using the medians of the posterior distributions for the parameters, the distribution of time to naturalisation for major weeds was Gamma (shape parameter=2.1, rate parameter=0.009), and for minor, insignificant and non-weeds, Gamma (shape parameter=4.9, rate parameter=0.07). The corresponding mean residence times before recorded naturalisation for these distributions are *c*. 73 years (95% C.I. 57–144 years) for major woody weeds and *c*. 243 years (95% C.I. 205–287 years) for naturalised plants whose weed status was considered minor or less. Note the considerable uncertainty around these estimates. The upshot is that for species that will ultimately be insignificant or minor weeds, several centuries of observing a plant as failing to naturalise will be required before it can be declared with any certainty that it will subsequently fail to do so (Figure 2). Major weeds, however, show evidence of naturalisation much faster with *c*. 80% expected to be recorded as naturalising within a century of introduction (Figure 2).

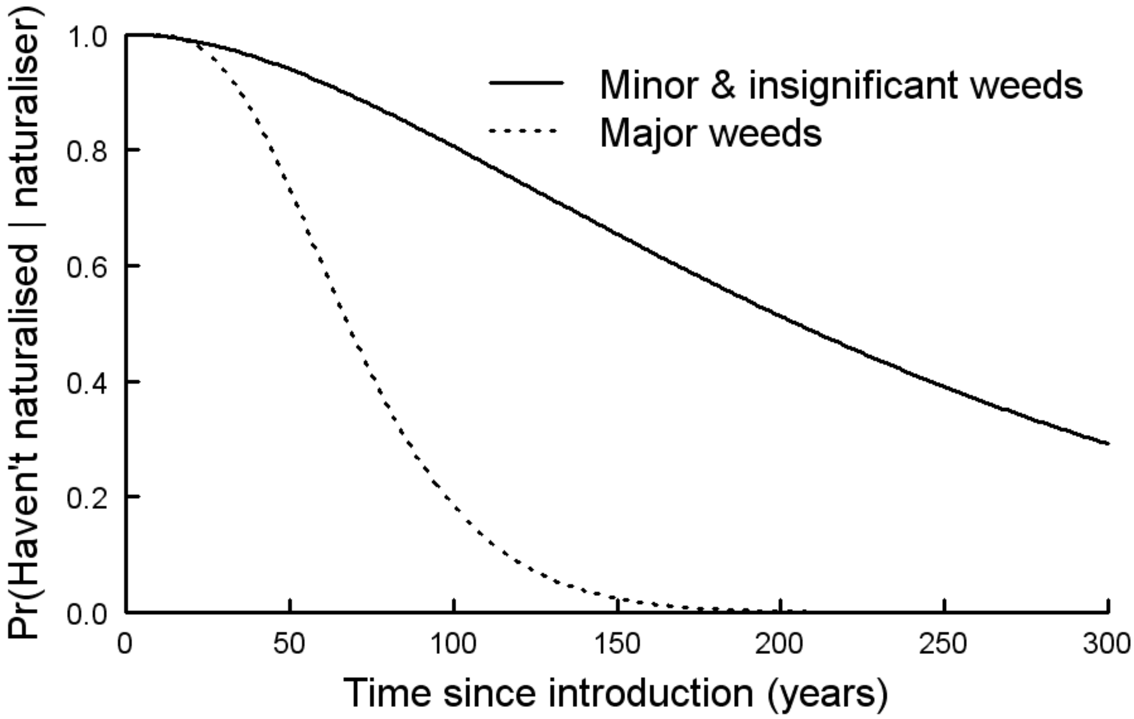

Figure 2. Estimated probability of naturalisation at a future date for woody perennials introduced to South Australia that are either major weeds (dashed line), or insignificant or minor weeds. Data sourced from Caley *et al*. (2008).

5.3 Estimating effects of not knowing invasiveness elsewhere

5.3.1 Directly from WRA system training dataset

If not knowing that a species is invasive elsewhere was truly uninformative, then we would expect the posterior probability of weediness to be unchanged from the prior. That is, the bootstrapped relationships should cluster around the line of equivalence (45º). However, it appears that not knowing whether a plant is a weed elsewhere is indeed informative in that it lowers the probability that a species is a weed (confidence limits based on 1,000 bootstrap samples don't include a 45º line of equivalence) (Figure 3). As might be expected, when the invasiveness of species elsewhere is uncertain, the results lie between the two certain cases (known to be invasive and known to be non-invasive). However, the results are much closer to the case where species are known to be non-invasive elsewhere.

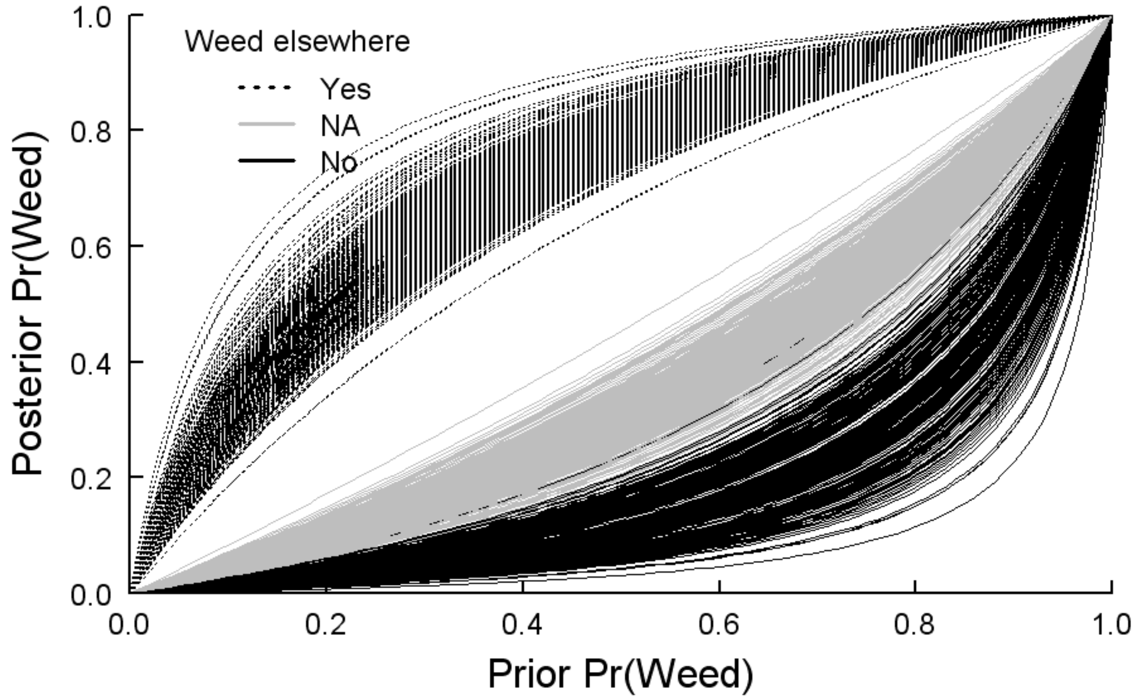

Figure 3. Relationship between the posterior probability of being a weed (either major or minor) in relation to the prior probability of being a weed for different levels of knowledge with respect to the 'Weed elsewhere' question ('Yes'=dotted, 'NA'=grey, 'No'=black). The variation within each level of knowledge arises from a form of simulation known as bootstrapping, whereby the training dataset was resampled with replacement, thus generating a range of possible relationships that could be expected in practice.

5.3.2 Via a statistical model classifier

The random forest constructed using all questions had a sensitivity of *c*. 93% and a specificity of *c*.56% [slightly better than chance]. The estimate of the overall classification error (taken from the "out of the bag" error rate generated during the model fitting process), whilst low (*c*. 16%), is largely meaningless as it applies only to this dataset which contains an inflated proportion of invasive species. In practice, the error rate would be expected to rise with an increasing proportion of non-invasive species being assessed due to the low specificity (Caley *et al.* 2006).

The random forest constructed without questions pertaining to being invasive elsewhere had a sensitivity of *c*. 91% and a specificity of *c*.51% [not really better than chance, and not much less than the model constructed using all questions].

6. Discussion

6.1 Results

The examination of the existing WRA training data suggests that the current approach (default settings assume the worst case) of increasing the estimated risk (via increasing the score) of a species being invasive if 'invasiveness elsewhere' is unknown, is on average slightly conservative, but correctly reflects the increased probability of these species being invasive compared with those not considered to be invasive elsewhere. In agreement with the empirical results, even if it is unknown whether a species has been introduced outside of its native range, a lack of evidence for naturalisation is associated with a reduced risk of the species being invasive. So, the absence of evidence of invasiveness in this case implies evidence of absence of invasiveness — in disagreement with the common maxim. This could be incorporated into assessments as a prior belief, thus avoiding the requirement for documentary evidence. One explanation for this apparent lack of information actually being of assistance in estimating risk is the differential reporting rates between plant species documented as being invasive when introduced outside their native range (e.g. via weed lists) with those subject to recorded introductions outside their native range for which the outcome is uncertain or unknown. A related explanation is that unrecorded introductions of species' outside their native ranges occur frequently.

For woody perennials known to be recently introduced elsewhere, but showing no signs of invasiveness, it will take a long time (centuries) under a wait and see approach before they can be declared as being unable to naturalise with any certainty. The implication is that weed risk assessments on woody perennials with unknown invasiveness elsewhere will have considerable "shelf life"— evidence will accumulate slowly over time that some species are less likely to become invasive than originally thought, while others will demonstrate invasiveness. The drawn out nature of the time to observed naturalisation arises from the many independent and additive waiting processes involved. These include reaching reproductive maturity, undergoing successful pollination, having propagules lodge in a suitable location, being detected and correctly identified, etc. The experimental trial approach (e.g. Pattison and Mack 2009), can fast-track some of these processes, and to some extent mirrors accelerated failure testing used in manufacturing. Such an approach may ultimately be the only method of resolving uncertainty in a species' ability to naturalise in the short term.

It was surprising how well the random forest model was able to classify species as either minor or major weeds with the questions pertaining to invasiveness elsewhere removed, although this sensitivity came at the expense of the classifier's specificity (the ability to correctly identify non-invasive species), which was poor (no better than guessing). The implication of this is that the positive predictive value of such a test when applied to prospective introductions for which the prior probability of being invasive is low will be poor. That is, of those species classified as invasive, only a small proportion will truly be invasive. The only slightly reduced performance of the random forest classifier in the absence of questions pertaining to invasiveness elsewhere suggests that this information may not be as critical as previously reported (e.g. Caley and Kuhnert 2006). The result was unexpected, though as always, a test on truly independent data would be much more robust and preferable. To elaborate, the experts generating the training dataset underlying the current system did so with the knowledge of whether a species being assessed was a known weed within Australia or not. Often more is known of the biology of weeds than non-weeds. Such knowledge could potentially cause a form of recall bias, a well recognised problem in medical case-control studies whereby cases (afflicted individuals) are more likely to recall exposures to postulated explanatory variables than controls (healthy individuals).

Virtue (earlier unpublished report as part of this project) identified the importance of context when undertaking weed risk assessment. Presenting estimates of weed risk as done in Figure 3 (posterior probability on y-axis dependent on chosen prior probability on x-axis) forces individuals to specify the context of the assessment, in terms of the prior probability of invasiveness. It also gives a graphical representation of the uncertainty in the estimated probability arising from variability.

6.2 Future work

There is a need to collect additional data on plant introductions to critique current and alternative weed risk assessment models, as the current training data are limited in extent (mainly for non-invasive plants), and may be in need of updating. This, however, is not as straightforward as initial appearances suggest. The current WRA system is essentially based on a case-control experimental design, whereby cases are known weeds and controls are widely introduced plants that have shown no weed tendencies. The type of data collected needs to suit the design underlying the classification model. For example, a case-control design requires that cases (weeds) and controls (non-weeds) are matched for opportunity to demonstrate invasiveness as a means of avoiding confounding. Plants introduced since Australia has been banning the importation of known weeds, and preventing the introduction of predicted weeds, will clearly be biased against species known to be weeds elsewhere and those predicted to be weeds if introduced to Australia. The suitability of these data for modelling needs to be carefully considered. Species introduced much earlier, such as those identified in nursery stock catalogues (see Mulvaney 1991) may be an avenue to pursue, as they were presumably subject to little or no screening. In the case of woody perennials introduced into temperate climes, the results here estimate that those having no record of invasiveness after 200 years could be considered highly unlikely to naturalise (and subsequently become invasive). An increasing number of plants introduced in the early19th century would be reaching this length of residence time, and could be included in a training data set. As with previous data, there is a risk of experts being biased in their responses to questions due to them knowing the species being assessed is not invasive. The solution would be to design such data collection and analysis in a manner so as to minimise potential biases.

7. References

Breiman, L (2001) Random forests. *Machine Learning,* 45(1): 5–32.

Breiman, L (2003) *Manual for Setting Up, Using, and Understanding Random Forest V4.0*.

Caley, P, RH Groves and R Barker (2008) Estimating the invasion success of introduced plants. *Diversity and Distributions,* 14: 196–203.

Caley, P and PM Kuhnert (2006) Application and evaluation of classification trees for weed risk assessment. *Austral Ecology,* 31: 647–55.

Caley, P, WM Lonsdale and PC Pheloung (2006) Quantifying uncertainty in predictions of invasiveness. *Biological Invasions,* 8: 277–86.

Crooks, JA (2005) Lag times and exotic species: The ecology and management of biological invasions in slow-motion. *Ecoscience,* 12(3): 316–29.

Daehler, CC (2009) Short lag times for invasive tropical plants: evidence from experimental plantings in Hawai'i. *PLoS ONE,* 4(i): e4462.

Gordon, DR, B Mitterdorfer, P Pheloung, S Ansan, C Buddenhagen, C Chimera, CC Daehler, W Dawson, JS Denslow, A LaRosa, T Nishida, DA Onderdonk, FD Panetta, P Pysek, RP Randall, DM Richardson, NJ Tshidada, J Virtue and PA Williams (in press) Guidance for addressing the Australian Weed Risk Assessment questions. *Plant Protection Quarterly*.

Gordon, DR, DA Onderdonk, AM Fox and RK Stocker (2008) Consistent accuracy of the Australian weed risk assessment system across varied geographies. *Diversity and Distributions,* 14(2): 234–42.

Hastie, T, R Tibshirani and J Friedman (2009) *The Elements of Statistical Learning: Data Mining, Inference and Prediction,* New York: Springer.

Hayes, KR and SC Barry (2008) Are there any consistent predictors of invasion success? *Biological Invasions,* 10(4): 483–506.

Kloot, PM (1986) *Checklist of the introduced species naturalised in South Australia*. Technical Paper No. 14. South Australian Department of Agriculture, Adelaide.

Kowarik, I (1995) Time lags in biological invasions with regard to the success and failure of alien species. In: Pysek, P (Ed.) *Plant Invasions: General Aspects and Special Problems.* Amsterdam: Academic Publishing.

Krivanek, M and P Pysek (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.

Lockwood, JL, P Cassey and T Blackburn (2005) The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution,* 20(5): 223–8.

Mack, RN (1996) Predicting the identity and fate of plant invaders: emergent and emerging approaches. *Biological Invasions,* 78: 107–21.

Mulvaney, MJ (1991) *Far from the garden path: an identikit picture of woody ornamental plants invading south-eastern Australian bushland*. Canberra, Australian National University.

Natural Resource Management Ministerial Council (2006) *Australian Weeds Strategy – A national strategy for weed management in Australia*. No. Australian Government Department of the Environment and Water Resources, Canberra, ACT. http://www.weeds.gov.au/publications/strategies/pubs/weed-strategy.pdf.

Nishida, T, N Yamashita, M Asai, S Kurokawa, T Enomoto, PC Pheloung and RH Groves (2009) Developing a pre-entry weed risk assessment system for use in Japan. *Biological Invasions,* 11(6): 1319–33.

Pattison, RR and RN Mack (2009) Environmental constraints on the invasion of *Triadica sebifera* in the eastern United States: an experimental field assessment. *Oecologia,* 158(4): 591–602.

Pheloung, PC, PA Williams and SR Halloy (1999) A weed risk assessment model for use as a biosecurity tool evaluating plant introductions. *Journal of Environmental Management,* 57: 239–51.

Pysek, P, DM Richardson, M Rejmanek, CL Webster, M Williamson and J Kirschner (2004) Alien plants in checklists and floras: towards better communication between taxonomists and ecologists. *Taxon,* 53: 131–43.

R Development Core Team (2008) *R: A language and environment for statistical computing*. Vienna, Austria, R Foundation for Statistical Computing.

Walton, C, N Ellis and P Pheloung (1998) *A manual for using the Weed Risk Assessment system (WRA) to assess new plants*. No. Australian Quarantine and Inspection Service, Canberra.

Williamson, M and A Fitter (1996) The varying success of invaders. *Ecology,* 77: 1661–5.

Wilson, JRU, DM Richardson, M Rouget, S Proches, MA Amis, L Henderson and W Thuiller (2007) Residence time and potential range: crucial considerations in modelling plant invasions. *Diversity and Distributions,* 13: 11–22.