Australian Government
Department of Agriculture

cebra
Centre of Excellence for Biosecurity Risk Analysis

CEBRA Report Cover Page

<table>
<tr><td>Project Title, ID & Output #</td><td colspan="4">Development of a Marine Spatial Analysis Model for Improved Biofouling Risk Assessment, 1402A, Output ID#4</td></tr>
<tr><td>Project Type</td><td colspan="4">Final Report</td></tr>
<tr><td>DAFF Project Sponsor</td><td colspan="2">Tim Chapman</td><td>DAFF Project Leader/s</td><td>Chris Starkey
Peter Stoutjesdijk</td></tr>
<tr><td>CEBRA Project Leader</td><td colspan="2">Mark Burgman</td><td>NZ MPI Collaborator</td><td>NA</td></tr>
<tr><td>Project Objectives</td><td colspan="4">The aim of this project is to investigate high-risk pathways for the arrival and survival/establishment of marine pests of concern, based on compatibility between Australian ports, biofouling temperature tolerance and distribution and the last ports visited by international vessels. Due to limitations in the availability of species level data for all biofouling species, this aim will be addressed through expert elicitation approaches to supplement the available data in a way that is logical and statistically consistent.</td></tr>
<tr><td>Outputs</td><td colspan="4">The project will provide an approach to assessing high-risk pathways for biofouling. The project’s outcomes will inform the Department’s risk assessment methods underpinning the proposed biofouling management arrangements for internationally arriving vessels. Irrespective of the risk management options implemented by the Department, they will require an underlying risk assessment methodology. The outcomes of this project could also inform other management actions. On a broader scale, the outcomes will improve understanding of important information gaps.
The project’s outputs include:
• A review of relevant literature
• Identification and recruitment of relevant expertise
• The development and implementation of an expert elicitation protocol
• Predictive modelling based on available data and expert judgements
• The final report detailing the risk tool and its application</td></tr>
<tr><td rowspan="2">CEBRA Workplan Budget</td><td>Year 2013-14</td><td>Year 2014-15</td><td>Year 2015-16</td><td>Year 2016-17</td></tr>
<tr><td></td><td>$120,000</td><td></td><td></td></tr>
<tr><td>Project Changes</td><td colspan="4">The timeline for data and expert availability was revised to the end of October 2014, and other time lines were pushed back one month in agreement with all participants. The expert’s project meeting was held on 31 October 2014 at the CSIRO in Canberra. Peter Caley, Chris Starkey, Simon Barry, Emma Johnston, Graeme Clarke and Tony Arthur revised the project aims, and jointly developed an experimental protocol for expert elicitation (see progress report distributed in December 2014). Draft and final reports were prepared and delivered according to this revised timeline.</td></tr>
<tr><td>Research Outcomes</td><td colspan="4">The results presented in this report represent a novel approach to estimating the risk posed by hull fouling on vessels. This has been a difficult problem. Introduction events are rarely observed so it not possible to construct direct estimates of the probability of introduction. Physiological models and data about the biogeography of species are poorly developed so the use of process-based models has not been feasible. While it is important to not overestimate the ability of experts to understand complex phenomena, the study nonetheless provides a transparent and systematic summary of their views, which is a significant advance on innuendo and anecdote.

The model as specified is well suited to risk ranking applications. It cannot be applied to determine absolute risk and therefore, given a threshold, whether a journey’s risk is acceptable or not. If there were a requirement to do this, further analysis using the marginal rate of introductions would be needed. Thus the model as it stands could directly support compliance targeting. To apply it to primary estimation of risk would require both additional analysis as well as work to integrate it logically with the Marine Growth Risk Assessment (MGRA) tool.</td></tr>
</table>

<table>
<tr><td rowspan="4">CEBRA Use only</td><td>Received by:</td><td>Date:</td></tr>
<tr><td>CEBRA / SRC Approval</td><td>Date:</td></tr>
<tr><td>DA Endorsement (   ) Yes (   ) No</td><td>Date:</td></tr>
<tr><td>Report published</td><td>Date:</td></tr>
</table>

Recommendations	This project developed a model that evaluates biosecurity risk (likelihood of marine pest establishment) based on the vessels movement history and the characteristics of Australian ports. It contributes to risk targeting for biofouling, and benefits both the Australian environment and the international shipping industry. The outputs from this project enable additional information on vessel biofouling risk to be incorporated into overall risk assessment systems, potentially resulting in more comprehensive risk assessment outcomes, and reduced cost to industry as risk management will be focussed around the highest risk vessels.
Related Documents	CEBRA project 1302A, *Evaluating spatial analysis tools for surveillance and monitoring in marine environments*, precedes this project. Documents related to this project should be reviewed in conjunction with this report.
Report Complete	Date: 30/06/15

CEBRA Use only	Received by:	Date:
	CEBRA / SRC Approval	Date:
	DA Endorsement ( ) Yes ( ) No	Date:
	Report published	Date:

BIOSECURITY
www.csiro.au

Development of an Expert-based Model for Improved Biofouling Risk Assessment

CEBRA Project No 1402A

Simon C. Barry, Peter Caley, Shuang Liu, Dean R. Paini, Jan Carey and Graeme Clark

June 2015

Centre of Excellence for Biosecurity Risk Analysis

Citation

Barry SC, Caley, P., Liu, S., Paini, DR., Carey, J., and Clark, G. (2015) Development of an Expert-based Model for Improved Biofouling Risk Assessment. CSIRO, Australia.

Copyright

Important disclaimer

Contents

Figures

Tables

Acknowledgments

Tony Arthur, Mark Burgman, Emma Johnson and Chris Starkey are thanked for contributing to valuable discussions and planning. We thank the numerous experts who generously gave their time for the project. The report was improved by the comments of two anonymous reviewers.

Executive summary

This report seeks to enhance the Department of Agriculture's existing biofouling risk assessment tools, which have been developed to underpin a proposed regulatory approach to biofouling on vessels entering Australian waters. The Invasive Marine Species Programme is in the process of implementing new biofouling management arrangements, and a risk assessment system (the Marine Growth Risk Assessment or MGRA) has been developed to accompany these. The MGRA will operate alongside existing quarantine pre-arrival processes, and consider information about vessel biofouling management such as presence and age of antifouling coating, history of recent hull survey and marine growth inspection, evidence of seawater pipe work treatment, duration of stay in overseas locations, and duration in Australian ports and coastal waters. However it does not have the capacity to identify high risk pathways based on compatibility between Australian ports, the temperature and salinity tolerance and distribution of invasive marine species of concern and the last ports visited by internationally arriving vessels. Here, we report on a project seeking to improve existing risk assessment methods by undertaking expert elicitation on various incursion scenarios, and statistical analysis of outputs.

The project involved twenty nine national and international experts each rating the relative risk of twenty comparative scenarios. These ratings were used to construct a statistical model emulator summarising the expert's views. The summary of these is as follows.

The estimated impact of anti-fouling coat condition on relative risk was large, as expected. Compared to a new anti-fouling coat, having anti-fouling at 50% of its service life increased the risk by a factor of about 2. Similarly, compared to a new anti-fouling coat, having anti-fouling at 100% of its service life increased the risk by a factor of approximately 4.

Moving from an enclosed river port to a coastal port lowered the estimated risk by 54% compared to moving between either coastal ports or river ports.

The environmental similarity of ports was inferred to be an important cue by experts. Each degree increase in the absolute difference in mean sea surface temperature from the port of origin changed the relative risk by 0.923 (*c.* 8% decrease). Similarly, each unit change in the absolute difference in mean water salinity from the port of origin altered the relative risk by 0.97 (a 3% decrease).

Each additional day of berthing increased the relative risk by 1.033 (a 3.3% increase).

There was no evidence that experts consistently factored in similarity in tidal range, or voyage duration, into their estimates.

The model emulator developed is well-suited to risk ranking applications. It cannot be applied to determine absolute risk and therefore, given a threshold, whether a journeys risk is acceptable or not. If there was a requirement to do this, further analysis using the marginal rate of introductions would be needed. Thus the model as it stands could directly support compliance targeting. To apply it to primary estimation of risk would require both additional analysis as well as work to integrate it logically with the MGRA tool.

1 Introduction

1.1 Introduction to hull fouling issue

Biological invasions are among the most serious causes of global biodiversity loss (Butchart *et al.* 2010). Introduced species can displace natives and interfere with fundamental ecological processes, causing impacts that are difficult if not impossible to reverse (Mack *et al.* 2000). In the marine environment, shipping is considered the dominant cause of invasion. Global-scale analyses have demonstrated strong spatial links exist between historical shipping patterns and invasion records (Seebens *et al.* 2013; Capinha *et al.* 2015), and shipping is a known vector for approximately 70% of marine species introduced by human-assisted pathways (Molnar *et al.* 2008). Potential translocation of biota is inherent to many human activities and is unlikely to ever be fully contained, but strategic management informed by risk assessment may significantly curtail the problem.

Shipping can facilitate bioinvasion through two major pathways: ballast water and hull fouling. Ballast water generally transports water-borne organisms, such as plankton, juvenile fishes and invertebrate larvae, and historically has been the focus of most ship-based risk assessments (Drake and Lodge 2004). Hull fouling, on the other hand, transports sessile organisms (usually algae or marine invertebrates) attached to hard surfaces on vessels, which release propagules or fragment to establish populations at new locations. It is difficult to partition the role of these two pathways in specific invasion events since many species can travel by either vector, although there are cases where circumstances strongly suggest hull fouling to be the cause (e.g. Apte *et al.* 2000). In the past ten years the importance of hull fouling has been increasingly recognised (Drake and Lodge 2007), partly due to more stringent regulations in ballast water management. Efforts are now being made to better understand risks associated with hull fouling (Sylvester *et al.* 2011) in order to fill this gap in risk assessment and mitigation.

The likelihood of a species being transported via hull-fouling is dependent on its physiological and life history traits. Hull fouling species require a life cycle amenable to attachment to hard surfaces (e.g. a pelagic larval stage and sessile adult stage), sufficient attachment strength to withstand voyages (Murray *et al.* 2012a), and environmental tolerance suitable for survival in new locations. Tolerance to antifouling paint has also emerged as a key factor (Piola *et al.* 2009), particularly since copper became the international standard on commercial vessels in 2008 after tributylin tin (TBT) was banned (Dafforn *et al.* 2011). Copper paint is an effective biocide against most species, but some can develop a degree of tolerance and colonise surfaces regardless of anti-fouling. These copper-tolerant species can then act as secondary surfaces for the settlement of other, less copper-tolerant species (Floerl *et al.* 2004), facilitating transport of a variety of taxa. The type and quantity of taxa transported also depends on the vessel type and design. Both commercial (Ruiz *et al.* 2013) and recreational (Murray *et al.* 2011) vessels contribute to hull fouling transport, but large vessels with sea chests may pose a particular threat because of the protection sea chests provide to fouling organisms (Coutts *et al.* 2003).

1.2 Scope of policy problem

This report seeks to address deficiencies in the department's existing biofouling risk assessment tools, which have been developed to underpin a proposed regulatory approach to biofouling on vessels entering Australian waters. The Invasive Marine Species Programme is in the process of implementing a new biofouling management arrangement, and a risk assessment system (the Marine Growth Risk Assessment or MGRA) has been developed to accompany these. The MGRA will operate alongside existing quarantine pre-arrival processes, and consider information about vessel biofouling management such as presence and age of antifouling coating, history of recent hull survey and marine growth inspection, evidence of seawater pipe work treatment, duration of stay in overseas locations, and duration in Australian ports and coastal waters. However it does not have the capacity to identify high risk pathways based on compatibility between Australian ports, the temperature tolerance and distribution of invasive marine species of concern and the last ports visited by internationally arriving vessels. Here, we report on a project seeking to improve existing risk assessment methods by undertaking expert elicitation on various incursion scenarios, and statistical analysis of outputs.

This research need is recognised nationally as both an area of research deficiency and priority by those state and territory jurisdictions with interests in biofouling management. The department's biofouling risk assessment tools are being used by several jurisdictions on a trial basis, and any improvements to these tools would also provide flow on benefits to these jurisdictions' biofouling research and management, thus ensuring a nationally consistent approach to biofouling management as supported by key industry stakeholders.

1.3 Challenges of species specific approaches

Bioinvasion risk assessment can be broadly divided into species-specific or environmental based approaches (Barry *et al.* 2008). Species-specific approaches estimate invasion risk per species, informed by occurrence records at potential donor and recipient locations, life history traits, and environmental tolerance at each life history stage (e.g. Hayes and Hewitt (2001)). They consider the degree to which potential invaders are suited to conditions in recipient environments, and may consider a range of additional biological attributes such as behaviour, habitat requirements, interspecific interactions, or density-dependent intraspecific processes. While theoretically valid, species-specific assessments are data intensive and often impractical unless limited to a few key species of interest (Barry *et al.* 2008). The necessary data for many species is often unavailable, or, where it is available, time consuming to collate. Furthermore, obtaining a short-list of high-risk species to assess can be a non-trivial exercise in itself, making the process as a whole difficult to resource, unless the risk assessment process has a large qualitative component (Hewitt[1] *et al.* 2011).

Another impediment inherent to the species-specific approach is the method of estimating environmental tolerance. Ideally, environmental tolerance is gathered from experimental data or physiological models, representing a species' fundamental niche. However, such data typically only exist for a small proportion of the species of interest, so it is common practice to estimate environmental tolerance from their distributional ranges instead (Thuiller *et al.* 2005). This can be problematic in that distributional data represents the realized rather than fundamental niche, and may not be indicative of environmental tolerance if a species distributions are significantly influenced by biotic interactions, dispersal barriers, or ongoing propagule supply that perpetuates a population beyond the range that it would otherwise viable (Pulliam 2000). This method can therefore lead to erroneous estimates of tolerance, and the alternative of gathering direct estimates of tolerance can be prohibitively laborious.

1.4 Report structure

This report is arranged as follows. Section 2 outlines the methodology, reviews approaches to elicitation and describes the scenarios that were used with experts. Results are presented in Section 3 and the report concludes with a discussion of these in Section 4.

2 Methods

2.1 Overview

An alternative to analysing individual species is to consider the estimation of aggregate risk for entire vessels for the establishment of marine species of concern. If this risk can be related to the characteristics of the vessels and the journeys they are undertaking it can be used as a risk assessment screening tool to support inspections and compliance monitoring. Because of the lack of relevant data, we estimate these risks by using experts. Experts can use their knowledge and experience to express their views about risk, and this can be used to make judgements. We could use experts to develop scoring systems directly but there is significant evidence that experts can find this difficult (Barry 2015). Instead we ask experts to consider hypothetical scenarios and then use statistical methods to build models of the experts.

2.2 Scenario calibration method

This report utilises scenario based calibration to produce an estimate of risk. The rationale behind this approach is detailed in Barry (2015). In summary, scenarios are constructed and experts provide their views about the risk endpoint of interest. These views are summarised statistically across the sample of experts. A model is built relating the risk to the attributes of the scenario, in this case the vessel and voyage characteristics, and the environmental features of donor and recipient ports.

A novel component of this analysis was use of relative assessments. The risk of translocation and establishment of species is very low on a per vessel basis. Initial discussion of this with experts identified that they would firstly not be particularly good at this, and secondly would find it very uncomfortable to put absolute numbers to the risk. To address this problem we have modified the elicitation to consider the relative risk of journeys. To understand the motivation for this consider the following.

Assume that the probability p_i of an establishment occurring from a vessel movement follows a logistic model:

$$\log\left(\frac{p_i}{1-p_i}\right) = \alpha + X_{1i}\beta_1 + X_{2i}\beta_2 + X_{3i}\beta_3 + \cdots + X_{ki}\beta_k$$

where the X’s are functions of the *k* attributes of the vessel and the source and destination ports considered to materially influence the risk of establishment. If p_1 is the risk from the source port to destination 1 and p_2 is the risk from the source port to destination 2 then:

$$\log\left(\frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}}\right) \approx \log\left(\frac{p_1}{p_2}\right)$$

given that p_i is close to zero. We therefore elicited from the experts the quantity $\frac{p_1}{p_2}$. Simple algebra gives that

$$\log\left(\frac{p_1}{p_2}\right) = \alpha + X_{11}\beta_1 + X_{21}\beta_2 + X_{31}\beta_3 + \cdots + X_{ki}\beta_k - (\alpha + X_{12}\beta_1 + X_{22}\beta_2 + X_{32}\beta_3 + \cdots + X_{k2}\beta_k)$$

Note that common terms, such as the intercept, cancel out. This will occur for all terms that are common between the journeys such as simple attributes of the source port and vessel characteristics. There are obviously additional uncertainties introduced by the use of experts. A number of points should be noted:

- The elicited response does not allow estimation of absolute risk for a particular journey. This affects the ability to use the resulting model to assess whether a particular journey is below a particular threshold (such as one derived for Australia’s Acceptable Level of Protection). Relative risk is still useful for ranking threats which is important for problems such as compliance targeting. Additional analysis (e.g. marginalising over the observed rate of species establishments in Australia) can be used to convert the assessments into estimates of absolute risk.
- The covariates will sometimes be functions of the difference between the source and destination ports. For example the environmental difference between source and destination port may be the major determinant of risk. When we ask experts to assess relative risks we are comparing two different journeys. This means there can be two different levels of “comparison” in each scenario. This does not present a problem, but is important to understand in interpreting the method.
- As outlined in Barry (2015), some experts may be more or less pessimistic. The relative nature of the analysis means that these effects cancel out and does not need to be included in the analysis.
- The elicitation process will obviously introduce uncertainties as the experts do not possess perfect knowledge. For instance for scenario *i* rated by expert *j* they will not provide $\frac{p_1}{p_2}$ but

  $$y^*_{ij} = \frac{p_1}{p_2} + b_{ij}.$$

  The term b_{ij} is a random effect that represents the uncertainty introduced by the expert’s lack of knowledge. The variation in b_{ij} will vary across experts as some will have more expertise than others. This will be included in the analysis.

2.3 Process

An initial workshop of project participants and Department of Agriculture was held to determine project scope, methods and timelines (participants and report in Section 5.5). This workshop identified that the project should, as much as was possible, not consider factors that were included in the MGRA tool. The primary focus was on environmental factors. Antifouling was included as it is a primary factor and not including it could potentially introduce excessive ambiguity to the process. Scenarios were developed and trialled in the project team before being deployed as described below.

2.4 Expert elicitation

2.4.1 Applying expert elicitation in biosecurity

Biosecurity management is inevitably an exercise in decision-making under uncertainty. A high level of uncertainty prevails within each step of the invasion process, including how human actions can alter the process of invasion (Liu *et al.* 2011). Although decision-makers would prefer to have empirical evidence to develop predictions of biosecurity management outcomes, often such data are not available and cannot be made available before an initial decision must be made. This is because the risk posed by biological invasions belongs to a category of “extreme risks,” which is characterised by low probability of occurrence and potentially high impacts (Burgman *et al.* 2012). The low probability means that the limited data that has been collected about invasive species and their impacts is not representative (Franklin *et al.* 2008). In such a situation of data paucity, it is becoming increasingly common to rely on *expert elicitation* (EE) to characterize uncertainty and fill data gaps (Burgman 2005; Kuhnert *et al.* 2010; Martin *et al.* 2012; Drescher *et al.* 2013).

EE is defined as a formal systematic process to obtain expert judgments on scientific questions (U.S. Enviornmental Protection Agency 2011). It is typically used to quantify ranges for poorly known parameters, but it may also be useful to develop definitions, conceptual models, etc. (Knol *et al.* 2010). It should be noted, that the results of an EE are not limited to the quantitative estimates or qualitative models derived, and they also include the rationale of the experts regarding the available evidence that they used to support their judgments and how they weighed different pieces of evidence (U.S. Enviornmental Protection Agency 2011).

The origins of EE can be traced to the advent of decision theory and decision analysis in the early 1950s. Over time it has been recognized as a powerful and legitimate method by government agencies, the private sector, academia, and other groups (U.S. Enviornmental Protection Agency 2011). For example, the U.S. Environmental Protection Agency started to explore the use of EE in the late 1970s and the Intergovernmental Panel on Climate Change (IPCC) has used EE for more than a decade to address specific components of the climate change issue (IPCC 2001).

A general EE approach includes five steps: deciding how information will be used, determining what to elicit, designing the process of eliciting judgments, performing the elicitation, and translating the elicited information for use (Martin *et al.* 2012). Even though there are such standard procedures and detailed guidelines having been developed (O'Hagan *et al.* 2006), the implementation of EE should be flexible and it is always necessary to customize it for specific applications (U.S. Enviornmental Protection Agency 2011). Factors such as the types of uncertainties considered, the intended use of the elicited information, and the available resources are believed to be important in determining the specifics of each EE application (Knol *et al.* 2010).

Information may be elicited directly or indirectly (Low Choy *et al.* 2009; Kuhnert *et al.* 2010). Direct elicitation requires experts to express their knowledge in terms of the quantities required by the EE analyst. For example, the expert may be asked to provide statistical summaries (e.g., a lower bound, a best estimate and an upper bound) or a full parametric probability distribution. Indirect elicitation requires experts to answer questions that relate to their experiences. Their responses are then encoded into the quantities needed. For example, the expert may be asked about expected economic cost given different efforts of eradication, which the analyst then translates into a corresponding probability distribution for a model parameter. Many experts feel more comfortable when information is being elicited indirectly (James *et al.* 2010).

If an elicitation process involves multiple experts, information can be elicited independently and then combined, or a group opinion can be sought (Martin *et al.* 2012). In either case, the process is subject to 'heuristics and biases' as well-documented by the psychological research on assessing uncertain information (Kynn 2008). In a recent study, researchers found experts appeared to use both a spatial proximity heuristic (they rated the species in their own country as having greater impacts) and an affect heuristic (economic impacts were believed to be more severe than equivalent environmental impacts) when assessing the potential impacts of marine invasive species (Davidson *et al.* 2013). When EE involves a group process, there would be extra concerns such as dominant group members and groupthink (Burgman 2005).

To overcome these limitations, structured approaches such as *the Delphi method* can help in allowing experts to adjust their own estimates in light of the answers of other group members while maintaining anonymity (Doria *et al.* 2009). This method involves getting 'position statements' from individual experts, circulating these, and allowing the experts to adjust their own opinions over multiple rounds. Typically, anonymity is maintained through written interactions, and more dynamic interactions can be achieved through modified Delphi methods with facilitated discussion (Burgman *et al.* 2011b).

Following the reviewer's suggestion, we've updated the corresponding paragraph on page 11 to, "In addition to the *Delphi method* (Linstone and Turoff 1975), there are other *behavioural approaches* that aim to create a consensus among experts (Drescher *et al.* 2013) . By comparison*, mathematical approaches* either apply probably theory or fuzzy set theory in order to aggregate information from different experts (Drescher *et al.* 2013)."

Regardless of the different approaches used, the outcome of aggregation usually includes a measure of central tendency (e.g. a best estimate) and at least a measure of variability (e.g. 'credible interval') (Aspinall 2010). A more sophisticated four-step procedure includes elicitation of a best guess or most likely value, the lowest value, the highest value, and a confidence interval (Speirs-Bridge *et al.* 2010). Given the frailties and limitations of expert judgments, it is advisable to 'train' experts by outlining a field's jargon and theoretical concepts and by using case studies and hypothetical scenarios to illustrate processes relevant to the questions at hand (Burgman *et al.* 2011a; McBride *et al.* 2012a; McBride *et al.* 2012b).

Only recently has EE been applied in the field of biosecurity management, and we review three case studies below. These show that EE can be applied for different invasive species (vertebrates, plant and disease related to animal health) and different purposes (model parameterization, constructing a conceptual model and quantifying uncertainties).

In order to design a statistically powerful surveillance program for invasive vertebrates on Barrow Island, Western Australia, researchers invited a group of six independent specialists with local knowledge to assist the parameterization of the statistical model for the surveillance system (Jarrad *et al.* 2011). The elicitation was held during two 1-day group workshops, and group consensus was sought in workshops such that the majority of opinion was captured. The EE process provided information and insight when little or no published information was available and it also built ownership by a stakeholder group.

EE was also applied in an effort to predict the potential distribution of an invasive riparian lippia *(Phyla canescens)* across a 26,000 km^2 catchment in eastern Australia (Murray *et al.* 2012b). A Bayesian belief network used to capture the mechanistic understanding of the distribution was generated from expert input at a 2-day workshop. The group of experts included four academics, one land management officer and six landholders. A facilitator encouraged individual expert input through breakouts of small groups (3-4 experts) combined with entire group participation for cross-feedback and group consensus. Experts were asked to identify and rank the key environmental variables associated with establishment and persistence of lippia, and they also discretised each of the key environmental variables into categories. Furthermore, the experts were asked to complete conditional probability tables by giving probabilities for lippia growth and spread. Last, feedback was sought on two occasions when experts involved in the original workshop and new experts were encouraged to critically analyse the completed model and the results of the research.

Using formal EE techniques, researchers were able to quantify the uncertainty of the government's own experts in estimating the costs of exotic livestock diseases in England (Gosling *et al.* 2012). Three elicitation workshops were held with 10 veterinarians and economists from the U.K. Department for Environment, Food and Rural Affairs. Two facilitators with experience in facilitating group elicitation sessions, expertise in statistical modelling, and extensive experience of risk analysis led the elicitation process and carried out calculations with the elicited judgments. The facilitators used a systematic approach to eliciting a range, median, and quartiles for each of the four parameters. In the elicitation process, the experts were asked to discuss the quantities of interest and relevant evidence and then arrive at a group consensus for their judgments on each parameter. After seeking feedback about the initial results, a Monte Carlo estimation procedure using 50,000 iterations was used to analyse the uncertainty in the total cost.

Recently EE using the method of Barry (2015) has also been successfully applied in the field of marine biosecurity management, and we review three case studies below. These show that EE can be applied for different invasive species (eradication feasibility, ballast water management, determining cause of range extension) and different purposes (model parameterization, constructing a conceptual model and quantifying uncertainties).

Crombie *et al*. (2008) used expert opinion to investigate the feasibility of the eradication of marine pests. They surveyed experts asking them to predict the probability of successful eradication of a number of marine pest scenarios given different levels of expenditure. Knight *et al*. (2007) also used expert opinion to determine the relative biological risk of exchanging ballast water at various locations compared to a port environment. Their results were incorporated into Australia's domestic ballast water management strategy. Darbyshire & Caley (2009) illustrated how a scenario based approach could be used to make inference on whether a range extension of a marine pest had an anthropogenic cause (e.g. biofouling, ballast water) as opposed to natural spread.

2.4.2 Choice of experts

The project team and our collaborators in the Department of Agriculture identified a small group of experts based on their track record, experience and knowledge in marine biofouling issues. Then we used the snowball technique to identify further experts (Reed 2008). In total we sent 82 invitations to the experts and 51 of them expressed interest in participating in the survey at an early stage of the project.

2.4.3 Structured elicitation protocol

We used a structured procedure (Figure 1) for expert elicitation, adapted from the workshop-based (Burgman *et al.* 2011b) and email-based (McBride *et al.* 2012b) procedures.

2.4.4 Survey procedure

We elicited from experts their estimate of the risk that an arriving vessel will result in the establishment of an exotic marine pest using a relative risk estimation approach. Specifically, the risk end point of interest was defined as "the establishment of an exotic marine pest for at least one year in the recipient ports, as a result of the vessel berthing as intended".

Two scenarios were presented to experts of a vessel travelling to either of two Australian ports (see Section 2.5 for scenario detail). The port of arrival differed between the scenarios ("Reference port" and a "Comparison port"), though with a similar port of origin ("port of last call"). The experts were asked to estimate their expectation for the relative risk of an establishment occurring in the Comparison scenario compared to the Reference scenario.

The relative risk to be calculated for each scenario is:

$$\text{Relative risk (Comparison vs. Reference)} = \frac{P(\text{Incursion in Comparison scenario})}{P(\text{Incursion in Reference scenario})}$$

where P(…) denotes the probability of the event occurring. Experts were first asked for their belief for the lower bound of the relative risk, followed by the upper bound, then finally their best estimate of the relative risk.

The expert elicitation process lasts about four months (early February to late May), and it can be divided into the four stages of: pre-survey preparation, first survey, group discussion (between the two survey rounds) and second survey.

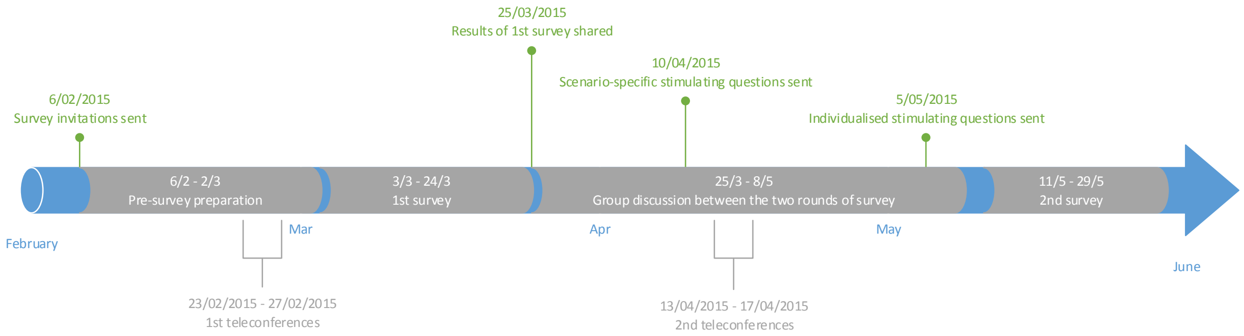


Figure 1. The procedure and timeline of the expert elicitation.

Pre-survey preparation

We sent 82 invitations and received 42 positives replies and 9 answers of "Maybe." All the experts were invited to take part in a teleconference that aimed to provide more information about the project. The project team hosted two teleconferences during the week of 22nd February, one for Australian and New Zealand participants, and one for overseas experts.

Over 30 experts participated in the two teleconferences, which each lasted for about an hour. During the teleconferences, the project team introduced the purposes and design of the project, demonstrated the approach of risk assessment with a sampling survey question and detailed the process of expert elicitation. We spent about 30 minutes on answering experts' questions during the teleconferences. Experts were also encouraged to continue to ask questions after the teleconferences by email. A meeting summary was shared with those who could not make it to the teleconferences.

1st survey round

We randomly assigned the 51 experts into two groups with 26 in Group 1 (G1) and 25 in Group 2 (G2), without any stratification. In early March, the project team sent two versions of the survey to each group, who had 10 business days to answer it. Multiple reminders were sent during this period and we also provided a week's extension for those who were too busy to finish answering the survey on time.

By late March, we received 14 survey answers from G1 and 15 from G2.

Group discussion between the two rounds of survey

On 25th March, we provided a summary of the survey results for both groups in Excel format and to those who answered the survey. Apart from the individual answers elicited for each scenario, we also presented summary statistics for the entire group (max, min, median, average and standard deviation). We also made box-whisker graphs, showing which expert's estimate was an outlier for a given scenario. For each of the 20 scenarios, we plotted all estimates against a baseline of relative risk ratio 1. We randomly assigned each expert a code (from Expert 1 or E01 to E14 for G1, and E01 to E15 for G2) to keep their answers anonymous.

In addition to the summaries, we also invited experts to comment on the results by email if they wished. This started the group discussion process between the two survey rounds.

In the following two weeks, participants provided some general observations about the results but no specific comments were made. To stimulate discussion, on 10th April we sent each group an updated version of the summary results with better data visualization. More importantly, for each scenario, under each plot of the group's estimates we also asked a stimulating question specific to the group's answers for that scenario. For example, "Estimates fairly consistent in Scenario A1, with the risk of establishment generally seen as being up to 10 times greater in Botany Bay than in Geelong. Expert 12's response is interesting because of his/her absolute certainty that Geelong and Botany Bay are identical in terms of the risk of establishment (i.e. lower bound = upper bound= best estimate). Such certainty is unusual in exercises of this type. (While the responses of Expert 11 look very similar to those from Expert 12, 11's bounds do differ slightly from the best estimate.) Could there be an as yet unlisted variable that might drive such certainty?" Apart from the scenario-specific stimulating questions, the project team also presented a preliminary analysis based on the answers to the first survey round. The aim of this draft was to show the experts how their contributions would be used in the next stage of the project and to inform the group discussion.

During the week of 12th April, the project team held four teleconferences with experts who completed the first survey round. Separate teleconferences were held for Australian/New Zealand and overseas participants from each of G1 and G2. About 70% of the experts attended the 2nd teleconferences, aimed at further stimulating group discussion of the survey results. During the teleconferences, we introduced the process of group discussion, presented the preliminary modelling results and again spent most of the hour on group discussion. The discussion focussed on the scenario-specific results and answering the stimulating questions.

As it was impossible to discuss all 20 scenarios, we chose to cover a limited number of them thoroughly in the limited time we had. Experts were encouraged to continue the discussion virtually by adding their answers and thoughts about the unfinished scenarios online after the teleconferences. During the teleconference and virtual group discussion, the project team did not reveal the survey answers given by any individual but experts could reveal their identities if they wished.

The project team provided a general summary of the teleconferences which they shared with those who could not participate in the teleconferences. We also provided a scenario-specific summary underneath each stimulating question in the Excel file, which was then shared with all experts via Dropbox. We also created text-boxes for each scenario and invited the experts to write down their comments.

On 5th May, the project team also sent individualized stimulating questions to all 29 experts, asking each of them to answer a specific question related to his/her estimates when comparing them to the rest of the group. For example, "The question for you is related to Scenario M2: Most people in your group agreed that relative risk ratio was larger than 1, but your estimate was the highest and you seemed to be quite certain about it. How would you convince the rest in your group that Cairns is much riskier than they thought? "

About 79% of the experts in G1 and 60% in G2 jointed the virtual discussion online.

2nd survey round

On 11th May, the project team asked the experts to start updating their estimates from the first round based on the new information they had gathered from the group discussion. Once again the experts had ten business days to finish their survey but we gave a one-week extension to those who were too busy to submit their answers on time.

By the end of May, five experts from G1 (36%) and seven experts from G2 (47%) submitted their second round of survey answers. On average, a bit over 40% of the survey participants updated their estimates.

2.5 Scenario scoring detail

Summary of variables

Each scenario contains a set of variables with different values available for each variable. The variables to be included were chosen on the basis of deliberations at a one-day workshop (Section 5.5). These variables are summarised in the Table 1.

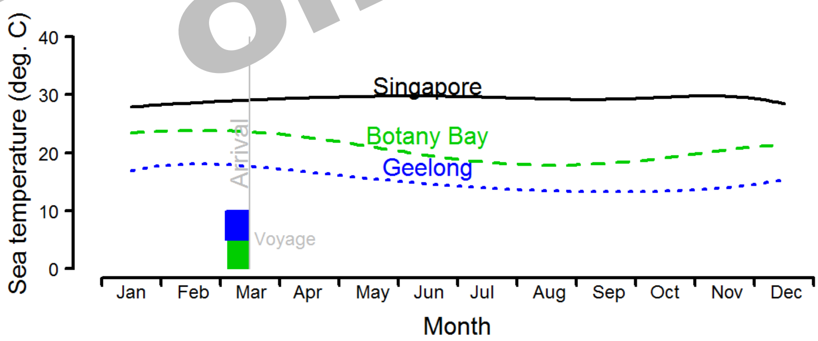
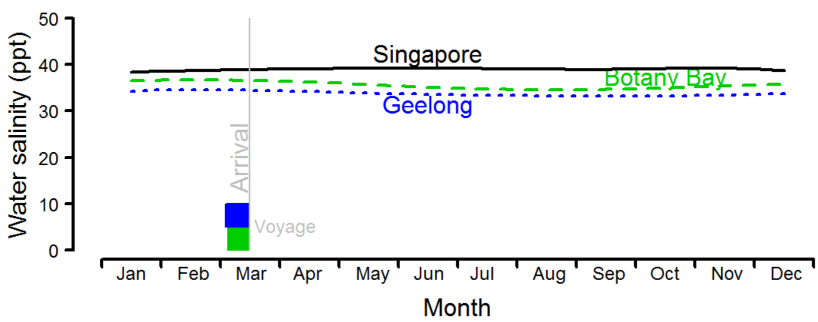
Table 1. Summary of variables used in scenario development.

Variable	Possible values
Port of last call	International ports, weighted for selection in sample based on current shipping patterns
Recipient port	Australian 1st ports of call
Harbour type	Coastal (natural, enclosed, breakwater) River (natural/basin/tide gates) Open roadstead
Mean tidal range	Continuous (m)
Mean monthly water temperature	Continuous (^{0}C)
Mean monthly water salinity profile	Continuous (ppt)
Donor port connectivity	Continuous (ship arrivals per day)
Age of vessel antifouling	New, 50% or Service Life, Out of Service Life
Voyage duration	Continuous (days)
Duration of stay in recipient port	Continuous (days)

An example scenario is provided in Table 2

Further details of the instructions to the experts on how to assess the scenarios, and complete details of scenarios is provided in the Supplementary Material.

Table 2. Example scenario given to experts.

Variable	Port of origin	Reference scenario	Comparison scenario
Port	Singapore	Geelong	Botany Bay
Latitude (Dec. Degrees)	1.3	-38.1	-34.0
Harbour type	Coastal natural	Coastal natural	Coastal natural
Voyage duration		11.5 days	12.5 days
Arrival month	March		
Antifouling age		50% of service life	50% of service life
Berthing duration		2 days	2 days
Connectivity (arrivals)	42,000 yr^{-1}		
Mean tidal range	3m	1m	1m
Monthly mean water temp. and range (^{0}C)	29.2 (27.9–30.2)	15.2 (13.2–18.2)	20.8 (17.8–23.8)
Mean water temp. w.r.t. port of origin		–14	–8.4
Monthly mean water temperature profiles			
Monthly mean water salinity and range (ppt)	39.0 (38.4–39.4)	33.8 (33.2–34.7)	35.6 (34.6–36.7)
Mean water salinity w.r.t. port of origin		–5.2	–3.4
Monthly mean salinity profile			

2.6 Analysis

2.6.1 Model fitting

The expert-derived estimate of relative risk (transformed to log10 scale) were modelled using a standard regression model assuming a Gaussian response which was checked using standard regression diagnostics shown in the Appendices (Section 5.3). As outlined in Section 2.2 the error variance within experts was allowed to vary across experts to allow for different levels of expertise.

The model was fitted in the software environment R version 3.0.3 using the gls() function within the mixed modelling package nlme (Pinheiro et al. 2015). The model relating the logarithm of the relative risk (RR) to predictor variables had the form:

$y_{ij}^{*} = (X_{11j} - X_{12j})\beta_1 + \cdots + (X_{k1j} - X_{k2j})\beta_k + b_{ij}$ (Equation 1)

Subscript i indexes the experts, j indexes the scenarios and k indexed the covariates associated with the journey. Factor variables where incorporated using treatment contrasts in R. The b_{ij} are assumed to be distributed as $N(0, \sigma_i^2)$, ie each expert has an uncertainty term with associated variance, i.e. the response is assumed to be heteroscedastic. The parameter $\sigma_i^2$is estimated during model fitting.

The R code to develop the model matrix and fit the model is given in Section 5.2.

2.6.2 Development of covariates and model selection

We were mindful not to over-parameterise the model given the nested nature (scenarios within experts) of the analysis. Because of this we chose to use the absolute difference in mean temperature and salinity as measures of environmental similarity. There were also a number of contrasts relating to port environment that were not estimable and these were removed. This arose partially through the restriction of using real ports in our scenarios. As some of these possibilities may occur in operational use of this model we collapsed the port environment variable into either “River” or “Coastal” and refitted the model. We also only investigated first order effects, and did not explore interactions between variables, although this is clearly an avenue for further analysis.

2.6.3 Model selection

We fitted a model including all terms that were selected in the initial workshop at the commencement of the project. We present this model as an un-biased estimate of these effects. We also used backwards elimination to derive a reduced model that only included significant effects. While this may introduce certain biases in parameter estimates it will also lead to a more credible (i.e. larger) prediction error.

2.6.4 Expert uncertainty

An experimental component of the elicitation was to get experts to provide ranges for their beliefs about relative risks. These data are complex and will need additional work to analyse. The relative nature of the analysis and the lack of natural ordering of the destination and comparison ports mean that the interpretation of highest and lowest is context specific. At this stage we provide a simple graphical summary of the range of responses to aid interpretation of results and a coarse signal to noise analysis to give perspective about the extent of the uncertainty.

2.6.5 Application of model to risk prioritization

The fitted model can be applied directly to risk prioritization tasks. As the model calculates the relative risk of any two journeys, it can also be applied to a set of journeys using one as a reference state to order the risk.

3 Results

3.1 Survey response

As expected, there was variation between experts in their estimates of risk for the scenarios. Despite this, the aggregated expert responses suggest there is information within the scenarios leading experts to assign different risks to different scenarios — there appears to be a degree of consensus emerging when comparing the results for each scenario (Figure 2).

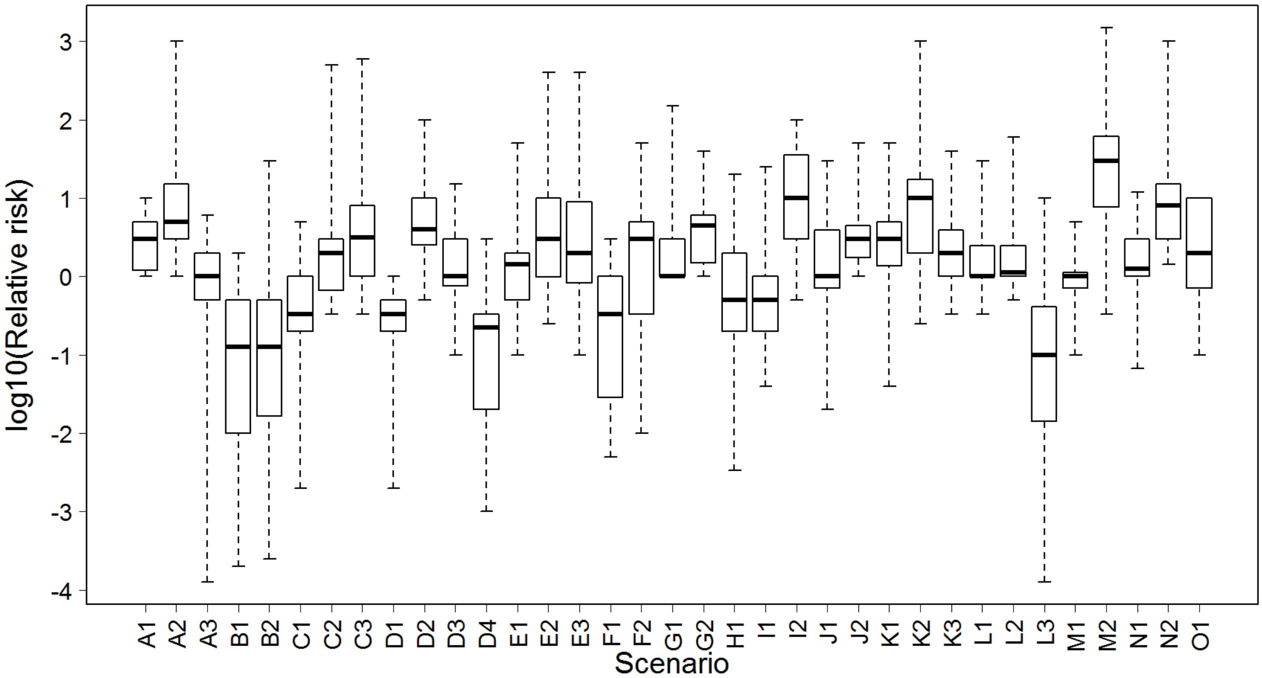


Figure 2. Box and whiskers plot for the aggregated responses of expert estimates of relative risk (on a logarithmic scale) for the various scenarios. All (n=29) experts assessed scenarios A1—B2. Scenarios C1—H1 were assessed by 14 experts and scenarios I1—O1 by 15 experts. Horizontal line is median expert response, boxes show the inter quartile range, and whiskers the response range.

Individual experts varied considerably in their range of estimates (Figure 3). For example, compare G1E09 (the 9th expert from group1) with G1E13 (the 13th expert from group 1). Note, however, that a major component of this is variation of the characteristics of the survey, i.e. this is signal and not noise.

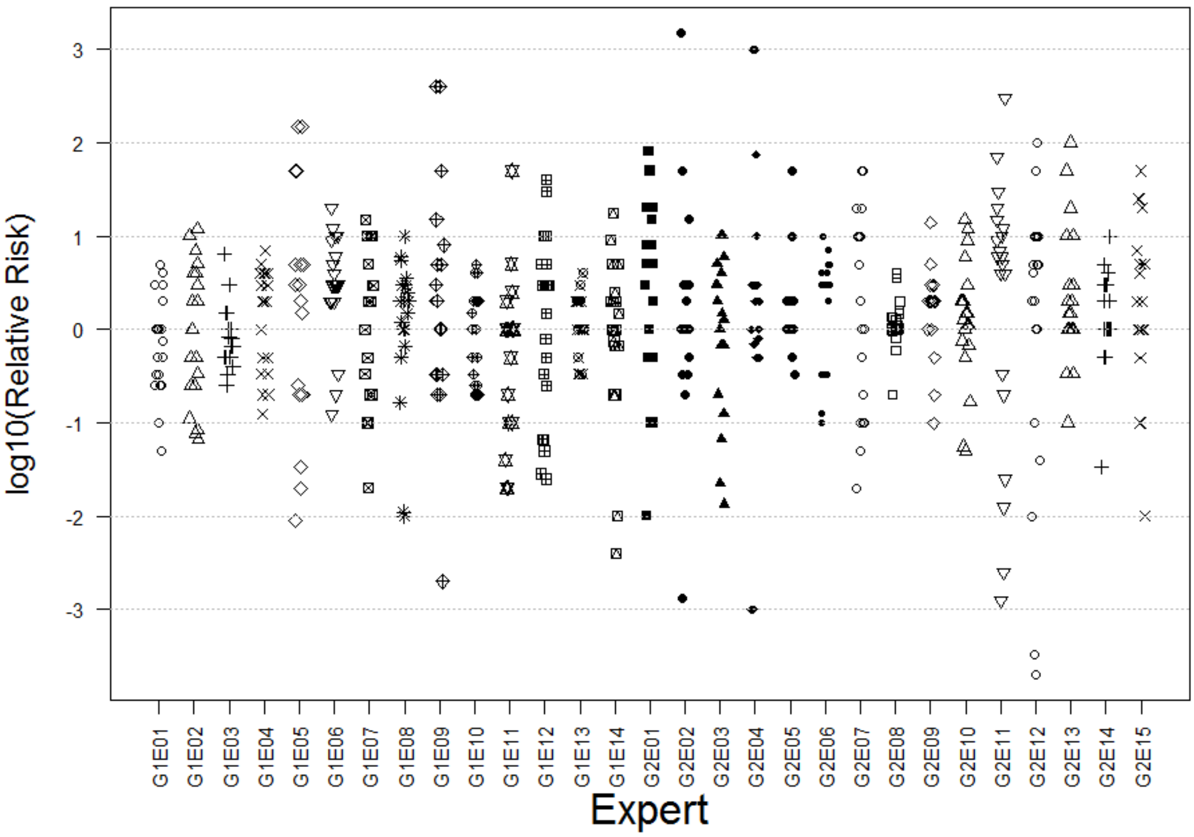


Figure 3. The "best" responses of experts on a logarithmic scale (base 10) for each scenario, grouped by expert. Note that the two groups (prefixes "G1" and "G2" on the horizontal axis labels) assessed a partially different set of scenarios.

3.2 "Signal to noise" based on experts uncertainty ranges

We can get a sense of the "signal to noise" within the expert's judgements by assessing the ratio of the variation in the "best" estimate, to the variation in the difference between either their best estimate and their 95% "upper" or their "lower" bound, respectively. If, for example, the expert derived "signal" was similar to the noise on the upside or downside we would expect this ratio to be approximately one. Specifically, we calculated:

$$Ratio = \frac{var(\log(x_{best}))}{var(\log(x_{best}) - \log(x_{95}))}$$

where $x_{95}$is either the upper or lower bound for the best estimate (x_{best}). When analysing data on a logarithmic scale, these ratios are *c.* 12 for the upper bound and 9 for the lower bound when averaged across all experts. This indicates that expert's belief "signal" is considerably larger than their estimated uncertainty in either direction. This is illustrated graphically in Figure 4.

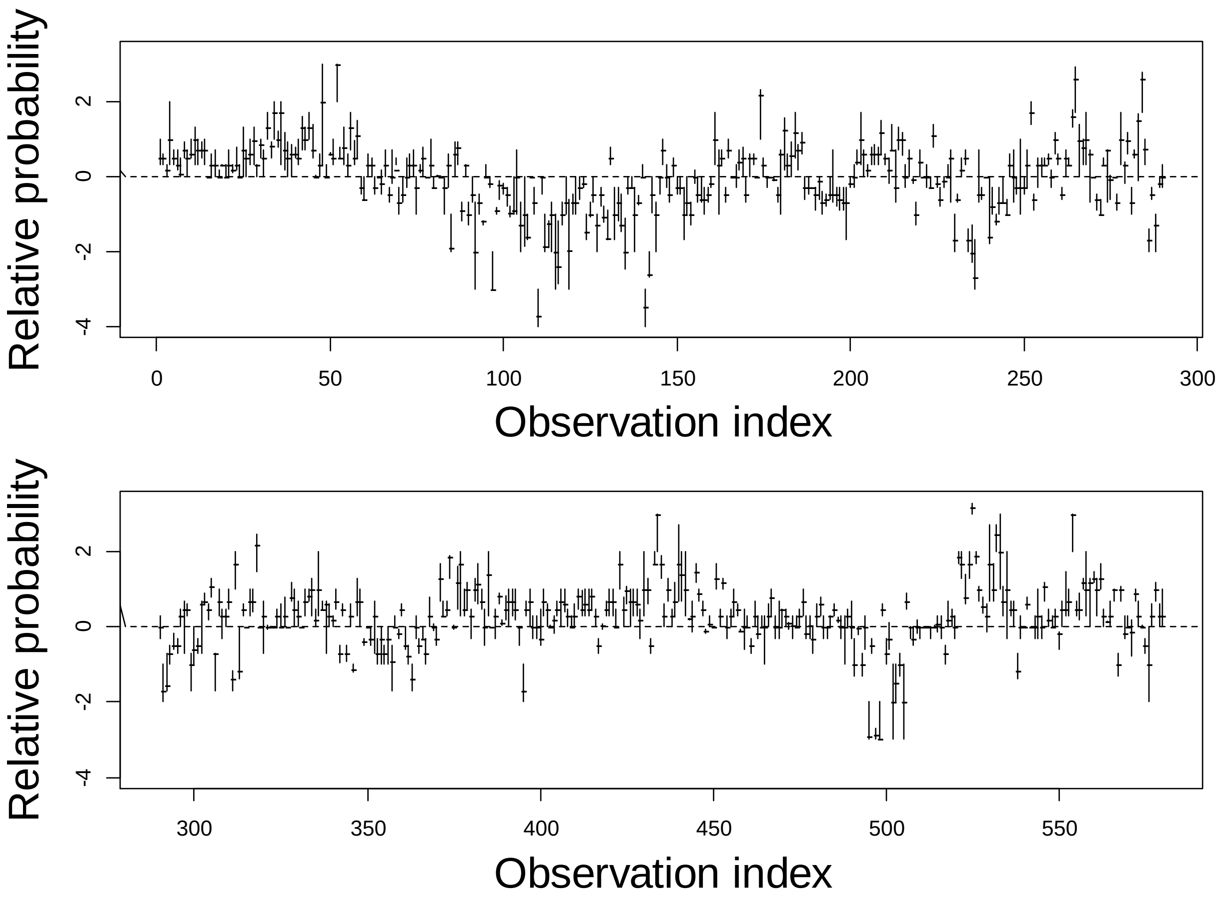


Figure 4. All experts estimates of relative risk on a logarithmic scale (base 10). Horizontal dashes show best estimates and vertical bars show upper and lower estimates. The horizontal dashed line is provided for reference only.

3.3 Development of model of expert response

Plotting the best estimates against the model fitted values reveals that despite the variation between experts, there is still a strong signal overall (Figure 5). One measure of the performance of the model emulator is its ability to capture the mean response of experts. On this measure, the fitted model emulator performs well in capturing variation (Pearson correlation coefficient between fitted and expert mean, r=0.92) in the mean expert relative risk across the scenarios considered (Figure 6). The mean relative risk across the scenarios spans two orders of magnitude, and the model emulator captures this variation reasonably well. This by no means demonstrates that the model emulator is correctly estimating relative risk, but rather, it demonstrates that the emulator has captured a significant component of the variation between scenarios. This arguably demonstrates there is an underlying model-based explanation for the expert responses. Whether this is a good model difficult to ascertain, although the direction of parameter estimates provides a preliminary check (see below).

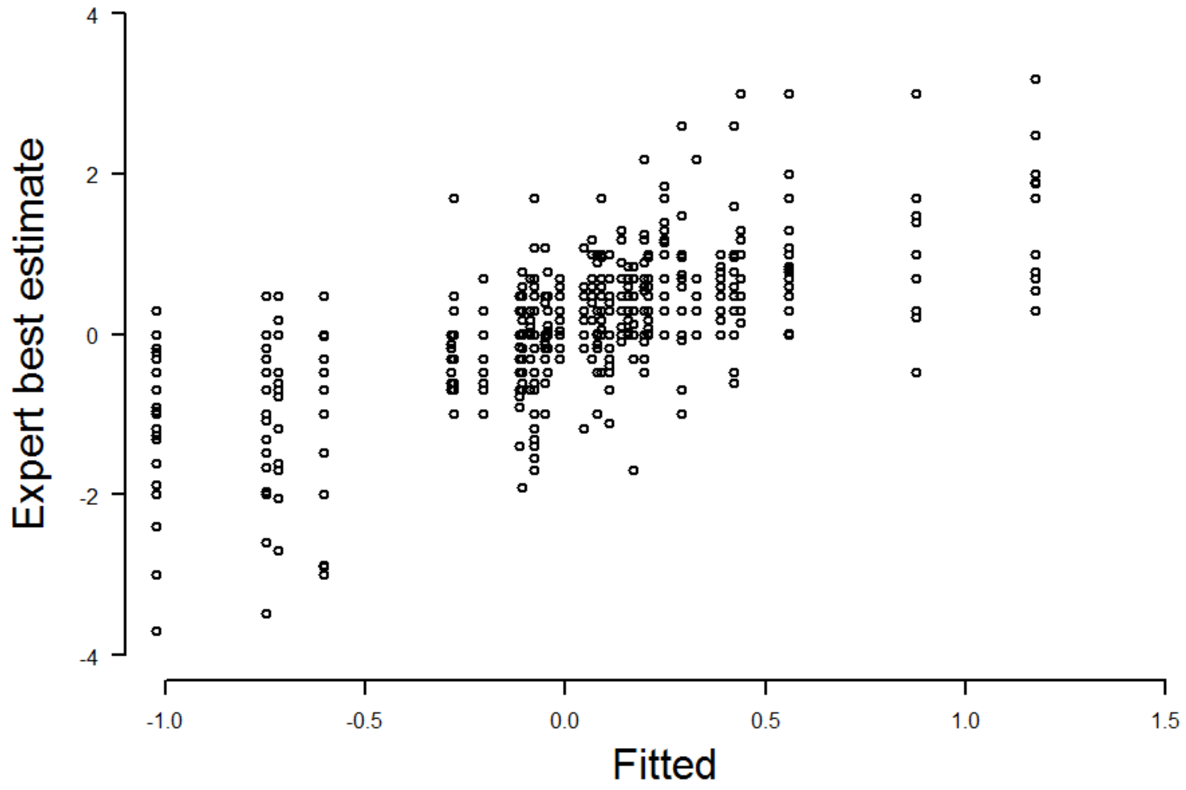


Figure 5. Best estimates of experts as a function of the fitted model values. Both axes are on a logarithmic (base 10) scale.

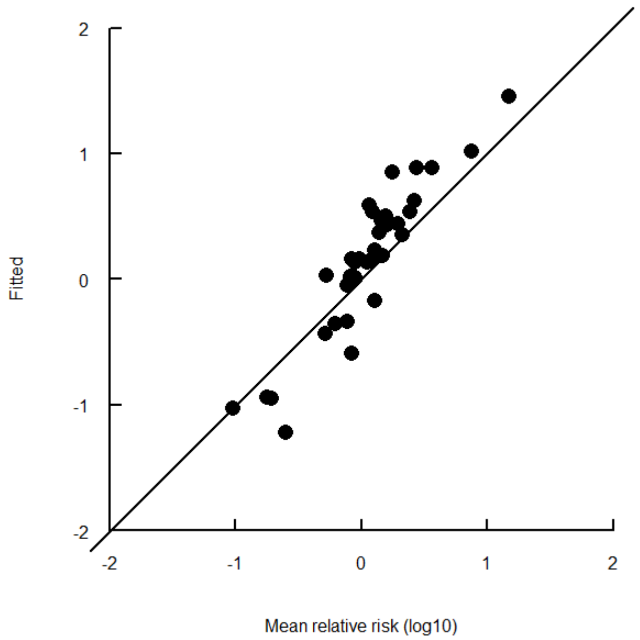


Figure 6. Relationship between fitted model response and mean of all experts (n=29) for each of the 35 scenarios assessed. The solid line represents equivalence between the fitted model response and the mean response.

Of methodological interest, there was only a weak (albeit significant) relationship (r^2=0.09) between the self-assessed range of possible values for the relative risk, and the model residual (Figure 7).

Figure 7. Relationship between absolute value of the model residual and the uncertainty interval provided by experts are part of the elicitation of their best estimate. Both measure are on a logarithmic scale (base 10).

3.4 Parameter estimates

A summary of the fitted model coefficients, their precision and significance is provided in Table 3. A key point to note is these coefficients relate to the logarithm (base 10) of the relative risk of establishment. The sign of significant coefficients was consistent with current beliefs regarding biophysical factors (Table 3**Error! Reference source not found.**). That is, being translocated to a ess enclosed port environment in relative terms lessens the risk of establishment, as does moving to a port with a greater difference in mean temperature and/or salinity. Also, decreasing relative effectiveness of antifouling was associated with large and positive increases in the relative risk of establishment, which was further exacerbated by the berthing duration.

The multiplicative effect of each explanatory variable on the relative risk of establishment, compared to the reference level for that variable, is obtained by raising the parameter estimate to the power of 10 (Table 4). For example, compared to a new anti-fouling coat, having anti-fouling at 50% of its service life increased the risk by a factor of *c.* 2 (=$10^{0.284}$) (Table 4). Similarly, compared to a new anti-fouling coat, having anti-fouling at 100% of its service life increased the risk by a factor of *c.* 4.

Moving from an enclosed port (e.g. river) to a less enclosed port (e.g. coastal or open roadstead) significantly lowered the estimated risk by 54% compared to moving between similar ports, although this effect was not consistent when moving in the opposite direction. In addition, discussions with experts revealed some believed that voyages travelling between freshwater ports would have lower risks arising from the journey duration of time spent in saltwater, although this belief was not unanimous.

The environmental similarity of ports was inferred to be an important cue by experts. Each degree increase in the absolute difference in mean sea surface temperature from the port of origin changed the relative risk by 0.923 (*c.* 8% decrease). Similarly, each unit change in the absolute difference in mean water salinity from the port of origin altered the relative risk by 0.97 (a 3% decrease).

Each additional day of berthing increased the relative risk by 1.033 (a 3.3% increase).

There was no evidence that experts consistently factored in similarity in tidal range, or voyage duration into their estimates (see Table 5 in Appendix).

Table 3. Summary of model parameter estimates (on the fitted logarithmic scale) from final fitted expert emulator.

Variable level	Value	Std. Error	t-value	p-value	Sig.
Port type change					
Same	0 (aliased)				
More enclosed	-0.08679	0.047882	-1.81252	0.070429	n.s.
Less enclosed	-0.33822	0.092462	-3.65798	0.000278	***
Absolute difference in mean temperature	-0.03484	0.006642	-5.24522	2.20E-07	***
Absolute difference in mean salinity	-0.01316	0.002501	-5.26362	2.00E-07	***
Anti-fouling					
New	0 (aliased)				
50% of service life	0.283862	0.068718	4.130828	4.15E-05	***
100% of service life	0.642734	0.058611	10.96603	1.57E-25	***
Berthing duration	0.013998	0.001532	9.138205	1.09E-18	***

Table 4. Summary of variables retained in the final model and their effects on the relative risk of establishment compared to the reference level.

Variable level	Effect size	Reference level
Port less enclosed	54% decrease	Same level of enclosure
Absolute difference in mean temperature	7.7% decrease	Per extra ^{0}C difference
Absolute difference in mean salinity	3.0% decrease	Per extra ppt difference
Anti-fouling 50% of service life	92.2% increase	New antifouling
Anti-fouling 50% of service life	4.4-fold increase	New antifouling
Berthing duration	3.3% increase	Per extra day of berthing

There remains considerable variation in the spread of residuals for different experts (Figure 8). This reflects fundamentally different views about the range of relative risks. Note that this is factored into the analysis and the inferences via the use of the heteroscedacstic regression model.

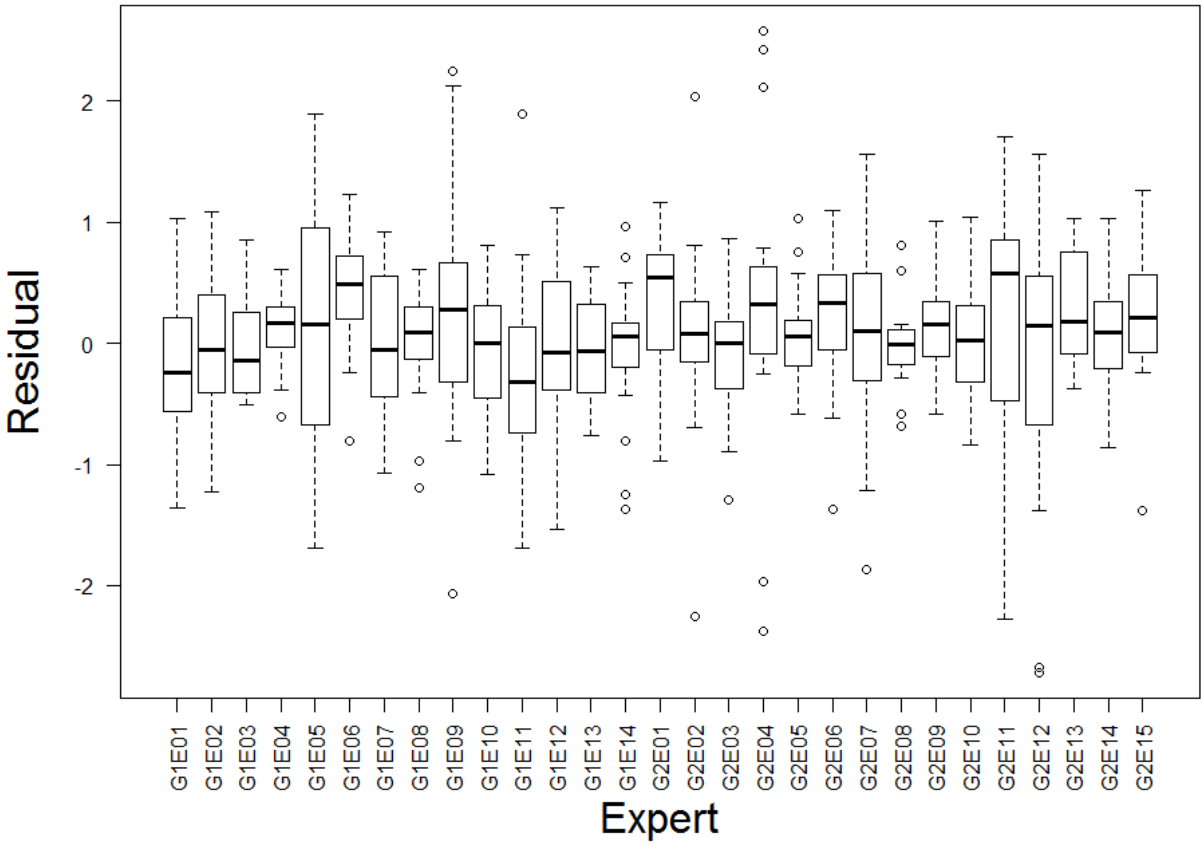


Figure 8. Boxplots of model residuals by expert. Note the heteroscedasticity across the experts.

4 Discussion

The results presented in this report represent a novel approach to estimating the risk posed by hull fouling on vessels for the establishment of marine pests. This has been a difficult problem. Introduction events are rarely observed so it is not possible to construct direct estimates of the probability of introduction. Physiological models and data about the biogeography of species are poorly developed so the use of process based models has not been feasible. While it is important to not overestimate the ability of experts to understand complex phenomena the study still provides a transparent and systematic summary of their views which is a significant advance on innuendo and anecdote.

Significant complexity remains. Hull fouling assemblages will often depend on a number of previous ports. There are potential for complex interaction between variables that are not factored into this analysis. But these challenges will exist no matter what technique is applied. In any application of these results these issues will need to be considered and managed as appropriate. For example, the model may be applied to a number of the vessels previous ports to develop an aggregate estimate of relative risk.

The model as specified is well suited to risk ranking applications. It cannot be applied to determine absolute risk and therefore, given a threshold, whether a journeys risk is acceptable or not. If there was a requirement to do this, further analysis using the marginal rate of introductions would be needed. Thus the model as it stands could directly support compliance targeting. To apply it to primary estimation of risk would require both additional analysis as well as work to integrate it logically with the MGRA tool.

Some aspects of the results where surprising. For example, the lack of a discernible effect for the influence of voyage duration on the risk of establishment. Overall, however, there was a surprising level of consistency in views. The estimated impact of anti-fouling on relative risk was large, as expected based on previous discussions with experts.

If an analysis based on the use of relative risks was to be repeated there have been a number of learnings. Consideration of the design space of the study was complicated by the journey risk depending on a comparison between ports and the elicitation depending on a comparison of journeys. This introduced subtleties in the design process that were not easily identified. In future we recommend using simulated data to assess the estimability of all parameter contrasts. It would also have been more natural to consider scenarios with two source ports and a single destination, as this configuration is how the model will mostly be used. It does not technically affect the results of this study but would be easier to communicate and explain.

5 References

Apte, S., Holland, B. S., Godwin, L. S., and Gardner, J. P. A. (2000). Jumping ship: a stepping stone event mediating transfer of a non-indigenous species via a potentially unsuitable environment. *Biological Invasions* **2,** 75-79.

Aspinall, W. (2010). A route to more tractable expert advice. *Nature* **463,** 294-295.

Barry, S. C. (2015). Scenario Calibration: the construction of well calibrated expert based scoring systems. **in review**.

Barry, S. C., Hayes, K. R., Hewitt, C. L., Behrens, H. L., Dragsund, E., and Bakke, S. M. (2008). Ballast water risk assessment: principles, processes, and methods. *Ices Journal of Marine Science* **65,** 121-131. doi: 10.1093/icesjms/fsn004

Burgman, M., Carr, A., Godden, L., Gregory, R., McBride, M., Flander, L., and Maguire, L. (2011a). Redefining expertise and improving ecological judgment. *Conservation Letters* **4,** 81-87. doi: 10.1111/j.1755-263X.2011.00165.x

Burgman, M., Franklin, J., Hayes, K. R., Hosack, G. R., Peters, G. W., and Sisson, S. A. (2012). Modeling Extreme Risks in Ecology. *Risk Analysis* **32,** 1956-1966.

Burgman, M. A. (2005). *Risks and decisions for conservation and environmental management*. (Cambridge University Press.

Burgman, M. A., McBride, M., Ashton, R., Speirs-Bridge, A., Flander, L., Wintle, B., Fidler, F., Rumpff, L., and Twardy, C. (2011b). Expert Status and Performance. *PLoS ONE* **6,** e22998.

Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond, R. E. A., Baillie, J. E. M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K. E., Carr, G. M., Chanson, J., Chenery, A. M., Csirke, J., Davidson, N. C., Dentener, F., Foster, M., Galli, A., Galloway, J. N., Genovesi, P., Gregory, R. D., Hockings, M., Kapos, V., Lamarque, J.-F., Leverington, F., Loh, J., McGeoch, M. A., McRae, L., Minasyan, A., Morcillo, M. H., Oldfield, T. E. E., Pauly, D., Quader, S., Revenga, C., Sauer, J. R., Skolnik, B., Spear, D., Stanwell-Smith, D., Stuart, S. N., Symes, A., Tierney, M., Tyrrell, T. D., Vié, J.-C., and Watson, R. (2010). Global Biodiversity: Indicators of Recent Declines. *Science* **328,** 1164-1168. doi: 10.1126/science.1187512

Capinha, C., Essl, F., Seebens, H., Moser, D., and Pereira, H. M. (2015). The dispersal of alien species redefines biogeography in the Anthropocene. *Science* **348,** 1248-1251. doi: 10.1126/science.aaa8913

Cooke, R. M. (1991). *Experts in uncertainty: opinion and subjective probability in science*. (Oxford University Press: New York).

Coutts, A. D. M., Moore, K. M., and Hewitt, C. L. (2003). Ships' sea-chests: an overlooked transfer mechanism for non-indigenous marine species? *Marine Pollution Bulletin* **46,** 1504–1515.

Crombie, J., Knight, E., and Barry, S. (2008). Marine Pest Incursions--a Tool to Predict the Cost of Eradication Based on Expert Assessments. (Australian Government, Bureau of Rural Sciences: Canberra).

Dafforn, K. A., Lewis, J. A., and Johnston, E. L. (2011). Antifouling strategies: History and regulation, ecological impacts and mitigation. *Marine Pollution Bulletin* **62,** 453-465.

Darbyshire, R., and Caley, P. (2009). Identifying significant range extensions of invasive marine pests – CCIMPE Range Decision Guidelines Project. (Bureau of Rural Sciences: Canberrra).

Davidson, A. D., Campbell, M. L., and Hewitt, C. L. (2013). The role of uncertainty and subjective influences on consequence assessment by aquatic biosecurity experts. *Journal of Environmental Management* **127,** 103-113. doi: http://dx.doi.org/10.1016/j.jenvman.2013.03.043

Doria, M. d. F., Boyd, E., Tompkins, E. L., and Adger, W. N. (2009). Using expert elicitation to define successful adaptation to climate change. *Environmental Science & Policy* **12,** 810-819. doi: 10.1016/j.envsci.2009.04.001

Drake, J. M., and Lodge, D. M. (2004). Global hot spots of biological invasions: evaluating options for ballast-water management. *Proceedings of the Royal Society B-Biological Sciences* **271,** 575-580. doi: 10.1098/rspb.2003.2629

Drake, J. M., and Lodge, D. M. (2007). Hull fouling is a risk factor for intercontinental species exchange in aquatic ecosystems. *Aquatic Invasions* **2,** 121-131.

Drescher, M., Perera, A. H., Johnson, C. J., Buse, L. J., Drew, C. A., and Burgman, M. A. (2013). Toward rigorous use of expert knowledge in ecological research. *Ecosphere* **4**. doi: 10.1890/es12-00415.1

Floerl, O., Pool, T. K., and Inglis, G. J. (2004). Positive interactions between nonidigenous species facilitate transport by human vectors. *Ecological Applications* **14,** 1724-1736.

Franklin, J., Sisson, S. A., Burgman, M. A., and Martin, J. K. (2008). Evaluating extreme risks in invasion ecology: learning from banking compliance. *Diversity and Distributions* **14,** 581-591.

Gosling, J. P., Hart, A., Mouat, D. C., Sabirovic, M., Scanlan, S., and Simmons, A. (2012). Quantifying Experts' Uncertainty About the Future Cost of Exotic Diseases. *Risk Analysis* **32,** 881-893. doi: 10.1111/j.1539-6924.2011.01704.x

Hayes, K. R., and Hewitt, C. L. (2001). Ballast-Water Risk Assessment - Volume 2. (CSIRO Division of Marine Research, Hobart, Australia:

Hewitt[1], C., Campbell[1], M., Coutts, A., Dahlstrom[1], A., Shields, D., and Valentine, J. (2011). Species biofouling risk assessment. (Commonwealth Department of Agriculture, Fisheries & Forestry: Canberra).

IPCC (2001). Chapter 6: Quantifying uncertainties in practice. In 'IPPC Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories'.

James, A., Choy, S. L., and Mengersen, K. (2010). Elicitator: An expert elicitation tool for regression in ecology. *Environmental Modelling & Software* **25,** 129-145. doi: 10.1016/j.envsoft.2009.07.003

Jarrad, F. C., Barrett, S., Murray, J., Stoklosa, R., Whittle, P., and Mengersen, K. (2011). Ecological aspects of biosecurity surveillance design for the detection of multiple invasive animal species. *Biological Invasions* **13,** 803-818. doi: 10.1007/s10530-010-9870-0

Knight, E., Barry, S., Summerson, R., Cameron, S., and Darbyshire, R. (2007). Designated exchange areas project – Providing informed decisions on the discharge of ballast water in Australia (Phase 2). . (Bureau of Rural Sciences: Canberra).

Knol, A. B., Slottje, P., van der Sluijs, J. P., and Lebret, E. (2010). The use of expert elicitation in environmental health impact assessment: a seven step procedure. *Environmental Health* **9**. doi: 10.1186/1476-069x-9-19

Kuhnert, P. M., Martin, T. G., and Griffiths, S. P. (2010). A guide to eliciting and using expert knowledge in Bayesian ecological models. *Ecology Letters* **13,** 900-914.

Kynn, M. (2008). The 'heuristics and biases' bias in expert elicitation. *Journal of the Royal Statistical Society Series a-Statistics in Society* **171,** 239-264.

Linstone, H. A., and Turoff, M. (1975). The Delphi method -- techniques and applications. (ed^(eds. (Addison-Wesley: Boston.)

Liu, S., Sheppard, A., Kriticos, D., Cook, D., and Liu, S. (2011). Incorporating uncertainty and social values in managing invasive alien species: a deliberative multi-criteria evaluation approach. *Biological Invasions* **13,** 2323-2337. doi: 10.1007/s10530-011-0045-4

Low Choy, S., O'Leary, R., and Mengersen, K. (2009). Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. *Ecology* **90,** 265-277. doi: 10.1890/07-1886.1

Mack, R. N., Simberloff, D., Lonsdale, W. M., Evans, H., Clout, M., and Bazzaz, F. A. (2000). Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications* **10,** 689-710.

Martin, T. G., Burgman, M. A., Fidler, F., Kuhnert, P. M., Low-Choy, S., McBride, M., and Mengersen, K. (2012). Eliciting Expert Knowledge in Conservation Science. *Conservation Biology* **26,** 29-38. doi: 10.1111/j.1523-1739.2011.01806.x

McBride, M. F., Fidler, F., and Burgman, M. A. (2012a). Evaluating the accuracy and calibration of expert predictions under uncertainty: predicting the outcomes of ecological research. *Diversity and Distributions* **18,** 782-794. doi: 10.1111/j.1472-4642.2012.00884.x

McBride, M. F., Garnett, S. T., Szabo, J. K., Burbidge, A. H., Butchart, S. H. M., Christidis, L., Dutson, G., Ford, H. A., Loyn, R. H., Watson, D. M., and Burgman, M. A. (2012b). Structured elicitation of expert judgments for threatened species assessment: a case study on a continental scale using email. *Methods in Ecology and Evolution* **3,** 906-920. doi: 10.1111/j.2041-210X.2012.00221.x

Molnar, J. L., Gamboa, R. L., Revenga, C., and Spalding, M. D. (2008). Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment* **6,** 485-492. doi: 10.1890/070064

Murray, C. C., Pakhomov, E. A., and Therriault, T. W. (2011). Recreational boating: a large unregulated vector transporting marine invasive species. *Diversity and Distributions* **17,** 1161-1172. doi: 10.1111/j.1472-4642.2011.00798.x

Murray, C. C., Therriault, T. W., and Martone, P. T. (2012a). Adapted for invasion? Comparing attachment, drag and dislodgment of native and nonindigenous hull fouling species. *Biological Invasions* **14,** 1651-1663. doi: 10.1007/s10530-012-0178-0

Murray, J. V., Stokes, K. E., and van Klinken, R. D. (2012b). Predicting the potential distribution of a riparian invasive plant: the effects of changing climate, flood regimes and land-use patterns. *Global Change Biology* **18,** 1738-1753. doi: 10.1111/j.1365-2486.2011.02621.x

O'Hagan, A., Buck, C. E., Daneshkha, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., and Rakow, T. (2006). *Uncertain judements: Eliciting experts' probabilities*. (Wiley: UK).

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2015). nlme: Linear and Nonlinear Mixed Effects Models. http://CRAN.R-project.org/package=nlme

Piola, R. F., Dafforn, K. A., and Johnston, E. L. (2009). The influence of antifouling practices on marine invasions. *Biofouling* **25,** 633-644.

Pulliam, H. R. (2000). On the relationship between niche and distribution. *Ecology Letters* **3,** 349-361. doi: 10.1046/j.1461-0248.2000.00143.x

Reed, M. S. (2008). Stakeholder participation for environmental management: A literature review. *Biological Conservation* **141,** 2417-2431.

Ruiz, G. M., Fofonoff, P. W., Ashton, G., Minton, M. S., and Miller, A. W. (2013). Geographic variation in marine invasions among large estuaries: effects of ships and time. *Ecological Applications* **23,** 311-320.

Seebens, H., Gastner, M. T., and Blasius, B. (2013). The risk of marine bioinvasion caused by global shipping. *Ecology Letters* **16,** 782-790. doi: 10.1111/ele.12111

Speirs-Bridge, A., Fidler, F., McBride, M., Flander, L., Cumming, G., and Burgman, M. (2010). Reducing Overconfidence in the Interval Judgments of Experts. *Risk Analysis* **30,** 512-523. doi: 10.1111/j.1539-6924.2009.01337.x

Sylvester, F., Kalaci, O., Leung, B., Lacoursiere-Roussel, A., Murray, C. C., Choi, F. M., Bravo, M. A., Therriault, T. W., and MacIsaac, H. J. (2011). Hull fouling as an invasion vector: can simple models explain a complex problem? *Journal of Applied Ecology* **48,** 415-423. doi: 10.1111/j.1365-2664.2011.01957.x

Thuiller, W., Richardson, D. M., Pysek, P., Midgley, G. F., Hughes, G. O., and Rouget, M. (2005). Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology* **11,** 2234-2250. doi: 10.1111/j.1365-2486.2005.01018.x

U.S. Enviornmental Protection Agency (2011). Expert elicitation task force white paper. (Prepared for the Science and Technology Policy council:

Appendix A

5.1 Statistics of the estimation of relative risk

Assume that the probability *p* of an establishment occurring from a vessel movement follows a logistic model:

$$\log\left(\frac{p_i}{1-p_i}\right) = \alpha + X_{1i}\beta_1 + X_{2i}\beta_2 + X_{3i}\beta_3 + \cdots + X_{ki}\beta_k$$

where the X's are functions of the attributes of the source and destination ports. If $p_1$is the risk from the source port to destination 1 and $p_2$is the risk from the source port to destination 2 then:

$$\log\left(\frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_p}}\right) \approx \log\left(\frac{p_1}{p_2}\right)$$

given that p is close to zero. We elicit $\frac{p_1}{p_2}$. Simple algebra gives that

$$\log\left(\frac{p_1}{p_2}\right) = \alpha + X_{11}\beta_1 + X_{21}\beta_2 + X_{31}\beta_3 + \cdots + X_{ki}\beta_k - (\alpha + X_{12}\beta_1 + X_{22}\beta_2 + X_{32}\beta_3 + \cdots + X_{k2}\beta_k)$$

Note that common terms, such as the intercept cancel out. This will occur for all terms that are common between the journeys such as simple attributes of the source port and vessel characteristics.

5.2 R Code

```
#########################################################################################
# Filename: Analysis.R
# Author:     Simon Barry & Peter Caley
########################################################################################
BASE.PATH <- "//nexus/Projects/Biosecurity/CEBRA_BIOFOULING"
source(file.path(BASE.PATH,"code/DataPrep_v01.R"))


library(nlme)


setup.frame <- function(x, levels=AllPortChanges)
{
  x$AF <- factor(x$AF)     # Antifouling (New, 50% of working life, 100% of working life)
  x$AF <- C(x$AF, treatment)
  x$PortChangeDirection <-
factor(x$PortChangeDirection,levels=c("Same","MoreEnclosed","LessEnclosed"))
  x
}
```

```
# Set up data frames
f1 <- setup.frame(dat.Ref)
f2 <- setup.frame(dat.Comp)

# Generate model formula
# Full model
form <- formula(~PortChangeDirection + abs(SSTAvgDiff) + abs(SalAvgDiff) + abs(TRDiff) +
Voyage + AF + Berth)

# Full model minus Voyage
form <- formula(~PortChangeDirection + abs(SSTAvgDiff) + abs(SalAvgDiff) + abs(TRDiff)  +
AF + Berth)         # 2 Port types - River, Coastal

# Full model minus Voyage + TRDiff
form <- formula(~PortChangeDirection + abs(SSTAvgDiff) + abs(SalAvgDiff) + AF + Berth)
       # 2 Port types - River, Coastal

# Generate model matrix
m1 <- model.matrix(form, data=f1)
m2 <- model.matrix(form, data=f2)
model.mat <- m2 - m1

#########################################################################################
# Current best model using gls()
mod <- lm(log10(Best)~-1 + model.mat,data=dat) # identifying what's not estimable ...
summary(mod)
fr.mod <- model.mat[,!is.na(coef(mod))]        # Remove non-estimable

# Fit gls() model
mod1 <- gls(log10(Best) ~ -1 + fr.mod,weights=varIdent(form=~1 | Expert),data=dat)
summary(mod1)
plot(mod1)
qqnorm(resid(mod1))
qqline(resid(mod1)) # still heavy tailed
abline(coef=c(0,1))

# Plot of residuals vs. fitted values
win.graph(h=7,w=10)
par(lwd=2,cex.lab=2,mar=c(6,6,2,2))
plot(mod1$fitted,residuals(mod1,type="normalized"),ylab="Standardized
residual",xlab="Fitted value")
abline(h=0)

# Plot of Expert estimates vs. fitted values
win.graph(h=7,w=10)
```

```
par(lwd=2,cex.lab=2,mar=c(6,6,2,2))

plot(mod1$fitted,log10(dat$Best),axes=F, xlab="Fitted", ylab="Expert best
estimate",ylim=c(-4,4),xlim=c(-1.5,1.5))

axis(1,lwd=2)

axis(2,las=2,lwd=2)


# Add fitted values to df

dat$mod1.fit <- mod1$fit


# Write values to csv file

write.csv(summary(mod1)$tTable,file=file.path(BASE.PATH,"Scenarios/Analysis/ModelCoeffici
entsFinal.csv"))

write.csv(summary(mod1)$tTable,file=file.path(BASE.PATH,"Scenarios/Analysis/ModelCoeffici
entsFull.csv"))


# Calculated observed means over experts and fitted means over experts

scen.ftd.mean <- tapply(dat$mod1.fit,dat$Scenario, mean)

scen.best.mean <- tapply(dat$Best,dat$Scenario,function(x) mean(log10(x)))


#######################################################################################

# Generating plots below here

win.graph()

par(cex.axis=1.5,cex.lab=2.0,mgp=c(2,1,0),lwd=2)


par(pty='s')

plot(scen.ftd.mean,scen.best.mean,pch=16,cex=2,ylim=c(-2,2),xlim=c(-
2,2),axes=F,xlab="Mean relative risk (log10)",ylab="Fitted")

axis(1,tck=0.02,pos=-2,lwd=2)

axis(2,las=2,tck=0.02,pos=-2,lwd=2)

abline(coef=c(0,1),lwd=2)


# Check out correlation between expert mean and fitted value  [SB hates this :)]

cor(scen.best.mean,scen.ftd.mean)

summary(lm(scen.ftd.mean ~ scen.best.mean))


# Plot of residuals by expert

win.graph(h=7,w=10)

par(las=2,mar=c(6,6,2,2),cex.lab=2)

plot(factor(dat$Expert),resid(mod1),xlab="",ylab="Residual")

mtext("Expert",1,las=1,padj=3,cex=2)


#######################################################################################
```

5.3 Model fitting and diagnostics

A summary of the full model coefficients is given in Table 5.

Table 5. Summary of model coefficients for full model.

Factor	Value	Std.Error	t-value	p-value	Sig.
Port type change					
Same	0 (aliased)				
More enclosed	-0.0798533	0.054124	-1.47538	0.140663	n.s.
Less enclosed	-0.33211555	0.093825	-3.53974	0.000433	***
Absolute difference in mean temperature	-0.03329717	0.007212	-4.617	4.81E-06	***
Absolute difference in mean salinity	-0.01313461	0.002512	-5.22879	2.40E-07	***
Absolute difference in tidal range	0.013090839	0.020071	0.652211	0.514528	n.s.
Voyage duration	-0.0014062	0.009605	-0.1464	0.883655	n.s.
Anti-fouling					
New	0 (aliased)				
50% of service life	0.274240061	0.07186	3.816296	0.00015	***
100% of service life	0.644957644	0.058815	10.96586	1.60E-25	***
Berthing duration	0.014059248	0.001554	9.046981	2.29E-18	***

The model fitted well across the range of estimated risks, as evidenced by the plot of model residuals versus fitted value (Figure 9).

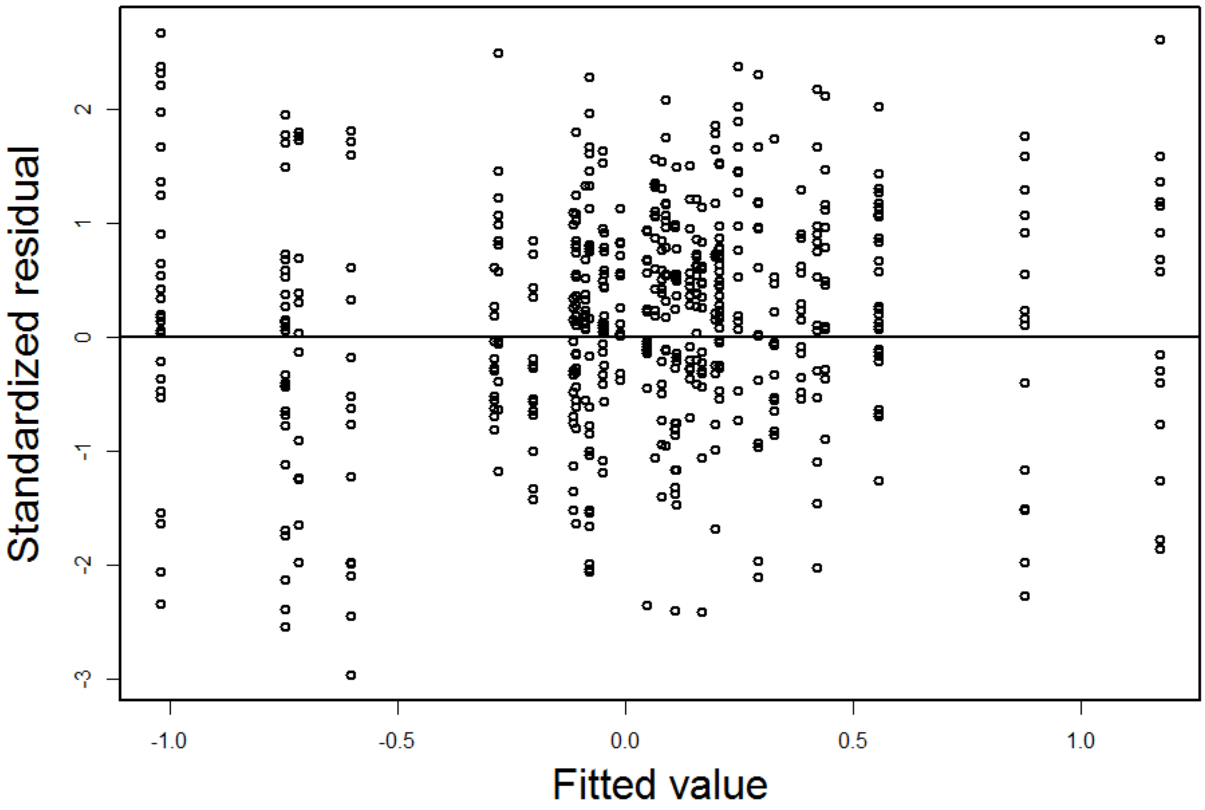


Figure 9. Plot of standardized residuals versus fitted values.

5.4 Material sent to experts

Please refer to Supplementary Material.

5.5 Workshop summary details

CEBRA biofouling project update

Summary

1. **Set the date for the workshop (mid-March)**
2. **Identify experts:**

   In consultation with DA determine appropriate stratum of experts. Note we can go more broadly but for DA decision making they can use a subset. Write to experts seeking their participation and gauging their availability for a workshop (noting this is not funded in this project so will need to be done in close collaboration with CEBRA).

Completion by 24/12

3. **Develop scenarios based on assessing relative risk as outlined in the notes.**
   a. **Consider using data on ship movements to Australia as a sampling frame, and for each movement match a "useful" comparison.**
   b. **Draft instruction to deal with "marginalisation" of unknown factors.**

c. **Develop cueing data (i.e. temp, salinity, substrate, hydrodynamics)**

Draft scenario and instructions developed by 24/12.

Full set of scenarios 06/02.

Deployment 20/2

Draft 26/11/2014

CEBRA Biofouling workshop summary notes

31 October 2014, Acton, Canberra

Present: Simon Barry, Emma Johnston, Chris Starkey, Graham Clark, Peter Caley, Mark Burgman, Tony Arthur, Dean Paini

Scribe: Peter Caley

Scope and background of project (SB)

Fundamental problems/challenges with current Species Distribution Modelling (SDM)

- causative vs. correlative determinants of species distributions
- problems translating data/models to novel environments
- Example of *Crassostrea gigas*, where increasingly complicated predictive models have ultimately had to be tuned to local distribution data – the only data of use is the distribution in OZ.

Expose of challenges of handling list of 40 spp.

- list may keep changing
- can we really generate estimates for 40 species, when our previous attempts for *c.* 7 species have struggled
- measurement of the aggregate level of risk may be more tractable – "how risky is a shipping arriving in Sydney from Singapore?"
- DoA resource allocation, port specific or OZ more generally?

MB noted the previous work of Rupert Summerson

- Temperature matching in ports
- Darren Kriticos work + machine learning
- General consensus that sufficient data lacking for marine species

SB – Marine data lacking

EJ – Especially port environments – terrestrial ⇔ marine interface

GC – Work showed temperature most important variable

SB – General discussion of CEBRA terrestrial project, noting potential cross over to marine environment of Elith paper.

GC questioned whether it was the intention to integrate the tool being developed with the MGRA. CS indicated that would be the plan.

SB then gave a precis of the MGRA

- Uses valid risk factors
- Scoring system sits behind
- “adds stuff up”
- Potentially not well calibrated (and no technical report sitting behind it)
- 5-year trial (1st four years not mandatory)
- Outcome of this work to build on it
    - “gate keeping” role
    - Nuanced () role
- Not integrating [risk] directly

GC noted a possible problem of duplicated variables in both systems. It was agreed that there will be inevitably some overlap. TA noted that reducing the amount of overlap would reduce the number of possible variables to be included in a scenario.

Question (TA) of how quickly species could reach sexual maturity? EJ noted it was variable – 1-2 weeks for some weedy species (e.g. Bryozoans). About 6 weeks for barnacles. Generally faster in the tropics.

Straw man Risk model

SB noted:

- There are lots of things to consider when assessing risk, hence the need to aggregate.
- Important to for communication to experts regarding aggregation of risk factors
- We’re not 2nd guessing the MGRA – can be different.

Conditioning vs. marginalisation

When we need to condition on risk factors (provide cue)

- E.g. “vessel of type X, berthed for 6 months previously at port Y”

OR

Average (marginalize) over possibilities

- “Consider a typical ship ...”
- Certain things cancel out which helps

Practicality issue – avoiding trying to elicit over too many things.

Question of estimating relative risk (easier) versus absolute risk (not well known). It was argued that relative risk would still be a major step forward, as it enables prioritisation.

Ensuing discussion

ET – Average over antifouling history + length of stay in previous port + Australia?

Question of how far back in time should we be evaluating over?

What levels of Antifouling to use? No AF versus “Best” AF?

Don’t want to throw out best variables (e.g. anti fouling) if available (GC).

SB – If the scenario gets too complicated, experts will start cueing off just a couple of variables. They will be integrating over other factors (e.g. sea chests) ... they may, for example, decide that sea chests are more important.

A bit nebulous how “bolted on” to MGRA?

MB – shouldn't be constrained.

CS – MGRA not tested ("line in sand")

ES – can test using New Zealand datasets

SB – need to calibrate fouling to Biosecurity risk

EJ – New Zealand port data + rec vessels + commercial available – need to use NZ data to help inform experts

MB – make model 1st.

SB – Conceptually, population of vessels entering Australia, uncertainty, need to marginalize over unknown factors. In a statistical sense, get uncertainty correct at the population level. Scenarios weighted to reflect kinds of patterns of shipping data.

EJ – discussed unpublished NZ data (Graham) [Simon can contact]

- ~200 ships + covariates
- Boosted Regression Trees to estimate risk factors

SB – Discussion of different ways of using experts ...

- Outline of a scenario
- Explanation of relative risk estimation
- Can then explore factors underlying relative risks (e.g. GLMMs)
- Vessel level risk assessment

MB – presented counter example of using frequencies to estimate absolute risk (e.g. how many of 1,000 visits would you expected to result in unwanted endpoint). Noted that relative risk will still suffer from anchoring.

Group discussion [considerable] around estimating absolute frequencies vs. relative risks

- scenario path vs. direct parameter estimation

SB on talked on issues of framing

- all 40 species or species more generally [recognising that risks are more diffuse that these species only].
- List can change, important species may be missing, different from CCIMPE list (currently defunct).
- Would strongly prefer generic approach as this future proofs work better
    - Experts may not know 40 species
    - Safe at a policy level

MB – as a start, explore the temperature tolerances of all 40 species on the same axis. Still prefers frequencies [SB – Need some scientific "theatre" to kick off on – this could be some].

EJ – expect a sigmoidal relationship between likelihood and the number of trips.

TA – not necessarily sigmoidal, unless we are talking about Establishment cf. Entry and there are Allee effects at play (see Figure 1).

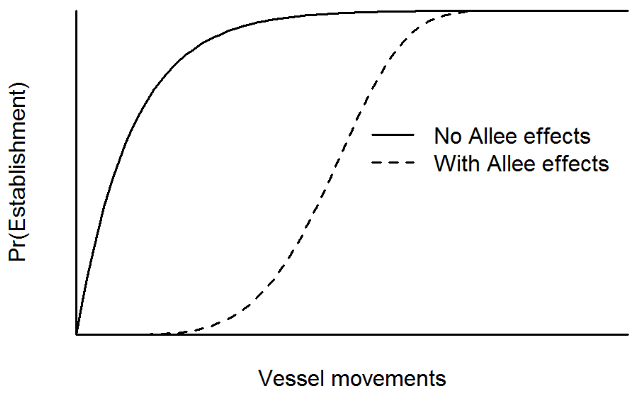


Figure 1. Schematic impact of Allee effects on the relationship between vessel movements to a port and the probability of establishment of an exotic species.

Discussion then moved to the definition what endpoint the likelihood was for (Entry? Establishment?). The Biosecurity literature uses Entry, Establishment and Spread (EES)

PC then tabled a report "**Identifying significant range extensions of invasive marine pests – CCIMPE Range Decision Guidelines Project**" written with Rebecca Darbyshire for the BRS, to illustrate the successful use of the then named Point of Truth Callibration (PoTCal) method for eliciting. Paper by Barry describing the method to be resubmitted shortly.

Discussion surrounding Allee effects

SB – Are experts cueing on volume? Can we use scenarios to assess?

MB – arrivals proportional to propagule pressure

SB – in practical sense, more vessels matter. What level of incursions matters?

MB – viruses

GC – Economically speaking, very few fouling species matter! Macro fouling more of a problem than micro fouling.

MB – introduced the concept of a values hierarchy

- Way of structuring what matters
- Measurable proxies for things that matter
- Avoidance of double counting

SB – "Poor man's approach"

- Impact definitions
- Framing for experts – what "matters" for industry and environment

General discussion on impacts that "matter"

GC – distribution of impacts

MB – how about "one of the species on the list establishes and persists for 12 months"?

EJ – What does the DoA regard as mattering?

Discussion of definition of establishment & incursion, noting that an established population may not persist …

PC – an incursion in generally usage means a reproducing population (i.e. Entry and Establishment).

Everyone agreed that the risk endpoint of interest would be "establishment for 12 months in port environment" though not necessarily reproductive.

SB – asking experts to integrate of recipient port conditions. On the question of relative vs. absolute risk, we can do a bit of both.

MB – nice to have externally validated information bounding approaches (upper, lower, belief etc.). Also, how to word questions to be closer to the truth? People tend to anchor on 1st question.

EJ – noted that 20% of Sydney harbour fish are seasonal (found in summer only).

Outline of relative modelling approach [SB]

Relative vs. absolute. Presented argument as to why relative is useful.

Let p be the probability of undesired endpoint. Using a logistic regression model:

$$\text{logit}(p) = \beta_0 + \beta_1 + \ldots + \beta_k = lp\,[say]$$

$$p = \text{ilogit}(p) = \frac{e^{lp}}{1 + e^{lp}}$$

Assuming the incursion events are rare, then $e^{lp} \approx 0$ and hence $p \approx e^{lp}$ and $\log(p) \approx lp$. This is useful, as for two scenarios, the log of the ratio of the incursion probabilities is:

$$\begin{aligned}\log\left(\frac{p_1}{p_2}\right) &= \log(p_1) - \log(p_2)\\ &= lp_1 - lp_2\end{aligned}$$

And the similar/same risk predictors will cancel. There is an argument that experts find p_1/p_2 easier to estimate, and have done so previously (see Darbyshire and Caley 2009).

Absolute risk requires marginalizing over all the predictors. Although rewriting reveals the seeming similarity between the two approaches.

$$\frac{p_1}{p_2} \quad vs \quad \frac{n_1/BD}{n_2/BD}$$

Where n is the number of adverse events and *BD* is the base denominator (e.g. 1,000). Answering the question "how much more dangerous is scenario #1 than #2?" avoids separate calculations.

MB suggested allowing the experts to choose either.

Presentation of scenarios

It is proposed to present seasonal temperature and salinity data to experts graphically (Figure 2). The associated question would be "What is the relative frequency of events [that matter] between the two scenarios?".

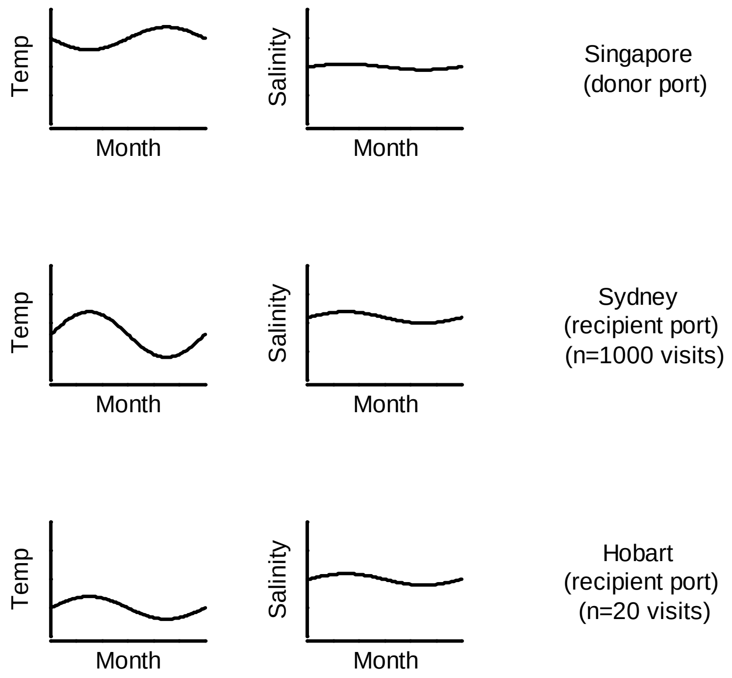


Figure 2. Schematic of proposed presentation of seasonal data to experts for donor and recipient ports.

TA noted that scenarios needed to be carefully constructed.

Discussion of what matters

SB – cueing? Looking for seasonal profiles, with minimum and maximum being the most important (EJ)

PC – easier to marginalize over season (i.e. scenarios aren't seasonal)

Does salinity matter? (SB)

EJ – the amount of hard substrate matters a lot (e.g. rock reefs, artificial structures). The extent can be estimated somewhat from shore modification.

SB – can it [predictor] be operationalized? E.g. do we know salinity profiles from donor ports?

MB – could classify more broadly as tropical/wet with flush (as proxy for changes in salinity), tropical/dry etc.

Factors to consider – Temp, Salinity, Hard substrate, Flush

General discussion on the issue of "connectedness" of the donor port.

SB noted that it may effectively cancel as on the source side. Can we leverage Lloyds data from AMSA?

GC mentioned use of ship logs data.

EJ noted the study of Ruiz *et al*. (2000) – connections still happening, and the hotspots paper of Drake and Lodge (2004).

EJ – NZ has nice data on connectedness (Flioel [spelling])

EJ – heavy metal contamination effects. Also, "portliness" of ports (low flushing) – deep water, retain larvae, slower flow etc. Disturbance (storms) of recipient port. Diversity of visits.

Port exchange/retention data (Simon to chase)

Obtaining experts – Snowballing

PC – Previous marine experts from range extension work

DP – MPSC members

EJ – connections

A whole pile versus subsets [meaning?]

Strata's

- Bioinvasion experts (typically taxon based)
- Biofouling managers
- Bioninvasion ecologists (invertebrates & algae) (better at general environmental matching)
- Modellers & empiricists
- Tropical & temperate
- Industry (e.g. Ashley Coutts)

The idea is to bring a subset of experts together for refinement of the method

MB – "double up" – estimate privately, meet to discuss, then repeat. Gave example of emailing 20 people with 100% response rate [enthusiastic birders?], then got back for second round.

SB – worried about diverse set of people in the one room (museum, academic, industry), particularly difficulty in managing highly structured set of interactions.

MB – reasoning for 2nd round interactions

- Discussion introduce new info
- Refine meaning of Q's (very important –95% of value)
- Better understanding of question
- Mark can facilitate
- Remote method (night parrot example)
- Empirical test of "Understanding question", 10,000 hours burnin test

TA – question of how meeting will interplay with statistical model (too much noise?)

SB – what do we define as expertise? When proposing experts, the chain of logic to support selections needs to be acceptable. Important to get DoA to sign off/bless.

MB on the group meeting – blinded to model outcomes. Private judgements (anonymous, avoid group think). Send questions & background material out before meeting. Groups of 6 working thru scenarios. Take questions home (2nd private judgement)

EJ – survey duration maximum of 2 hours [noted 2016 marine bioinvasion conference in Sydney]

Facilitator – Shuang

Invitations – pick a day in February (call it now).

References

Darbyshire, R., and Caley, P. (2009). Identifying significant range extensions of invasive marine pests – CCIMPE Range Decision Guidelines Project. (Bureau of Rural Sciences: Canberrra).

Drake, J. M., and Lodge, D. M. (2004). Global hot spots of biological invasions: Evaluating options for ballast–water management. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **271,** 575-580.

Ruiz, G. M., Fofonoff, P. W., Carlton, J. T., Wonham, M. J., and Hines, A. H. (2000). Invasion of coastal marine communities in North America: apparent patterns, processes, and biases. *Annual Review of Ecology and Systematics***,** 481-531.

CONTACT US

t 1300 363 400
+61 3 9545 2176
e enquiries@csiro.au
w www.csiro.au

AT CSIRO WE SHAPE THE FUTURE

We do this by using science to solve real issues. Our research makes a difference to industry, people and the planet.

As Australia's national science agency we've been pushing the edge of what's possible for over 85 years. Today we have more than 5,000 talented people working out of 50-plus centres in Australia and internationally. Our people work closely with industry and communities to leave a lasting legacy. Collectively, our innovation and excellence places us in the top ten applied research agencies in the world.

WE ASK, WE SEEK AND WE SOLVE

FOR FURTHER INFORMATION

Oceans and Atmosphere
Simon Barry
t +61 02 6216 7157
e simon.barry@csiro.au
w www.csiro.au/en/Research/OandA

Health & Biosecurity
Peter Caley
t +61 02 6216 7063
e peter.caley@csiro.au
w www.csiro.au/en/Research/BF