PRATIQUE

ENHANCEMENTS OF PEST RISK ANALYSIS TECHNIQUES

Best practice for quantifying uncertainty and summarising and communicating risk

D 3.2 Protocol for quantifying and communicating uncertainty in the PRA scheme accessed via a hyperlink in a project web page and integrated into the web-based EPPO DSS for PRA

D 3.4 Protocol for summarising and communicating overall risk in the PRA scheme accessed via a hyperlink in a project web page and integrated into the web-based EPPO PRA scheme

Author(s): Gritta Schrader, Richard Baker, Damien Griessinger, Andy Hart, Wil Hennen, Johnson Holt, Jon Knight, Adrian Leach, Alan MacLeod, David Makowski, John Mumford, Françoise Petter, Willem Roelofs
(see also acknowledgements)

Partner(s): WU, Imperial, EPPO, INRA, Fera, JKI

Submission date: 19 July 2011

EU Framework 7 Research Project
Enhancements of Pest Risk
Analysis Techniques

(Grant Agreement No. 212459)

SEVENTH FRAMEWORK PROGRAMME

PROJECT OVERVIEW: PRATIQUE is an EC-funded 7th Framework research project designed to address the major challenges for pest risk analysis (PRA) in Europe. It has three principal objectives: (i) to assemble the datasets required to construct PRAs valid for the whole of the EU, (ii) to conduct multi-disciplinary research that enhances the techniques used in PRA and (iii) to provide a decision support scheme for PRA that is efficient and user-friendly. For further information please visit the project website or e-mail the project office using the details provided below:

Email: pratique@csl.gov.uk
Internet: www.pratiqueproject.eu

Authors of this report and contact details

Name: Gritta Schrader
Partner: JKI
E-mail: Gritta.Schrader@jki.bund.de

Name: Richard Baker
Partner: Fera
E-mail: richard.baker@fera.gsi.gov.uk

Name: Damien Griessinger
Partner: EPPO
E-mail: dg@eppo.fr

Name: Andy Hart
Partner: Fera
E-mail: andy.hart@fera.gsi.gov.uk

Name: Johnson Holt
Partner: Imperial
E-mail: : j.holt@imperial.ac.uk

Name: Wil Hennen
Partner: WU
E-mail: Wil.Hennen@wur.nl

Name: Jon Knight
Partner: Imperial
E-mail: j.d.knight@imperial.ac.uk

Name: Adrian Leach
Partner: Imperial
E-mail: : a.w.leach@imperial.ac.uk

Name: Alan Mac Leod
Partner: Fera
E-mail: alan.macleod@fera.gsi.gov.uk

Name: David Makowski
Partner: INRA
E-mail: makowski@grignon.inra.fr

Name: John Mumford
Partner: Imperial
E-mail: j.mumford@imperial.ac.uk

Name: Françoise Petter

Partner: EPPO
E-mail: fp@eppo.fr

Name: Willem Roelofs
Partner: Fera
E-mail: willem.roelofs@fera.gsi.gov.uk

Disclaimer:

This publication has been funded under the small collaborative project PRATIQUE, an EU 7th Framework Programme for Research, Technological Development and Demonstration addressing theme: [kbbe-2007-1-2-03: development of more efficient risk analysis techniques for pests and pathogens of phytosanitary concern call: fp7- kbbe-2007-1].

Best practice for quantifying uncertainty and summarising and communicating risk

Table of Contents

Background to this deliverable and introduction

This paper represents the Deliverables D 3.2 “Protocol for quantifying and communicating uncertainty in the PRA scheme accessed via a hyperlink in a project web page and integrated into the web-based EPPO PRA scheme” and D 3.4 “Protocol for summarising and communicating overall risk in the PRA scheme accessed via a hyperlink in a project web page and integrated into the web-based EPPO PRA scheme”. As described in the PRATIQUE 18 month report, “these deliverables have been merged since the tasks are so closely related, operate on the same time scale and it is much more efficient to manage them together”. Reflecting this, the milestones (M3.3 and M3.7) related to these two tasks have already been merged. The methods to summarise risk and to quantify and communicate uncertainty have been combined in the models. The risk rating and the level of uncertainty for each question are captured at the same time and propagated through the EPPO pest risk analysis (PRA) to deliver an overall risk level which is already integrated with the degree of uncertainty. This is done with the help of CAPRA, the electronic version of the EPPO Decision Support Scheme for PRA (EPPO DSS for PRA; EPPO, 2009 and revised version) created by

PRATIQUE.

The objective of this merged deliverable is to provide methods that can be used to combine risk ratings to produce an overall summary of pest risks and to capture uncertainty a) in the responses to every question in the pest risk assessment, b) when summarising each section and c) for the overall risk. For this:

- “innovative approaches will be used to structure and weight the combination of scores from different parts of the assessment in ways that reflect the real mechanisms that determine risk. Probabilistic techniques, e.g. Bayesian Belief Networks (BBN), will be developed and tested together with formal methods to elicit structure and weighting information, e.g. conditional probabilities, from PRA experts in all parts of the project. Innovative protocols will be developed for combining scores that are simple and practical for routine use but which reflect better the real mechanisms determining risk and therefore provide more meaningful information for decision-makers.” (Subtask 3.4.2)
- “A sensitivity analysis for each case study will be conducted to identify the parts of the pest risk assessment that contribute most to overall uncertainty.” (Subtask 3.2.2)

(Text in quotation marks cited from the Description of Work for Work Package 3).

Following a detailed review and evaluation of the different methods available (see Annex 2, Milestone 3.3/3.7) for these purposes, the most promising techniques were developed and tested at the PRATIQUE/EPPO workshop in Hammamet, Tunisia, in November 2010, in the EPPO Panel on PRA Development (December 2009, May 2010, January 2011), in several EPPO Expert Working Groups on PRA for certain organisms (*Drosophila suzukii, Agrilus anxius*), and in the EPPO Panel on Phytosanitary Measures in February 2010. The most useful methods have been selected for incorporation into the EPPO DSS for PRA.

The EPPO DSS for PRA uses a risk rating system, like many other risk assessment schemes worldwide (see PRATIQUE deliverable D1.1 collection of schemes for pest risk analysis), which enables experts to provide qualitative assessments for key topics arranged as a logical series of questions with an aggregation step at the end where the scores are summarized to provide an overall judgement of risk. The aggregation stage for the EPPO DSS for PRA has so far not been standardized and no method for summarizing and combining risk ratings is provided in previous versions of the scheme (EPPO, 2009) other than asking for an overall expert judgement (see below). However, since a good method for aggregating risk ratings and uncertainties is essential to guarantee that the overall judgement matches the expected risk as closely as possible, is reproducible, transparent and easy to use, an important goal of PRATIQUE was to improve the EPPO DSS for PRA in this respect.

This deliverable consists of 3 sections and three annexes.

The first section explains the concept of uncertainty because of the many different definitions and categorizations.

In section 2 brief descriptions are provided for the five methods that were selected and tested for their usefulness to rate, summarize and visualise the factors affecting pest risk and to assess uncertainty. The five methods are: Bayesian Belief Networks (BBNs), the Knowledge-based Approach (KBA), the PRA Risk Rating and Uncertainty Visualiser, the Risk Matrix Models (MMs) and the Invasive Risk Impact Simulator (IRIS).

Section 3 discusses the different methods described in section 2, summarizes the results of the testing phase, considers their advantages and disadvantages and presents the methods that were finally selected.

In Annex 1, the methods that were selected (PRA Risk Rating and Uncertainty Visualiser, MMs and IRIS) are described in detail. This corresponds to the "protocol for quantifying and communicating uncertainty and for summarising and communicating overall risk in the PRA scheme accessed via a hyperlink in a project web page" (a combination of the two titles of deliverables D 3.2 and D 3.4.

In Annex 2, the Milestone 3.3/3.7 can be found, that was included in the 18 month report.

In Annex 3, an example is given, showing the integration of new methods into CAPRA using *Drosophila suzukii* as a test organism.

This deliverable used the EPPO DSS for PRA (EPPO, 2009) as a starting point for the evaluation of methods for pest risk assessments. Changes to the scheme were implemented in a revised written version of the EPPO DSS for PRA and in CAPRA.

In the EPPO DSS for PRA it is important to acknowledge that some questions are more important than others (MacLeod and Baker, 2003). In addition, there are questions whose answers may influence the relevance or weight of other questions, or even have a significant influence on the overall outcome, regardless of other questions. For example, question 1.11 (of the EPPO DSS for PRA from 2009) asks whether consignments arrive at a suitable time of year for pest arrival. Clearly, the answer to this question will determine the relevance of all other questions. Therefore, when alternative methods were evaluated and developed, knowledge about the differences in the weight of questions, dependencies, and overall influence was taken into account.

Section 1: The concept of uncertainty

Uncertainty is inherent to PRA. Section 3.1 of the (revised) International Standard on Phytosanitary Measures (ISPM) No. 2 (FAO, 2007) contains a short paragraph on uncertainty, and states that "*Uncertainty is a component of risk and therefore important to recognize and document when performing PRAs. Sources of uncertainty with a particular PRA may include: missing, incomplete, inconsistent or conflicting data; natural variability of biological systems; subjectiveness of analysis; and sampling randomness. Symptoms of uncertain causes and origin and asymptomatic carriers of pests may pose particular challenges. The nature and degree of uncertainty in the analysis should be documented and communicated, and the use of expert judgement indicated. If adding or strengthening of phytosanitary measures are recommended to compensate for uncertainty, this should be recorded. Documentation of uncertainty contributes to transparency and may also be used for identifying research needs or priorities.*"

This text, which was written after the publication of ISPM11 (FAO, 2004), provides a clear definition of uncertainty that was missing from ISPM No. 11, where the topic is described vaguely and does not provide any guidance on how to capture uncertainty, stating in section 2.4 that "*the estimation of the probability of introduction and its economic consequences involves many uncertainties*" and that "*the assessment of the probability and consequences of environmental hazards of pests of uncultivated and unmanaged plants often involves greater uncertainty than for pests of cultivated or managed plants. This is due to the lack of information, additional complexity associated with ecosystems, and variability associated with pests, hosts or habitats*". In the same paragraph, the importance of uncertainty is underlined, without recommending methods for assessing or incorporating it into a PRA. However, it is stated that uncertainty should be captured where possible: "*It is important to document the areas of uncertainty and the degree of uncertainty in the assessment and to indicate where expert judgement has been used*".

The current EPPO DSS for PRA (EPPO, 2009) already requires uncertainty to be considered by asking the assessor to add a rating of the level of uncertainty (low, medium, or high) when answering each question. In addition, at the end of the assessment, an overall summary of uncertainty is required: "*Estimation of the probability of introduction of a pest and of its economic consequences involves many uncertainties. In particular, this estimation is an extrapolation from the situation where the pest occurs to the hypothetical situation in the PRA area. It is important to document the areas of uncertainty (including identifying and prioritizing of additional data to be collected and research to be conducted) and the degree of uncertainty in the assessment, and to indicate where expert judgement has been used. This is necessary for transparency and may also be useful for identifying and prioritizing research needs.*" However, no guidance is provided to indicate what is meant by low, medium, and high uncertainty (see Deliverable D3.1, the revised EPPO DSS for PRA and CAPRA for improvement of this situation) and how to combine these ratings to achieve a consistent overall rating of uncertainty.

Types of uncertainty

There are many different definitions and categorizations of uncertainty. Uncertainty can be considered to be a "*general concept that reflects our lack of sureness about something or someone, ranging from just short of complete sureness to an almost complete lack of conviction about an outcome*" (NRC, 2000). Uncertainty is thus a measure of the limitations in our knowledge about factors that influence risk.

Different taxonomies of uncertainty have been distinguished (Morgan & Henrion 1990; Cullen & Frey 1999; Morris and Sayers, 2002) but there are two basic kinds of uncertainty: variability and ignorance.

Natural variability - refers to the randomness observed in nature (standard term to be used) and is also referred to as:

- Aleatory uncertainty (meaning to 'gamble')
- External uncertainty
- Inherent uncertainty
- Objective uncertainty
- Random uncertainty
- Stochastic uncertainty
- Irreducible uncertainty
- Fundamental uncertainty
- Real world uncertainty

Ignorance refers to the state of knowledge of a physical system and our ability to measure and model it (standard term to be used) and is also referred to as:

- Epistemic uncertainty (meaning 'knowledge')
- Functional uncertainty
- Internal uncertainty
- Subjective uncertainty
- Incompleteness

There are fundamental differences between these uncertainties: Ignorance can, at least in theory, be reduced by further study whereas variability is an inherent characteristic of the system that is being studied. One can obtain a better estimate of variability but its magnitude cannot be reduced. Another difference is that, unlike variability, ignorance is subjective: it depends on what knowledge is available and how that knowledge is obtained. The differentiation between variability and ignorance is not always as clear cut as we think: what is variability in one problem can be uncertainty in another. In many traditional risk assessments, no distinction has been made between the two types.

Table 1 gives a useful typology of uncertainty from climate change science and describes the typical approaches used to address each type (IPCC, 2005). Both unpredictability and structural uncertainty issues form a significant part of PRA though most of the uncertainty involved in the PRA process is derived from value uncertainty. This category is addressed using observation data and various quantitative methods.

Table 1. A simple typology of uncertainties (from IPCC, 2005)

Type	**Indicative examples of sources**	**Typical approaches or considerations**
Unpredictability	Projections of human behaviour not easily amenable to prediction (e.g. evolution of political systems). Chaotic components of complex systems.	Use of scenarios spanning a plausible range, clearly stating assumptions, limits considered, and subjective judgments. Ranges from ensembles of model runs.
Structural uncertainty	Inadequate models, incomplete or competing conceptual frameworks, lack of agreement on model structure, ambiguous system boundaries or definitions, significant processes or relationships wrongly specified or not considered.	Specify assumptions and system definitions clearly, compare models with observations for a range of conditions, assess maturity of the underlying science and degree to which understanding is based on fundamental concepts tested in other areas.
Value uncertainty	Missing, inaccurate or non-representative data, inappropriate spatial or temporal resolution, poorly known or changing model parameters.	Analysis of statistical properties of sets of values (observations, model ensemble results, etc); bootstrap and hierarchical statistical tests; comparison of models with observations.

According to Van Leeuwen and Vermeire (2007), risk assessment in practice is hampered by four types of uncertainty:

- Lack of data and information
- Measurement uncertainties
- Observation conditions, e.g. due to climate or soil type
- Inadequacies of models

Other sources of uncertainty include:

- Uncertainty about distribution shape – often several different distributions may be plausible for the same model input and may show similar goodness of fit to sample data.
- Sampling uncertainty – when a sample is used to estimate distribution parameters or to derive an empirical distribution, there is uncertainty about their relationship to the true parameters or distribution for the larger population from which the sample was drawn.
- Extrapolation uncertainty – when it is necessary to extrapolate beyond the range of a dataset, or from one type of data to another (surrogacy), there is uncertainty about how closely the extrapolated values match the true values that are being estimated. It should be noted that this is one of the major sources of uncertainty in PRA as an assessor will use information from the area where the pest is present and extrapolate it to the PRA area.

- Model uncertainty – there is often uncertainty about which of several alternative model structures best represents real mechanisms and processes.
- Uncertain dependencies – there may be uncertainty about the nature, direction and magnitude of dependencies in the assessment.

Clearly, these are all uncertainties of the second kind (i.e. ignorance). In the current EPPO DSS for PRA, this kind of uncertainty has been addressed by explicit qualitative statements about the level of uncertainty associated with the answer to each question. However, as already mentioned above, no guidance was given so far on how these statements on uncertainty should be propagated through the analysis. Many questions are specifically designed to deal with temporal and spatial variability. Despite the fact that many questions result in statements about variability, it was unclear how these statements are combined into the final pest risk assessment. With this deliverable, methods are provided to address, document and take proper account of uncertainty in pest risk assessment.

Section 2: Five approaches to summarise risk and handle uncertainty

Twenty eight different methods for summarizing risk and the handling of uncertainty were assessed in detail. A full analysis of these methods can be found in Annex 2. From this analysis, two methods were taken forward for further study and development (though one of them includes several elements/approaches from other methods). In addition to these two methods, based on the experience gained in early work on this deliverable, three other approaches were further developed, in collaboration with the European Food Safety Authority (EFSA) funded project ("Prima phacie"[1], see below) and the feedback of PRA experts and practitioners. All five methods are described below.

Short descriptions of methods for summarizing risk and for quantifying and communicating uncertainty

For these descriptions, reports of the EPPO Panel on PRA Development in 2009, 2010, and 2011, the workshop in Hammamet, Tunisia, in November 2010, the Milestone 3.3/3.7 (Annex 2), and related documents were used.

[1] In late 2009, an EFSA-funded project (Prima phacie: Pest risk assessment for the European Community plant health: A comparative approach with case studies; MacLeod et al. 2010) began work to review and test methodologies for conducting pest risk assessment by means of case studies on different pests. The project is being conducted by an international consortium of 11 partners from which many are also partners in the PRATIQUE project. Consequently synergies have been established between the two projects resulting in three other methods being developed.

Bayesian Belief Networks (BBNs)

Bayesian (Belief) Networks (BBNs; Pearl, 1985; Jensen, 2001) are graphical models that use nodes to represent variables and arcs to demonstrate dependencies between the linked variables. Often the random variables are discretised into states or categories. For discrete variables, relations between a node and its parents are specified by conditional probability tables. Bayesian networks allow users to calculate the conditional probabilities of the nodes in the network given that the values of some of the other nodes have been estimated.

Prior probabilities are assigned to nodes without parents (for example, any of nodes in the top line of Figure 1) to specify the prior knowledge. When new evidence is entered, it is propagated throughout the network and its implications are calculated. Bayesian Networks can be constructed from expert knowledge where both the network structure and the conditional probability tables (CPTs) of the network are based on the knowledge of a domain expert. Each time a child node (for example, the Total Volume node in Figure 1) is identified, conditional probability tables need to be defined. Alternatively, the network characteristics can be acquired from data.

The ability of BBNs to deal with continuous data is limited. Continuous variables will either need to be made discrete or transformed into normal distributions to make sure that the solutions remain analytically tractable. BBNs represent one branch of Bayesian modelling, the other major approach being hierarchical simulation-based modelling.

BBNs can be used to improve consistency, to combine questions, and to deal with uncertainties. A BBN defines various events, the dependencies between them, and the conditional probabilities involved in those dependencies.

Fig. 1 shows an example of a BBN. This example is a representation of a part of the entry section, where it was considered that the total volume of pest arriving is dependent on the concentration of the pest on the pathway at origin, the volume of the commodity, the frequency, the survival during transport, and the multiplication during transport. Each box represents a node (parent nodes or child nodes); causal relations between the nodes are represented by arrows, the direction of the arrow indicating the direction of the causality. Causalities are determined by experts in conditional probability tables (see below). In the EPPO DSS for PRA parent nodes are mainly represented by the individual questions in the scheme. Child nodes correspond to the clusters (grouping of questions) or to summary conclusions for entry, establishment, spread and impact and, if appropriate, the overall conclusion. In the example, a three level rating (instead of the five level of the current scheme) was chosen. As shown in the example, in a BBN the assessor is asked to distribute probabilities for the different possible scores instead of giving a score and a level of uncertainty, e.g. in the example given for volume of trade (1.6) 70% is attributed to low, 25% for medium and 5% for high. Given the answers for the parent nodes and based on the conditional probability tables, the model calculates a distribution of probabilities for the different rating levels in the child node. Similarly, if the answer is known for a child node, the answers for the parent node can be derived back.

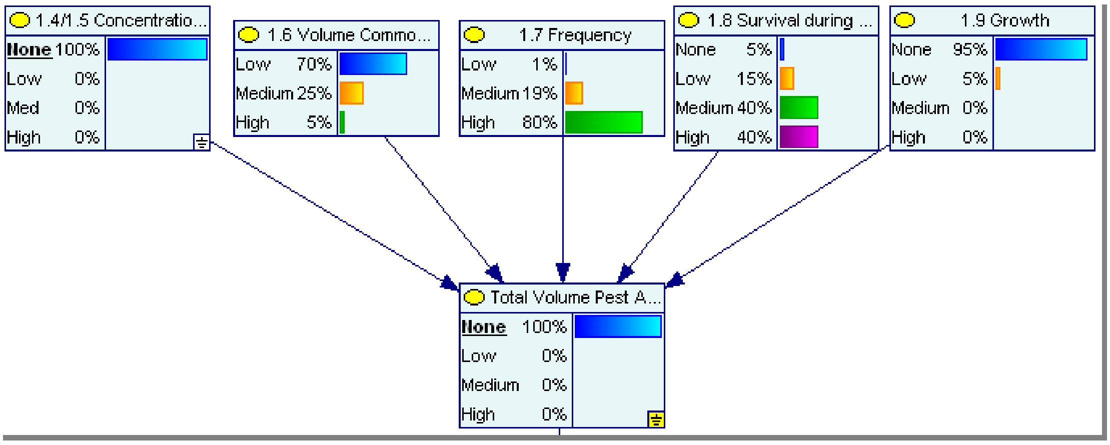

Fig. 1. Graphical representation of a BBN

A BBN can only be defined when a model structure is agreed. The model shown in Fig. 2 follows the current (2009) structure of the EPPO DSS for PRA (i.e. spread is part of the first section).

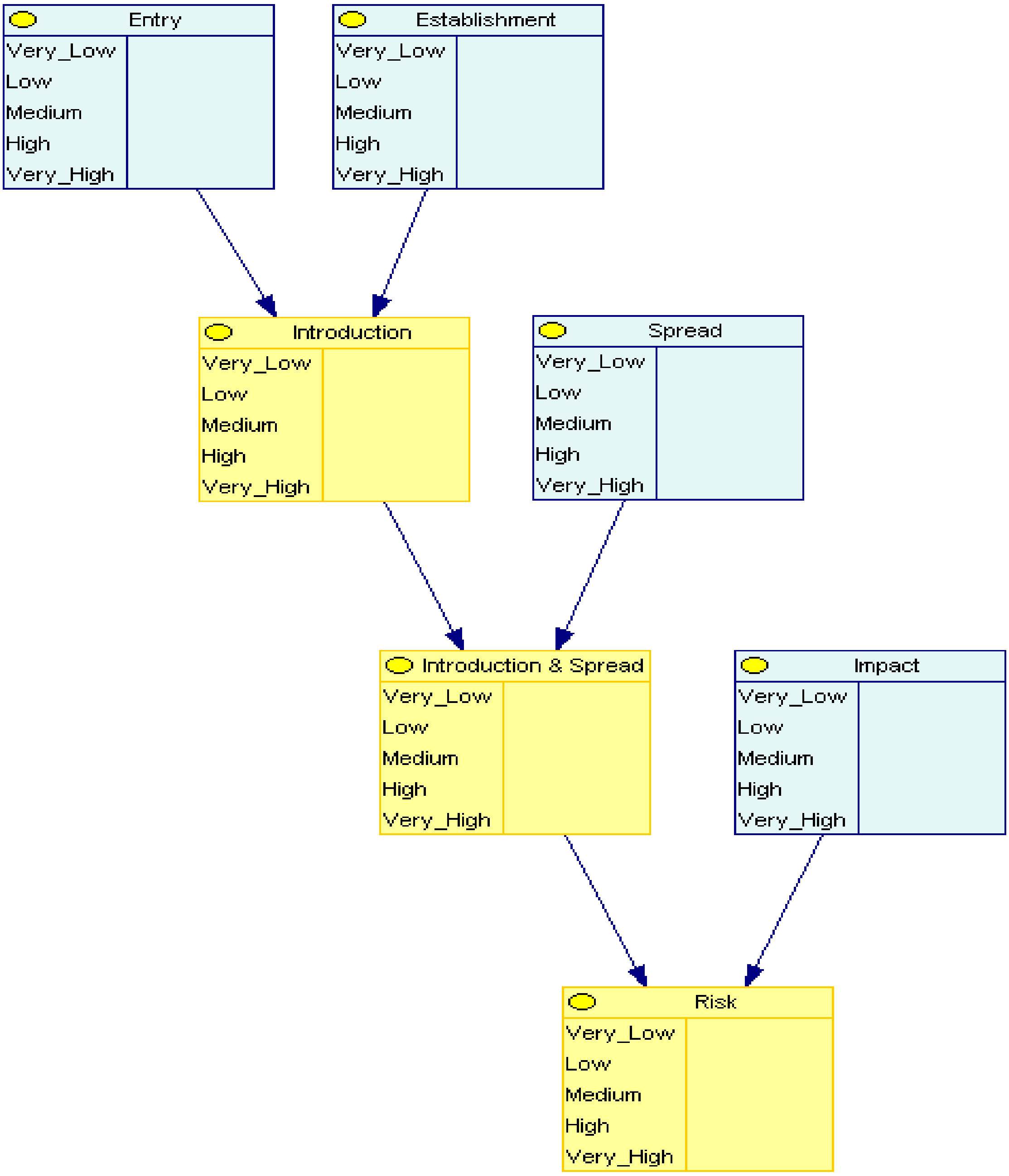

Fig. 2. Model structure for a BBN at the highest level of the EPPO DSS for PRA.

Conditional Probability Tables (CPTs) are used to assign probability values to a combination of scores. An example of a CPT is provided here, see Table 2 (with 3 score levels). The advantage of CPTs in comparison with, for example risk matrices, is that they allow for uncertainty about the combination of scores (combining medium and low could lead to low or medium risk), overcoming the issues identified by Cox (2008). BBNs therefore not only allow for uncertainty in the score assignment but also propagate this uncertainty together with uncertainties about score combinations.

Table 2. Conditional Probability table for entry and establishment.

Entry	Low			Medium			High		
Establishment	Low	Medium	High	Low	Medium	High	Low	Medium	High
Low Risk	90%	75%	20%	75%	10%	0%	10%	0%	0%
Medium Risk	10%	25%	60%	25%	80%	20%	60%	20%	0%
High Risk	0%	0%	20%	0%	10%	80%	30%	80%	100%

The CPT from Table 2 can be visualised as an uncertainty risk matrix shown in Fig. 3.
Green (top left) combines low probability of entry with low probability of establishment. It then shifts to yellow (medium) and red (very high) at the bottom right where very high probability of entry is combined with very high probability of establishment. Uncertainty is visualised by the uniformity of colours: the more the field is uniformly coloured (e.g. bottom right completely uniformly coloured) the lower the uncertainty. It is also useful to visualize inconsistencies (discontinuity of colour gradient).

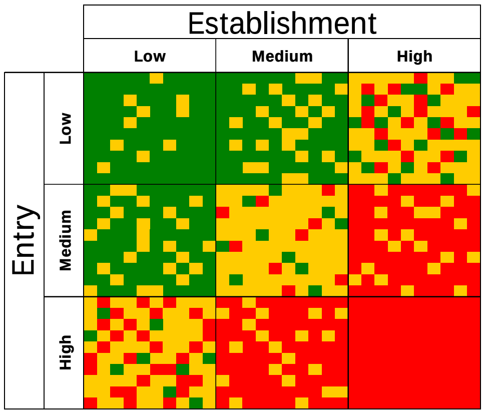

Fig. 3 uncertainty risk matrix

The knowledge based approach (KBA)

The knowledge-based approach (KBA) combines several of the approaches mentioned in Annex 2 (e.g., categories as scalars on a linear scale (2), rating transformations including sums and products (13-19), threshold approaches (20-22), etc). A KBA was based on the current (2009) qualitative EPPO DSS for PRA (Fig. 4) although it was found that ideally the number of categories per question should be increased. In the KBA, the qualitative rating values are transformed to quantitative ratings, not by the assessor but by means of rating-transformation algorithms.

The KBA is intended to mimic the way an expert would combine ratings from different questions to obtain a rating summary for each component of the risk assessment and for the risk assessment as a whole. The ways of combining ratings are translated into algorithms that calculate a final score that is correlated to a level of risk. The approach is based on expert knowledge – thus its name.

During the review process of the different methods for combining and summarizing risks, it was considered that this method could be used to summarize the question ratings in the EPPO DSS for PRA provided that algorithms can be determined that mimic the way an expert combines and summarizes answers to the questions. The model was also enhanced with a

Monte Carlo simulation[2] to take rating uncertainty into account. Since KBA models are usually based on the knowledge of one expert that is considered to be a "top" expert, KBA should not be developed by too many experts.

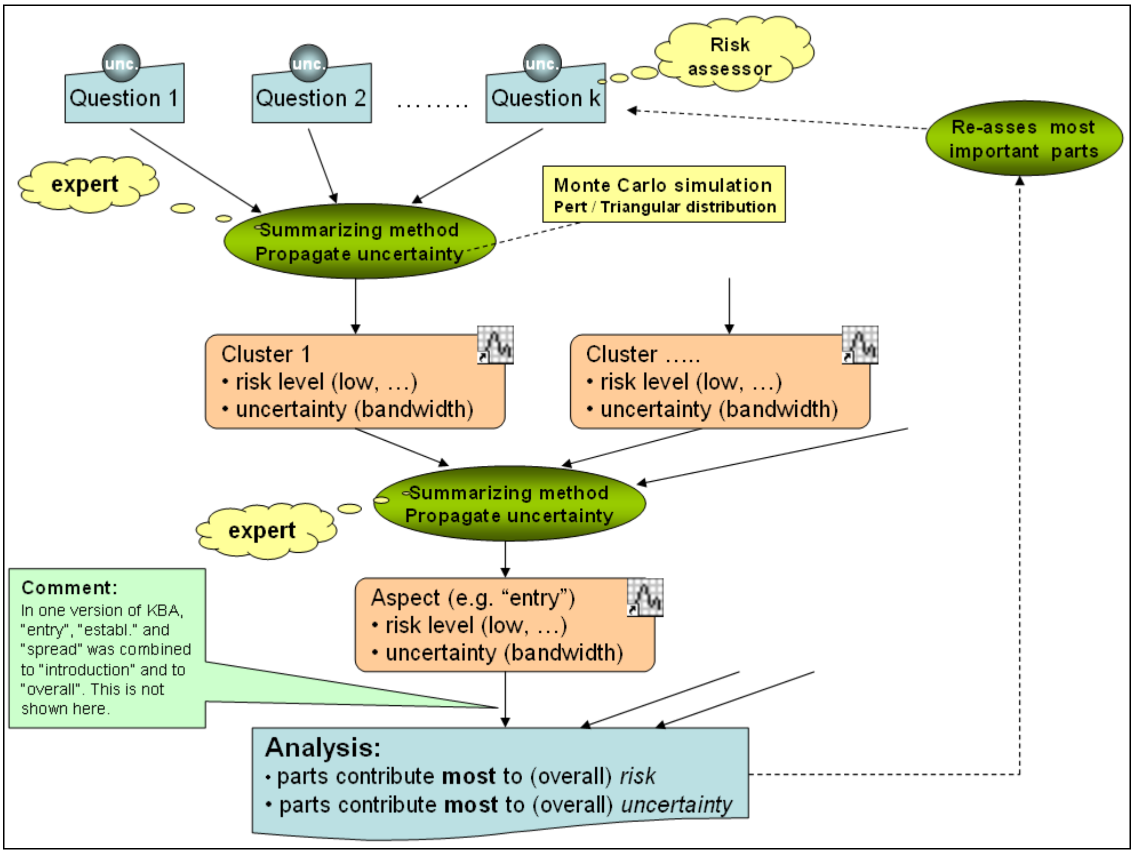

Fig. 4. General outline of the KBA approach.

A knowledge component is regarded as a 'piece' of expert knowledge that can be used to summarize the scores. Knowledge components in the EPPO DSS for PRA can be:

- If…then rules or heuristics (i.e. experience-based techniques)
- The weight of an individual question, a cluster or a risk factor (some questions are more important than others)
- The influence of a rating of an item (a question or a cluster) on the weight of another item
- The dominance of items (e.g. minimum or maximum algorithm): if an answer to a question is very likely then it will have a preponderant importance in the global score.

[2] In Monte Carlo simulation, values are generated randomly from all uncertain variables. Without simulation or use of random numbers, calculations should always result in the same outcome. For each uncertain variable, a probability distribution is defined from from which a possible value is generated at random. After selection of all the input values, a calculation scheme is used to calculate one or more output values based on these selected input values. This process is repeated a number of times (e.g. 1000 trials). The result of all trials can be displayed as a frequency chart (histogram). The distribution gives information about the mean, the most-likely result and minimum and maximum values for these trials. For more information, see e.g. Poulter (1998), Vose (2000) and Hardaker et al. (2002).

- The bonus of an item for another item (some questions will have an answer that will add to the risk but, if a low rating is given, this should not reduce the global risk).

This knowledge is transformed into a mathematical formula.

In order to develop the first prototype, the knowledge components were determined based on the understanding of the sections on entry, establishment and spread of the EPPO PRA DSS by two experts, Wiebe Lammers (Dutch Plant Protection Service) and Françoise Petter (EPPO), in December 2008. For each section (entry, establishment and spread) clusters (i.e. groupings) of questions were identified. It was considered that the clusters should correspond to the subsections of the EPPO decision-support scheme, e.g. for the entry part:

- Cluster 1: Probability of the pest being associated with the individual pathway at origin,
- Cluster 2: Probability of survival during transport or storage
- Cluster 3: Probability of the pest surviving existing pest management procedures
- Cluster 4: Probability of transfer to a suitable host or habitat.

The clustering was applied to the two other sections in the same way - following the structure of ISPM No. 11, FAO, 2004).

It needs to be noted that a cluster-outcome is not a probability. KBA tries to mimic the assessor's combination of questions and clusters in a cognitive way (e.g. heuristics) and not in a probabilistic way.

In almost all questions of the EPPO DSS for PRA, the rating is based on the choice between five risk levels where the highest level always corresponds to the highest risk. For the KBA, the middle point (0.5) had to be fixed; the experts were consequently asked where in the scale they would be indifferent. The choice was made by experts that the indifference point would be at the third level. Consequently, the following numerical translations from qualitative ratings to quantitative ratings were proposed:

- 1st level (e.g. very unlikely): 0.05;
- 2nd level (e.g. unlikely): 0.25
- 3rd level (e.g. moderately likely): 0.5;
- 4th level (e.g. likely): 0.75;
- 5th level (e.g. very likely): 0.95.

In addition, the level of uncertainty for each question should be given (low, medium or high) and, depending on this level, the distribution of the answers around the mean point is calculated (using a triangular or Pert distribution). The calculation of the final outcome is made by running 250 iterations (Monte Carlo simulation). The user of the model can change this number of iterations. The model gives as the outcome an entry/ establishment/spread risk mean percentage with a band width of value depending on the uncertainty (e.g. entry risk 70% (57-77%)). The model allows the questions for which uncertainty has a major influence on the global outcome to be identified as well as the questions that contribute most to the risk. This can be used by the assessors to identify the questions for which searching for additional data to reduce uncertainty would be most effective.

After testing and applying the first prototype, the model and in particular the knowledge underlying the mathematical algorithms was further developed and a revised version was prepared.

PRA Risk Rating and Uncertainty Visualiser

The PRA Risk Rating and Uncertainty Visualiser is a tool that graphically displays risk and uncertainty scores according to the position and size of bubbles for each question in each section (see Fig. 5). It was developed within the PRATIQUE Project (Mumford *et al.* 2010) in an Excel format and was integrated in CAPRA in order to be independent from Excel. It is simply a tool that displays what the expert has answered - no calculation is involved. This method was developed following the realization that assessors and panels need to visualize and summarize their responses. It provides a visual summary for each section (entry, establishment, spread and impact). This tool is designed to help assessors validate the risk ratings for each section of the PRA and to provide a visual feedback to assessors to improve consistency. For example, in the Entry section, if all individual responses are rated as low or medium but the overall Entry response is rated as High or Very High then there is a lack of consistency that requires explanation or reconsideration. Another advantage is to allow assessors to visualize implicit weighting. For example, using the visualiser in the Entry section, an assessor can easily see which questions have been rated highest for each species/pathway combination and hence can identify where management actions may play the greatest overall role in reducing the likelihood of entry. A grey bar represents the overall rating for the section being visualised; the width of this bar indicates the uncertainty associated with the overall rating. It does not represent a mean score or other type of average score of individual questions. The method is described in detail in Annex 1.

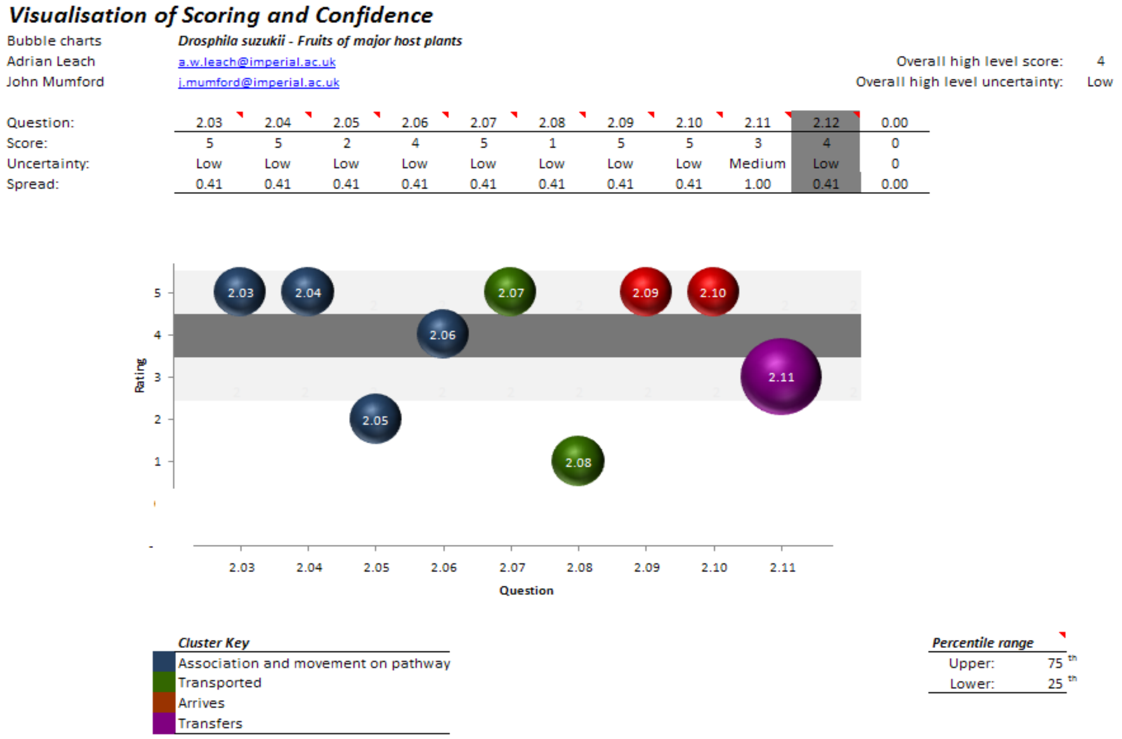

Fig. 5. PRA Risk Rating and Uncertainty Visualiser for the Entry section.

Rule-based Matrix Models (MMs)

The MM approach is another method that consistently combines risk elements (rating and uncertainty) to summarize the overall risk for each section of the scheme. The decision to develop the rule-based matrix models (MMs) emerged after ontologies (an example is given in Fig. 6) were prepared to visually describe the different sections of the EPPO DSS for PRA (see also Deliverable D3.1). The logical sequence of events leading to entry, establishment, spread, and impact was represented graphically in a flow diagram. Originally, it was only intended to present the visualiser and the ontologies for the different sections as outcomes of PRATIQUE. However, the ontologies were an essential precursor to the MM approach and triggered their development by the Prima Phacie project. It was then further elaborated and adapted to the needs of PRATIQUE. The ontologies enable experts to visualize the logical sequence and relationship of the different links between questions and provide a logical hierarchical structure to the scheme. Once these ontologies had been agreed, the MMs were designed. MMs summarize a set of rules (e.g. if the answer to one question is very unlikely and to another question is likely, then the result is, e.g. moderately likely, etc.). By providing fixed rules, they provide a practical solution to the complexity that would result when trying to fill in BBN CPTs for multiple combinations of questions and 5 levels of risk ratings. Uncertainty is captured by the assessor by distributing the rating across the different rating classes according to the degree of uncertainty (see below). By including beta distributions the models presented are semi-quantitative and they allow the aggregation of individual components of risk. Different models have been developed for entry, establishment, spread and impact (both environmental and other economic impacts). The MMs were developed based on “ontologies” for each section of the EPPO DSS for PRA (see Fig. 6 as an example for the establishment part), where questions from the scheme were arranged at the base of the hierarchy and are linked via intermediate nodes to provide an overall assessment for entry, establishment , etc. This hierarchical structure is similar to a BBN but the nodes are combined by rules rather than by calculating a probability distribution. Only two nodes can be combined at a time and no uncertainty exists about the combination of nodes although beta distributions are applied to represent uncertainty ratings associated with ratings to questions. The MMs were initially based on the EPPO DSS for PRA (EPPO, 2009) adapted by the EFSA Panel on Plant Health (Appendix C of EFSA, 2010) that was used in Prima phacie, and then adapted to the revised scheme. Further development of the MMs has been undertaken by Prima phacie and PRATIQUE to reflect the different approaches taken at question and sub-question level. MMs take the responses to the individual questions or sub-questions (the ratings being transformed to a 1 to 5 scale point) and the level of uncertainty to generate a combined risk representation. MMs have also been used to combine PRA section summaries, e.g. Entry with Establishment. The combined assessment of entry and establishment will not be combined with the assessment of magnitude of the risk; these two components of risk will therefore be kept separate.

Depending on the questions and their relationships, different types of matrix combinations can be used to combine the ratings from different questions. For example, the sphere “Climate & abiotic” results from the combination of questions 3.11 and 3.12 using a minimum matrix. Minimum matrices may be appropriate when the response to either question acts as a powerful constraint to the final result. With this method, the choice of matrix allows different types of interaction between risk components to be represented, according to the weight that assessors gave to each component.

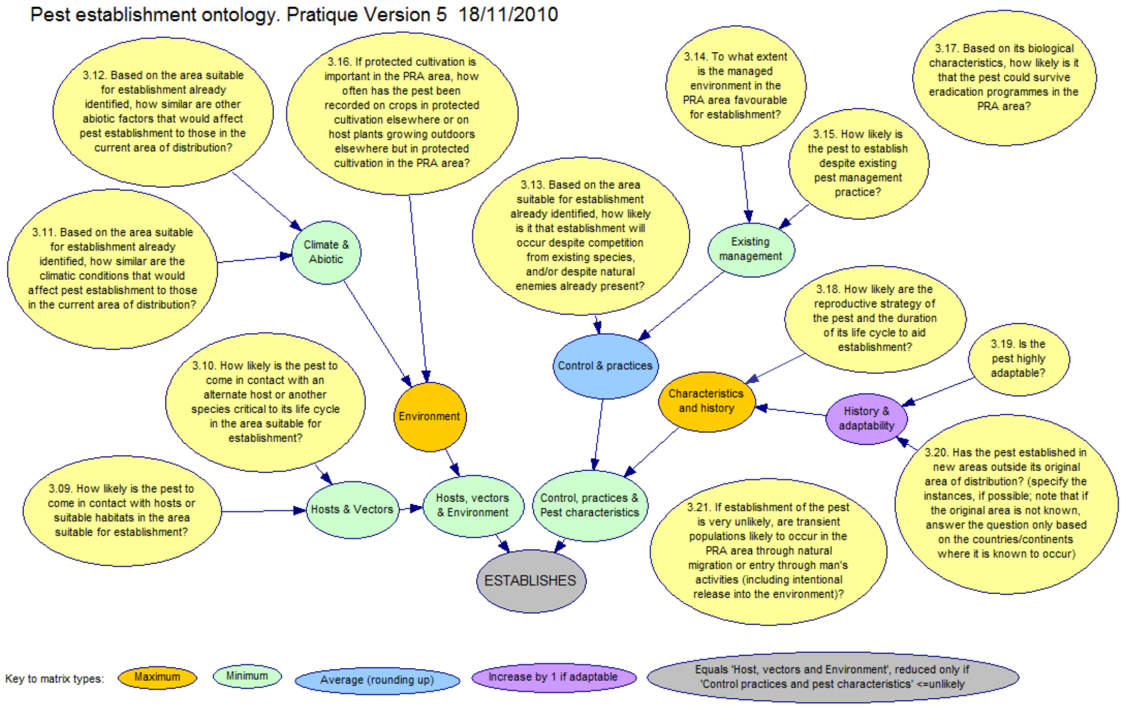

Fig. 6. Ontology of the establishment section (Version from 18 November 2010).

In the matrix model, the representation of uncertainty was adapted from the International Panel on Climate Change (IPCC) definitions IPCC, 2005). Low, medium and high uncertainty were defined as expressing 90, 50 and 35% confidence, respectively, that the rating selected is the correct one. Based on the level of uncertainty given by the assessor, the representation of uncertainty was then defined using standard distributions (Beta or Truncated Normal). Figure 7 shows a beta distribution used to display three levels of uncertainty around the rating "moderately likely" (see Fig. 7).

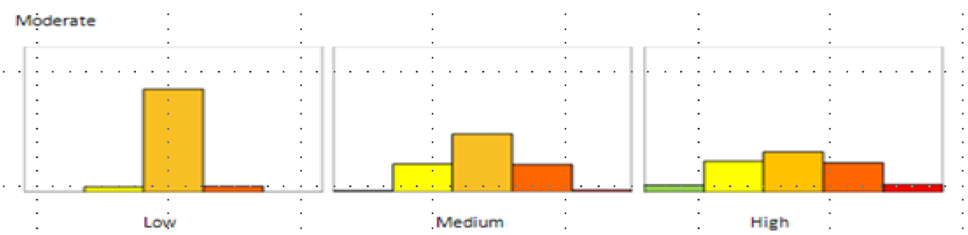

Fig. 7. Distribution of uncertainty: illustration of the distribution of a chosen rating of moderately likely with three different levels of uncertainty (low, medium, high).

The models have been implemented in GeNie2 and Excel; the ontology of the models is graphically depicted in both versions. The Excel version allows the user to select the matrix types from a drop-down menu; changing matrix types in GeNie2 is a more complicated process. GeNie2 software is available as a free download from http://genie.sis.pitt.edu/.

The method is described in detail in Annex 1.

Invasive Risk Impact Simulator (IRIS)

IRIS is a tool developed to interpret the subjective overall scores given by risk assessors for entry, establishment, spread and impact (Mumford *et al.* 2010). It is an Excel based system for converting scores of likelihood/impact (and the expert's confidence in these scores) from EPPO style PRA templates into potential expected annual costs with a time horizon of approximately five years. The qualitative scores are tied to quantitative scales that are explained to the expert prior to conducting the risk assessment.

Entry, establishment, spread and impact can be combined multiplicatively because the final output, potential expected impact (€/year), is directly contingent on the four previous steps. Entry, Establishment and Spread ratings use probability/proportion bandings based on Intergovernmental Panel on Climate Change guidelines (2005; Table 3).

For entry, establishment, and spread, a linear scale was used, whereas for impact, by contrast, a $\log_{10}$ scale was applied. See Annex 1 for an explanation.

Table 3. Probability and proportions attached to qualitative scores used by IRIS (after Qualitative description IPCC, 2005)

Entry Establishment (Probability)	Spread (Proportion)	Min	Max
Very unlikely	Minimal	0	0.1
Unlikely	Minor	0.1	0.33
Moderately likely	Moderate	0.33	0.66
Likely	Major	0.66	0.90
Very likely	Massive	0.9	1

The Impact scale uses five $\log_{10}$ impact classes suggested by the AS/NZS Risk Management Standard (Table 4). This scale can be adapted to different areas.

Table 4. Example of an impact scale using order-of-magnitude steps where Potential Loss has been set to €10 billion per year.

Qualitative description	Probability of occurrence in X years
Massive	€1billion - €10 billion pa
Major	€100 million - €1 billion pa
Moderate	€10 million - €100 million pa
Minor	€1 million - €10 million pa
Minimal	€100,000 - €1 million pa

Experts are asked to attach an uncertainty (or confidence) level to their answer (low, medium and high).

Based on the ratings and levels of uncertainty provided by the assessors, IRIS uses a stochastic simulation process to present a cumulative risk profile for the species being assessed. It generates a graphical representation of risk distribution that is readily understandable (as shown in Figure 8) and is particularly helpful for comparing risks between species (Figure 9). As expert confidence in each of the four component ratings increases (i.e. uncertainty decreases), the range of expected outcomes narrows and the slope becomes increasingly steep.

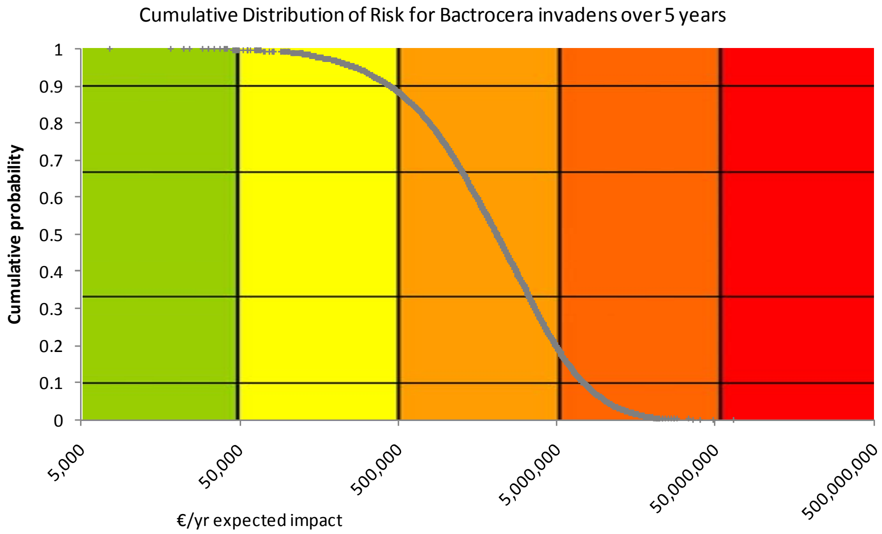

Fig. 8. Cumulative distribution of risk for *Bactrocera invadens* (from Hammamet workshop).

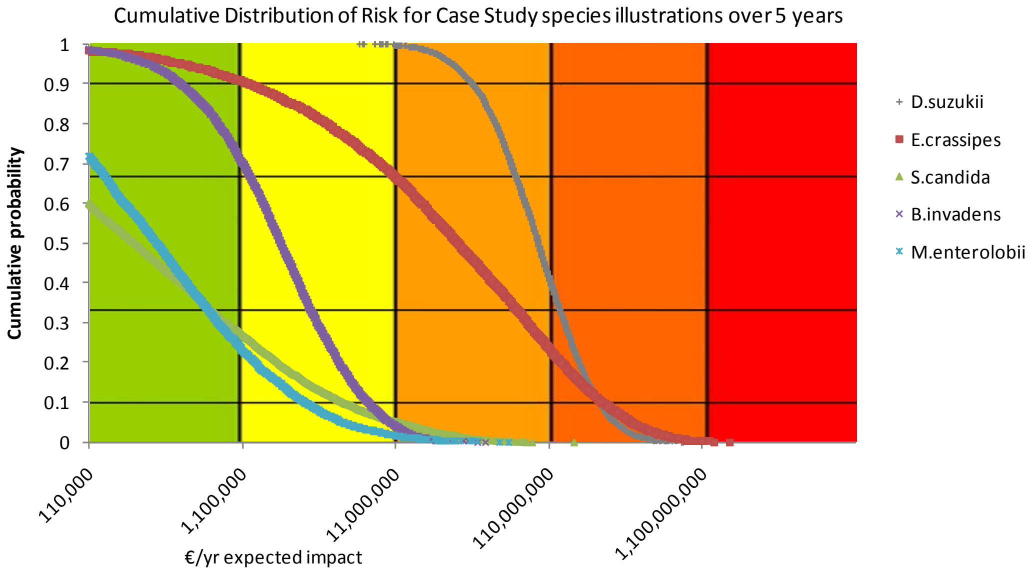

Fig 9. Set of cumulative distributions allowing comparison of different species risks (from Hammamet workshop). IRIS allows the comparisons of threats from very different taxonomic groups (three insects, a water weed and a nematode)

At the Ermenonville PRATIQUE meeting (February 2011), various PRATIQUE consortium members expressed an interest in presenting multiple species in both scatterplot form and in the exceedance (cumulative probability) graphs. The exceedance charts have been the format more regularly used to summarise IRIS outputs for individual species. Because of the profusion of overlapping data points, multiple species represented in the scatterplot graph could not be easily interpreted. It was suggested that hoops which enclosed a certain percentage of all the realisations could be used to summarise each species, which would avoid the need for plotting all the realisations. Figure 10 shows the scatterplot summary for three levels of “inclusion” (25% red; 50% green; 75% blue) against the individual realisations.

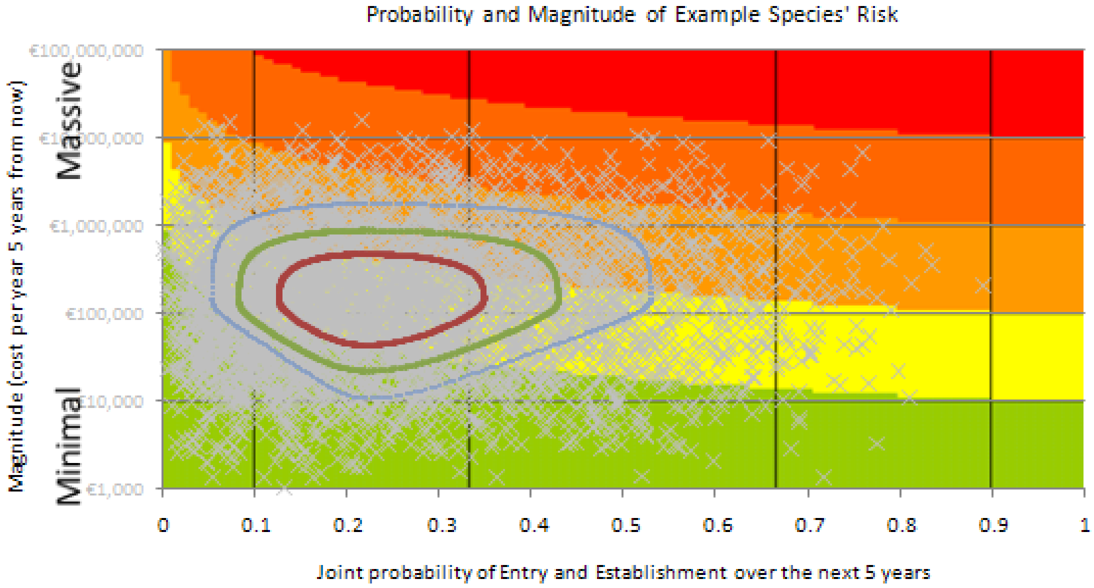

Figure 10. Scatter plot showing three hoops defining the boundaries of inclusion of 25% (red), 50% (green) and 75% (blue) of realisations against the backdrop of the grey realisation crosses (these would not be shown in multi-species representations).

Figure 11 shows a multi-species comparison using the IRIS' "hoops" method. This IRIS hoop visualisation method is informative not only in showing uncertainty (larger ovals indicate greater uncertainty) but also the axis on which the uncertainty is expressed. In Figure 11, the small size of *Buddleja davidii* compared with others indicates a low degree of uncertainty expressed by the assessors. The larger relative size of other species, especially *Bombus terrestris*, indicates greater uncertainty in the assessments. For *Sargassum muticum* the uncertainty is greater in the magnitude dimension than the likelihood dimension giving the hoop a vertically stretched appearance, whereas *Bubo bubo* is wider than it is tall, indicating that there was more uncertainty in the combined Entry and Establishment dimension than the magnitude dimension of Spread and Impact. The position of the centre of each hoop indicates the median IRIS value in each dimension. Species occupying the bottom left hand side present the least threat while those like *B. davidii*, occupying the top right hand side of the chart present the greatest threat.

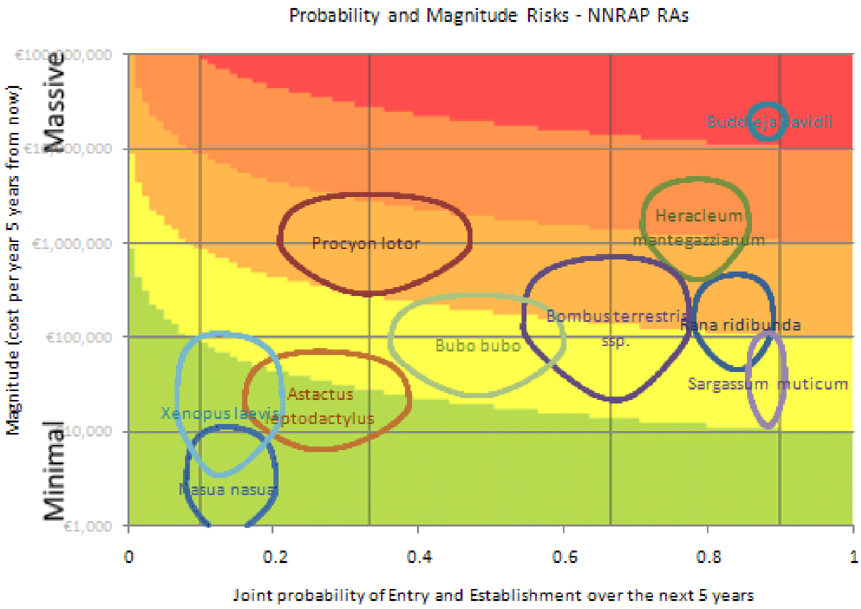

Figure 11. Multispecies comparisons facilitated by IRIS hoop graphic (the hoops in this example enclose 25% of the central part of the distribution - 12.5% each side of the median in each dimension). Note the asymmetry of the hoops for most species, indicating differences in uncertainty between the two dimensions. The bottom left of the graphic shows the species presenting the least threat while those in the top right hand corner present the greatest potential impacts. The colour coded background indicates level of threat of each species, from Green (minimal impact) to Red (massive impact).

In IRIS, a rating is chosen with an associated level of confidence. The level of confidence is tied to a continuous probability distribution which has greater variance when confidence is lower. The shapes of the probability distributions corresponding to the different levels of confidence are explained to the assessors. Entry and establishment are combined in a joint probability (likelihood product) as are spread and impact (magnitude product). Both are then combined multiplicatively resulting in probability distribution of potential costs (€ per year) five years into the future.

IRIS allows a form of sensitivity analysis on the uncertainty of ratings for Entry, Establishment, Spread and Impact. The overall effect of selecting different rating values or the uncertainty surrounding those ratings can be seen by re-running IRIS with different values.

The method is described in detail in Annex 1.

Section 3 Discussion of methods and final selection

The five approaches described in section 2 were subjected to testing by EPPO PRA Development Panel members, members of EPPO expert working groups on PRA, workshop participants (in particular the EPPO/PRATIQUE workshop in Hammamet, Tunisia, in November 2010) and individuals working for PRATIQUE.

The most important criteria (see also Annex 2) used to select the methods for the EPPO DSS for PRA were:

- Transparency, including comprehensibility by risk managers:
  transparency refers to access to the information that is used in decision-making and risk management and is generally considered to be a fundamental feature of risk assessment methods. It implies openness, communication and accountability with regards to the assumptions made and relates to how easy it is for outsiders (referees, other risk analysts, risk managers) to comment on the assessment/analysis and to understand it.

- Rigour in handling uncertainty:
  This criterion aims to assess how well the method accounts for uncertainties that affect the risk estimate. Various techniques are available to account for uncertainties and those can be applied to determine the reliability of model predictions by capturing and propagating the assessors' various uncertainties about variables and mechanisms that affect the risk. Models that do not properly account for uncertainties make it impossible to judge the reliability of the output as a basis for making decisions.

- Consistency:
  Consistency is an important feature of risk assessment methods because it ensures that the same data will lead to the same result and is, as far as possible, independent of the panel that carries out the assessment. The old EPPO DSS for PRA (EPPO, 2009) can lead to a lack of consistency at two levels. Firstly, because the scheme is a qualitative rating system, assessors select a rating for each question and different assessors may provide different ratings for the same question. The second level of inconsistency originates from the fact that the answers to all the questions need to be combined to obtain an overall estimate of risk. Because there are no guidelines for how to integrate the ratings, even when assessors give the same rating for each question in the rating system, the integration step may lead to inconsistencies between assessors (or assessments).
    - Ideally, we would like to use a method that is able to provide consistent results, not only between assessors (e.g. the same decisions would be reached by other assessors provided with the same information) but also "within assessors (e.g. different pests that in reality present similar risks are actually assessed as such).
    - However, inconsistency caused by the first type cannot be eliminated and the only way of dealing with this is by treating inconsistencies between assessors as an additional source of uncertainty. The same applies to lack of agreement between experts on how the ratings should be combined. A method that is consistent, should also be able to deal with uncertainties, but in the selection process of the different methods we ignored the

first type of inconsistency and we only focus on whether the method will lead to consistent results if experts provide the same ratings for every question or risk model parameter.

- Scientific defensibility:
  All of the proposed methods involve the introduction of certain assumptions (particularly with regards to how ratings are combined and the way they account for uncertainties). Even simple averaging of the ratings in a scoring-based system assumes that every question has the same influence on the risk.
  PRATIQUE has developed a series of biologically meaningful, mechanistic approaches but clearly, many of the assumptions (including distributions assigned to model variables) remain to be validated. Some validation of the overall model results have been carried out by comparisons with the independent judgement of assessors for a series of test cases, both historical and during project workshops. These approaches allow for future validation using scientific experiments to estimate model variables and can therefore be regarded as heading towards a more science-based approach for risk assessment.

- Ease of use in routine risk assessments:
  There is a clear difference in the development stage of a model (see above) and the actual use of a model for routine risk assessments. Some of the methods are easier to use than others, even when the model can be packaged in user-friendly software for the analyst. There are relations with other criteria. If the mental model of the user aligns with the model, it is easier to use because the user understands this model. Transparency and ease of comprehension are of importance in this respect. Methods with an appealing visualisation are therefore easier to use than methods with poor visualisation possibilities.

Bayesian Belief Networks (BBNs)

Bayesian methods are transparent, rigorous with regard to uncertainty, consistent and scientifically defensible but require more resources to implement and/or use. Nevertheless, the main advantage of BBNs is that they are capable of propagating uncertainties. Assumptions are explicit and effects can be monitored. The main disadvantage is that expert elicitation is needed to set up and validate conditional probability tables (CPTs) which can rapidly become very large and complex if multiple factors are combined, particularly when each factor consists of many categories. Some investigations of BBNs at PRA question level were made and tested individually, but BBN testing was primarily undertaken at the section level, particularly to combine the likelihood of entry and establishment.

After the PRATIQUE meeting in the Hague in March 2009, 20 participants volunteered to complete CPTs combining the conclusions of entry with the conclusions of establishment to give the likelihood of introduction and combining introduction with spread and finally introduction and spread with impact, to give an overall level of risk. Out of those 20 experts, 13 experts provided CPTs that passed consistency checks (medium + medium has higher risk than medium + low). These inconsistencies arose due to the fact that many of the experts had not been introduced to BBNs and CPTs before the exercise.

A simplistic test of the BBN was conducted by 2 experts who only used the conclusion sections of 7 case studies to convert the summary text into probability distributions The 7 case studies included: 4 insects, 1 fungus, 1 bacterium and 1 invasive plant, selected as they span the range from high to low risk, and have PRAs that include risk ratings and uncertainty scores for each section. Despite the simplicity of the analysis, the low number of experts involved, the results provided two important insights: firstly, the results indicated that it is possible to build a BBN that can be a useful tool for PRA. More importantly, the results indicate that risk analysts have very different views on risk and that this should be captured by any risk assessment method. Not only where there clear differences on how different parts of the risk assessment scheme should be combined (as indicated by large variations in CPTs; see below) but more importantly also in the assignment of probabilities to scores. This indicates that uncertainty between assessors should be captured carefully and reported within assessments.

A question-level BBN based on pre-determined CPTs and on a model structure using GeNIe was then presented to the Panel on PRA Development in May 2010, which tested the model in subgroups with the PRA on *Tetranychus evansi*, to check if the outcomes of the model are consistent with the conclusions of the EPPO Expert Working Group that conducted the PRA. The model that was tested had not been calibrated due to a lack of PRA expert involvement. As a result, the Panel noted various problems that indicated that the model was not reflecting their expertise. The Panel concluded that although the BBN approach is interesting, it needs to be worked on further by the help of experienced risk assessors.

It was noted that experts had different opinions on the combination of risk resulting in differences in the CPTs that can easily be visualised when displaying the CPTs as risk matrices (see Fig 12).

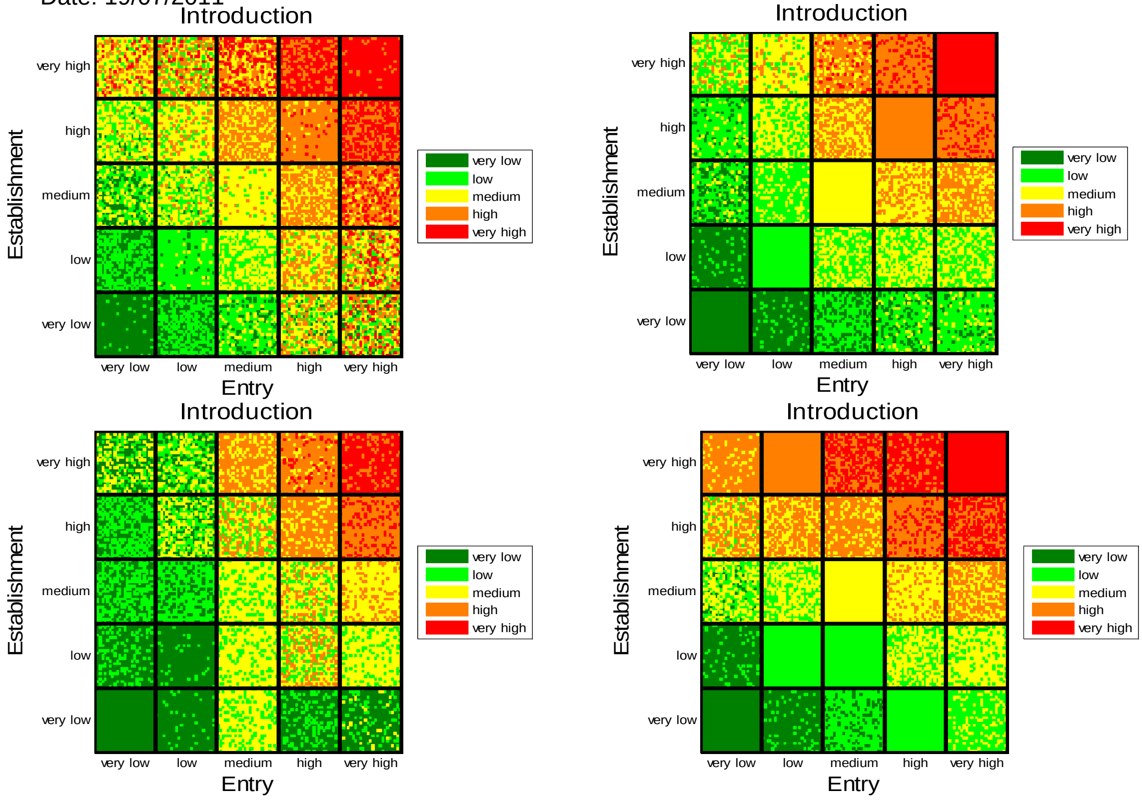

Fig. 12. Visualisation of the differences between conditional probability tables for entry and establishment (= introduction) prepared by six different experts.

The graphical representation in Fig. 9 gives a visualisation of the inconsistencies, and helps experts to challenge their perceptions of combining risk and to understand the process and increases consistency. However, it was also noted by the volunteers that the exercise was difficult and that doing it a second time would likely end with a different result. CPTs should be developed by experts running and using the EPPO DSS for PRA scheme. Due to the complexity, it was considered impossible to produce such tables at the level of individual questions whilst keeping the method flexible. Instead, it was decided to put several restrictions about the model in place, leading to the MM approach. Although the combination at higher level was perceived to be easier to achieve than at individual question level, the attempt to prepare conditional probability tables for the combination of entry and establishment was not successful, more involvement of PRA experts would have been necessary for further development.

The BBN approach was not developed further primarily because of the significant time required from PRA experts to build CPTs, particularly those at question level, and test them for logic and consistency. In addition, because experts may have very different views on how answers should be combined, compiling a unique CPT would need a considerable amount of work, communication and agreement. CPTs are thus considered to provide the major bottle neck for the development of BBNs for the EPPO DSS for PRA if no restrictions are introduced about the model. However, the developments so far have been taken up by the EFSA Art. 36 project “Prima phacie” where they continue to be explored.

The knowledge based approach (KBA)

The first prototype of the KBA was tested by comparing the model results with the results of the evaluations made on different species by eight EPPO PRA Expert Working Groups. These tests were conducted between December 2008 and March 2009 and triggered revision of the prototype. The second prototype (following the PRATIQUE project meeting in The Hague in October 2009) was tested on the eight species from the EPPO PRA Expert Working Groups, plus four other species (*Xanthomonas axonopodis* pv. *allii, Tetranychus evansi, Metamasius hemipterus, Heracleum sosnowskyi, Eichhornia crassipes, Aulacaspis yasumatsui, Raoiella indica, Lysichiton americanus, Meloidogyne enterolobiiIris Yellow Spot virus, Eutypella parasitica, Hydrocotyle ranuncoloides*). The concerns raised during this test, which was conducted in October 2009, prevented any further development of the model.

The main concerns were:

- The final outcome of the model for Entry did not vary greatly when changing from a large volume to a very low volume. In the first prototype of KBA it did, but this dominating effect of volume had to be removed from the model as a result of an expert meeting. Unfortunately, after removal, the volume effect did not seem correct. The reason seems to be the transformation of the ratings to numerical values from 0.05 to 0.95, which may be one of the main obstacles for KBA. For volume of import this does not match the perception of experts regarding the difference between very low volume and high volume. In fact, the difference between the lowest numerical score (0.05) corresponding to “very low volume” and the highest (0.95) corresponding to “high volume” is a multiplication factor of 19 where the perception of experts is that it should be much more, e.g. a multiplication factor higher than 1000. The question arose whether the scales could be different depending on the question. In fact, in the KBA model this can easily be adapted.
- The scale would probably also need to be different for different pests (e.g. to take into account the survival rate of bacteria and insects)
- The mathematical transformations made are not easily understood (e.g. the choice between summation and multiplication) and it is not easy to understand the significance of the percentages calculated by the model and how they relate to risk. It was noted that in the revised model, ratings of the different clusters are combined by multiplication and no longer by summation as was the case in the first model. This was considered more appropriate.
- Agreements on which questions should be weighted and the amount of weighting are very difficult to reach and weighting may even vary depending on the type of pest.

The different experts involved in the development of the model could not agree on the knowledge components and about the further time investment needed to try to solve this. However, elements of the KBA can be found in other methods developed in PRATIQUE and the exercise on clustering questions proved to be very useful when drawing up the ontologies for the matrix models. Further, the KBA has a built-in sensitivity analysis and an uncertainty analysis that might be useful for other methods.

PRA Risk Rating and Uncertainty Visualiser

The PRA Risk Rating and Uncertainty Visualiser was tested at the EPPO/PRATIQUE Hammamet workshop in November 2010 and was considered to be a very useful way of visualising the answers given by the assessors and consolidating the conclusions on the different sections of the PRA. It was also seen to be useful when summarising risk, as the visualisation of answers shows the “main trend” of answers. It was suggested that links should be provided so that it can operate in all sections of the scheme. Explanations should be added on how to interpret the bubbles (large bubbles correspond to a high uncertainty). The questions were identified by their number; a popup system was suggested so that the text of the questions can be seen. Questions related to the different sub-sections appear in different colours but the experts found it difficult to identify outliers. It was suggested that a system independent from EXCEL 2007 should be provided. All these amendments proposed have been implemented in the final version, which has been integrated into CAPRA.

Rule-based Matrix Models (MMs)

The MMs proposed for Entry and Establishment were tested by the model developers on several species and the results were found to be consistent with expert opinion. Consequently, it was proposed that this technique is taken forward for further testing and possible inclusion in the revised EPPO DSS for PRA. One of the advantages of risk matrices (MMs) is that they are familiar to risk managers.

The experts at the Hammamet workshop in November 2010 considered that MMs were not only useful for presenting a summary of risks for individual sections but could also be used to visualise this risk and its distribution and to investigate how a change in a rating or level of uncertainty may impact the overall risk. Nevertheless, the reasons for choosing between the different matrix combination rules (maximum, minimum, average) in the models were not considered to be transparent and it was suggested that explanations should be provided. The ontology and choice of matrices have been reviewed at the Panel on PRA development in January 2011 and the choices made between the different types of Matrices were agreed. Using fixed beta distributions was seen as a pragmatic way to represent uncertainty in the MMs, although it was noted that, for some questions, the expert can be more than 90% certain, and may wish to override the model. Although the Genie software allows this, additional guidance is required (see Annex 1). It was recognised that MMs provide a baseline for comparison and a method for checking the consistency between pests and also between assessors. It was also noted that expert judgements when summarising risk for each section should always be made first so that the MMs are used to aid expert judgement rather than replace it. Guidance has been developed and included in the computer programme CAPRA.

The Panel on PRA Development recommended that the preparation of MMs at a higher level (section level) should also be provided and this has been implemented for the Entry and Establishment sections using a matrix proposed at the PRATIQUE meeting in Ermenonville, France, in February 2011, see Table 5.

Table 5. Section level matrix for Entry and Establishment

Likelihood of establishment	**Very high (5)**	Moderate	Moderate	High	Very high	Very high
	High (4)	Low	Moderate	Moderate	High	High
	Moderate (3)	Low	Low	Moderate	moderate	Moderate
	Low (2)	Low	Low	Low	Low	Low
	Very low (1)	Very low	Very low	Very low	Very low	Very low
		Very Low	**low**	**moderate**	**high**	**Very high**
		Likelihood of entry				

It was decided to use this summary matrix (instead of a BBN at the higher level, see explanations above) as this was easier to develop and preferred by the experts testing the different approaches.

Finally, it has to be noted that although risk matrices should be used with caution, this approach has several advantages (Milestone 3.3/3.7, Annex 2; EFSA, 2010):

- Risk matrices are easy to understand.
- They provide risk assessors with a transparent way for combining scores.
- They can be used to improve the consistency of pest risk assessment.
- They allow the estimation of overall risk levels which can be used to rank different pests.
- They can be used to perform uncertainty analysis with qualitative scores.

The application of MMs has been integrated into CAPRA.

Invasive Risk Impact Simulator (IRIS)

At the EPPO/PRATIQUE Hammamet PRA workshop in November 2010, participants expressed their concerns that, to use IRIS effectively, detailed knowledge of crop value and maximum losses are needed. An alternative was required (and provided) that avoids a "quantitative" x-axis (i.e. Euros) but has the usual five level rating. It was agreed that IRIS is a "semi-quantitative" tool (quantitative with the quantitative part hidden for users), depending on what information is available. The visualisations provided by the tool were welcomed but need clear guidance to help risk managers to interpret results and understand how the graphs are generated. It was not clear how economic and environmental impacts are combined in IRIS. Finally the experts considered that it would be useful to test IRIS with organisms that have not already invaded our region to see how it works and to provide examples of their correct application.

The EPPO Panel on PRA Development preferred IRIS to be used as an additional tool (i.e. not directly integrated into CAPRA) that can be provided to risk assessors but considered that more guidance and examples need to be included to ensure that appropriate input values are selected and the results are correctly interpreted by risk managers.

Expert judgement

Even if model tools and additional guidance are provided to risk assessors, there are still questions for which it has not proved possible to devise detailed guidance and an expert judgement is required to determine a rating or provide an overall score supported by a written justification. Uncertainty in such judgements can be reduced by involving more than one expert in making the judgement (i.e. PRAs prepared in a group rather than by an individual). External review processes also help in reducing uncertainty and providing consistency.

In addition, it should be noted that the models are decision-support not expert-replacement tools. Risk assessors should use the models in conjunction with expert judgement and compare the models with their own assessment. Therefore, a warning is provided in CAPRA when using the models. "Although these models have been developed and tested on a number of examples, they may not be adapted to all pests or situations/scenarios. When an expert is confident in his judgement, this should be preferred in cases where the outcome of the model does not match this judgement and this expert will be asked to provide feedback to EPPO Secretariat so that the model may be reconsidered." In validations carried out so far it has been possible to identify the reason why discrepancies exist between the model and expert judgment. It was usually because certain risk elements have unusually high or low significance for those pests. Thus the models provide a result based on consistent weighting of risk elements, but expert judgement is always to be preferred. The models will have played a useful role if they cause the assessors to reconsider why a species is an exception.

Conclusions

PRATIQUE investigated 28 methods for summarising risk and capturing uncertainty when undertaking qualitative assessments of pest entry, establishment, spread and impact. The methods varied in their transparency (including comprehensibility to risk managers), consistency, rigour in handling uncertainty, ease of use and defensibility. The methods selected (the risk rating and uncertainty visualiser, rule based matrix models (MMs) and the optional invasive risk impact simulator (IRIS)) not only displayed the optimal combination of these factors but, following testing, were warmly received by pest risk analysts. The "daily use" of the models/approaches when conducting PRAs will help to further improve and test the approaches; this experience can only be gained over the years. The results will be followed up by EPPO.

Acknowledgements

The authors would like to thank the many groups and individuals, within the PRATIQUE consortium and in the wider PRA community, who have contributed to the development of these techniques through comments and suggestions at meetings, workshops and in expert working groups. Special thanks go to DirkJan van der Gaag, Dutch Plant Protection Service.

References

AS/NZS (2004). Standards Australia and Standards New Zealand. AS/NZS 4360:2004, Risk Management, Sydney, NSW. ISBN 0 7337 5904 1.

Cox, L.J. (2008). What's wrong with risk matrices? Risk Analysis 28 (2), 497-512

Cullen, A.C. and Frey, H.C. (1999). Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. Plenum Press, New York.

EFSA (2010). Guidance on a harmonised framework for pest risk assessment and the identification and evaluation of pest risk management options by EFSA. EFSA Panel on Plant Health (PLH). EFSA Journal 8 (2),1495-1561.

EPPO (2009). http://archives.eppo.org/EPPOStandards/PM5_PRA/PRA_scheme_2009.dc

FAO (2004). Pest risk analysis for quarantine pests including analysis of environmental risks. International Standards for Phytosanitary Measures. Publication No. 11. Rev. 1. FAO, Rome.

FAO (2007). Framework for Pest Risk Analysis. International Standards for Phytosanitary Measures. Publication No. 2. Rev. 1. FAO, Rome.

IPCC, 2005. Guidance Notes for Lead Authors of the IPCC Fourth Assessment Report on Addressing Uncertainties. Intergovernmental Panel on Climate Change. http://www.ipcc.ch/pdf/supporting-material/uncertainty-guidance-note.pdf

Jensen, F.V. (2001) Bayesian Networks and Decision Graphs. Springer-Verlag, New York.

Van Leeuwen, C.J., and T.G. Vermeire (2007) Risk Assessment of Chemicals, 2nd edition. Springer, Netherlands.

MacLeod, A., Anderson, H., Van Der Gaag, D. J., Holt, J., Karadjova, O., Kehlenbeck, H., Labonne, G., Pruvost, O., Reynaud, P., Schrader, G., Smith, J., Steffek, R., Viaene, N., Vloutoglou, I. (2010) Prima phacie: a new European Food Safety Authority funded research project taking a comparative approach to pest risk assessment and methods to evaluate pest risk management options, EPPO Bulletin 40 (3), 435-439.

Morgan, M.G. and Henrion, M. (1990) Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Cambridge.

Morris, M.W. and P.B. Sayers (2002) Assessing Uncertainty When Predicting Extreme Flood Processes. HR Wallingford.

Mumford, J.D., Booy, O., Baker, R.H.A., Rees, M., Copp, G.H., Black, K., Holt, J., Leach, A.W., Hartley, M. (2010) Invasive species risk assessment in Great Britain. Aspects of Applied Biology, 104: 49-54.

NRC (2000) 'Risk analysis and Uncertainty in Flood Reduction Studies'. National Research Council (US). National Academic Press

Pearl (1985). "Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning". In Proceedings of the 7th Conference of the Cognitive Science Society, University of California, Irvine, CA, pp. 329-334, August 15-17.

(More references can be found in Annex 1 and 2.)

List of Annexes

In Annex 1, the methods that were selected (PRA Risk Rating and Uncertainty Visualiser, MM and IRIS) are described in detail. This corresponds to the "protocol for quantifying and communicating uncertainty and for summarising and communicating overall risk in the PRA scheme accessed via a hyperlink in a project web page" (a combination of the two titles of deliverables D 3.2 and D 3.4.

In Annex 2, the Milestone 3.3/3.7 can be found, that was included in the 18 month report.

In Annex 3, an example is given, showing CAPRA at work with *Drosophila suzukii.*

Section	Title	File Name
Annex 1	PRA Risk Rating and Uncertainty Visualiser, Matrix Models, and Invasive Risk Impact Simulator	Annex 1_D3_2_D3_4 final.doc
Annex 2	Best practice for quantifying uncertainty and summarising and communicating risk PM No. M3.3: Best practice for quantifying and communicating uncertainty identified M3.7: Best practice for summarizing & communicating risk identified	Annex 2_M_3.3_M_3.7_Uncertainty_and_Summarising_Risk_FINAL.doc
Annex 3	PEST RISK ANALYSIS FOR : *Drosophila suzukii*	Annex 3_Drosophila_suzukii test final 2011-05-19.pdf