Fooled by local robustness: an applied ecology perspective

MOSHE SNIEDOVICH

Department of Mathematics and Statistics, University of Melbourne, Melbourne VIC 3010 Australia

Abstract. In this short discussion, we point out that it is apparently as easy to be fooled by robustness as it is to be fooled by randomness. Our objective is to bring to the attention of applied ecologists that radius-of-stability robustness models are models of local robustness. As such, these models are utterly unsuitable for the treatment/management of a severe uncertainty characterized by a vast uncertainty space and a likelihood-free quantification of the uncertainty. This observation is particularly pertinent to applications of info-gap decision theory in ecology, conservation biology, and environmental management, where the objective is to identify decisions that are robust against a severe uncertainty of this type.

Key words: global; info-gap decision theory; local; radius of stability; robustness; uncertainty.

INTRODUCTION

To motivate the discussion, consider the following abstract of an article that was published recently in Ecological Applications (van der Burg and Tyre 2011; emphasis added):

"Wildlife managers often make decisions under considerable uncertainty. In the most extreme case, a complete lack of data leads to uncertainty that is unquantifiable. Information-gap decision theory deals with assessing management decisions under extreme uncertainty, but it is not widely used in wildlife management. ... We provide a novel combination of the robust population management approach for matrix models with the information-gap decision theory framework for making conservation decisions under extreme uncertainty."

One's first reaction to this text would be: hang on! Might not this proposition be a bit too good to be true? After all, identifying management decisions that are robust against an uncertainty so extreme, an uncertainty that is unquantifiable due to a complete lack of data, is an onerous task!

Suppose then that one were to find reassurance in the following statement, which not only proposes an "info-gap uncertainty analysis," but gives it a ringing endorsement (Moilanen et al. 2006:23; emphasis added):

"In summary, we recommend info-gap uncertainty analysis as a standard practice in computational reserve planning. ... Information-gap decision theory provides a standardized methodological framework in which implementing reserve selection uncertainty analyses is relatively straightforward."

More reassurance would presumably be found in the following assertion in an article that was also published in Ecological Applications (Regan et al. 2005:1473; emphasis added):

"The power and novelty of the info-gap approach is in the ability to explore the sensitivity of the decision to a wide range of different types of parameter, functional, and structural errors and uncertainties simultaneously, given that we do not know the extent of uncertainty in the system at the outset."

But, most reassuring of all would probably be the following statement, which seems to lend support to the above propositions to use the info-gap methodology in ecosystem management (Hildebrandt and Knoke 2011:7):

"We will discuss information-gap decision theory (IGDT; Ben-Haim 2006) in greater detail, because this technique has obtained the greatest popularity in ecosystem management so far."

As for the kind of robustness that info-gap decision theory is claimed to provide, consider the following extract from an article that was also published in Ecological Applications (Nicholson and Possingham 2007:252; emphasis added):

"For example, the best management option may be one that ensures that a species does not exceed a given risk of extinction under the highest possible level of unfavorable uncertainty. The decision may not minimize the extinction risk when uncertainty is ignored, but it is the option least likely to fail because of uncertainty in model structure or parameter estimates."

The question obviously is: how can a likelihood-free model, such as info-gap’s robustness model, possibly identify an option that is least likely to fail?

Similarly, consider this claim (Chisholm 2010:1981; emphasis added):

"For example, the best management option may be one that ensures that a species does not exceed a given risk of extinction under the highest possible level of unfavorable uncertainty. The decision may not minimize the extinction risk when uncertainty is ignored, but it is the option least likely to fail because of uncertainty in model structure or parameter estimates."

Communications
“One possible approach would be the application of tools from non-probabilistic decision theories, such as info-gap decision theory (Ben-Haim 2006). Whereas classical decision theory approaches generally optimise the expected value of the decision variable, the info-gap approach instead minimises the probability of falling below a certain threshold (i.e., it maximises robustness to failure).”

How can a non-probabilistic decision theory, such as info-gap decision theory, possibly minimize the probability of falling below a certain threshold?

Still, putting this matter aside for the moment, the questions that we examine in this discussion are these:

1) Does info-gap decision theory have the capabilities to perform the tasks attributed to it in the applied ecology literature, or are the proposals to use this theory too good to be true?

2) Is info-gap’s robustness model indeed as novel as its advocates proclaim it to be?

The discussion that follows spells out clear-cut answers to these questions. In a nutshell, info-gap’s robustness model is a model of local robustness. As a matter of fact, it is a reinvention of a staple model of local robustness: the radius-of-stability model (circa 1960). As such, info-gap’s robustness model seeks decisions that are robust against small perturbations in a nominal value of the parameter of interest. Namely, it does not seek decisions that are robust against the severe uncertainty it claims to address, because this type of uncertainty requires a model that pursues global robustness.

The objective of this discussion is then to bring to the attention of applied ecologists a number of facts about info-gap decision theory with the view to give these answers content. More details on this topic can be found in Sniedovich (2007, 2010, 2012a, b).

Robustness

Let \( s \in \mathbb{S} \) denote the state of an ecosystem (e.g., population size of a species), where the state space, \( \mathbb{S} \), denotes the set of all possible/plausible states of the system; and let \( x \in \mathcal{X} \) denote a decision associated with the system, where the decision space, \( \mathcal{X} \), denotes the set of admissible decisions.

We distinguish between two types of states, namely “acceptable” and “unacceptable” states, letting \( S(x) \subseteq \mathbb{S} \) denote the set of states associated with decision \( x \) and \( A(x) \subseteq S(x) \) denote the set of “acceptable” states associated with decision \( x \in \mathcal{X} \). It goes without saying that we have an overwhelming preference for the system to be in an “acceptable” state.

In this context, given the variability of state \( s \) over \( S(x) \), the robustness of decision \( x \) is a measure of its ability to perform well over \( S(x) \), namely to render the system in an acceptable state. This is illustrated in Fig. 1 where the state spaces of two decisions, \( x' \) and \( x'' \), are shown as squares \( S(x') = S(x'') = \mathbb{S} \), and the respective sets of acceptable states are represented by the shaded areas \( A(x') \) and \( A(x'') \).

Now, to determine the robustness of the decisions so as to establish which is the more robust, we need to have on hand a measure of robustness. That is, we need a recipe, or a definition, that will make the term “robustness” determinate in this context. For example, in cases where the state \( s \) is a realization of a random variable, we may define the robustness of decision \( x \in \mathcal{X} \) as follows:

\[
R(x) := \Pr \left( s \in A(x) \mid x \right) \quad x \in \mathcal{X}
\]  

where \( \Pr \left( s \in A(x) \mid x \right) \) denotes the probability that the random variable takes value in \( A(x) \) according to its probability distribution, which may depend on \( x \). The larger the value of \( R(x) \), the more robust the decision.

And in cases where the variability of \( s \) over \( S(x) \) is not governed by a probability distribution, we might define the robustness of decision \( x \) as follows:

\[
R(x) := \frac{\text{size}(A(x))}{\text{size}(S(x))} \quad x \in \mathcal{X}
\]

where \( \text{size}(D) \) denotes the “size” of set \( D \subseteq \mathbb{S} \). We shall refer to this recipe as the size criterion. In particular, if \( S(x) \) consists of finitely many elements, then we can let \( \text{size}(D) = |D| \), where \(|D|\) denotes the cardinality of set \( D \).

These two criteria represent a global approach to robustness. Namely, they take account of all the elements of \( S(x) \), which therefore renders them particularly suitable for situations where the decision maker wants, or is required, to assess the behavior/performance of decision \( x \) over the entire state space \( S(x) \) associated with it.

In contrast, there are applications where the decision’s behavior/performance over its entire state space \( S(x) \) is of no concern. Instead, what is of concern is the behavior of the decision in the neighborhood of a given nominal state, call it \( \delta \in S(x) \). For instance, we may want to establish whether a sufficiently large neighborhood of \( \delta \) consists of acceptable states, or whether a large part of
this neighborhood consists of such states. Measures of robustness of this type are termed local.

It is important to note that, in general, a decision that is globally robust is not necessarily locally robust and vice versa. Furthermore, the local robustness of decision \( y \) typically depends on the location of the nominal state \( x \) in the state space \( S(x) \). This is illustrated in Fig. 2 where the circles centered at the nominal states represent the respective neighborhoods over which the (local) robustness analysis is conducted.

Observe that all the elements of the neighborhood of \( \hat{s}_1 \) are “acceptable,” namely, the neighborhood is a subset of \( A(x) \). In contrast, a significant part of the neighborhood of \( \hat{s}_2 \) is not contained in \( A(x) \). Based on this simple observation, we can argue that decision \( y \) is more (locally) robust in the neighborhood of \( \hat{s}_1 \) than in the neighborhood of \( \hat{s}_2 \). Since info-gap robustness model is a model of local robustness, in this discussion we focus on models of local robustness. And to set the stage, we consider a classic model of robustness that can deal both with local and global robustness.

**Wald’s Maximin Model**

It is hard to overstate the importance of Wald’s maximin model (Wald 1950), both as a tool of thought and as a practical “rule,” in the broad area of robustness analysis. Indeed, it is the foremost modeling paradigm and technique used in classical decision theory, robust optimization, robust control, robust statistics, and other fields. Informally, it can be stated as follows:

**Maximin Rule:** Rank alternatives according to their worst outcomes. Hence, select the alternative whose worst outcome is at least as good as the worst outcomes of all other alternatives.

There are many ways to formulate this rule mathematically and the model chosen for a particular application is determined on grounds of the features of the application in question. For reasons that will become clear below, it is instructive to introduce first a general version of the model. At a later stage, we consider a simpler version. We refer to the following general version of the model as the full Monty model:

\[
 z^* := \max_{y \in Y} \min_{s \in \Pi(y)} \{ g(y, s); C(y, s) \ \forall s \in \Pi(y) \} \tag{3}
\]

where \( Y = \) set of alternatives; \( \Pi(y) = \) set of possible/plausible states associated with alternative \( y; g(y, s) = \) outcome (numeric scalar) associated with the alternative-state pair \( (y, s) \); and \( C(y, s) = \) list of constraints on the \( (y, s) \) pairs.

The iconic

\[
 \max_{y \in Y} \min_{s \in \Pi(y)}
\]

operation indicates that robustness is sought with respect to the outcome \( g(y, s) \), whereas the “\( C(y, s), \forall s \in \Pi(y) \)” clause indicates that robustness is also sought with respect to the constraints imposed on the \( (y, s) \) pairs.

Now, consider the rather simple case where the outcome \( g(y, s) \) is independent of \( s \), hence robustness is sought only with respect to the constraints on the \( (y, s) \) pairs. In this case, \( g(y, s) = f(y) \) for all \( s \), where \( f \) is a real-valued function on \( Y \). Since in this case the

\[
 \min_{s \in \Pi(y)} \quad \text{operation is superfluous},
\]

the full Monty model is simplified to

\[
 z^* := \max_{y \in Y} \{ f(y); C(y, s) \ \forall s \in \Pi(y) \}. \tag{4}
\]

We shall refer to this maximin model as the mathematical programming (MP) model. Take note that it is possible to express the full Monty model (Eq. 3) itself as an MP model (Eq. 4). In fact, this is done routinely in the areas of game theory and robust optimization (e.g., Kouvelis and Yu 1997, Bertsimas and Sim 2004, Ben-Tal et al. 2009a, Sniedovich 2012a). So, although the MP model (Eq. 4) has a far simpler formulation than the full Monty model (Eq. 3), the former is as powerful as the latter.

**Radius of Stability**

Radius of stability models are models of local robustness designed to deal with small perturbations in the nominal value of the parameter of interest. Informally, the radius of stability of decision \( x \in X \) at a nominal state \( \hat{s} \in S(x) \) is the radius (size) of the largest neighborhood around \( \hat{s} \) all of whose elements are “acceptable.” Namely, it is the radius (size) of the largest neighborhood around \( \hat{s} \) that is contained in \( A(x) \). More formally, let \( N(\rho, \hat{s}) \) denote a neighborhood of radius \( \rho \geq 0 \) around \( \hat{s} \), that is, let it denote the set of all the states in \( S \) that are within a distance \( \rho \) from \( \hat{s} \) according to some suitable measure of distance on \( S \). Then the radius of stability of decision \( x \) at \( \hat{s} \) is defined as follows:

\[
 \rho(x, \hat{s}) := \max_{\rho \geq 0} \{ \rho; s \in A(x) \ \forall s \in \mathcal{N}(\rho, \hat{s}) \}. \tag{5}
\]
This is illustrated in Fig. 3, where the neighborhoods around two nominal states are represented by nested ellipses. The radii of stability are the radii of the ellipses that are tangent to the boundary of \(A(x)\) (shown in heavy lines).

The global and local perspectives on the two cases are as follows:

**Global**: Decision \(x'\) is much more robust than decision \(x''\) because \(A(x')\) contains \(A(x'')\) and it is much larger.

**Local**: The radius of stability of decision \(x''\) at \(\hat{s}\) is much larger than the radius of stability of \(x'\) at \(\hat{s}\). Hence, according to the radius of stability model, decision \(x\) is more (locally) robust at \(\hat{s}_1\) than at \(\hat{s}_2\). Fig. 4a illustrates the difference between local robustness, as determined by the radius of stability model, and global robustness (e.g., as determined by the size criterion).

The fact that decision \(x'\) is more robust locally than decision \(x'',\) whereas globally the converse is true, should come as no surprise because, the nexus between local and global robustness notwithstanding, the two concepts designate two fundamentally different approaches to robustness. The distinction between global and local robustness is reminiscent of the distinctions between local and global weather, local and global news, local and global anesthetic, and so on. The implication is then that it is incumbent on anyone proposing a robustness analysis to determine a decision’s robustness, to make it unambiguously clear what type of robustness does this analysis seek: local or global.

As for the connection of this issue to Wald’s famous maximin model, consider this:

**THEOREM 1**: The radius of stability model (Eq. 5) is an instance of Wald’s maximin model. In particular, it is an instance of the MP model (Eq. 4).

**Proof.** For a given decision \(x \in X\), consider the case where \(y = \rho, Y = [0, \infty), f(y) = y, \Pi(y) = \phi(y, \hat{s})\), and \(C(y, s)\) consists of the single constraint \(s \in A(x)\). Substituting these values in Eq. 4 we obtain the radius of stability model (Eq. 5).

With this in mind, let us now examine where exactly does info-gap decision theory, notably its robustness model, fit in this picture.

**INFO-GAP DECISION THEORY**

According to info-gap decision theory (Ben-Haim 2001, 2006, 2010), the robustness of decision \(x \in X\) is defined as follows:

\[
\hat{\rho}(x, \hat{s}) := \max_{\rho \geq 0} \{ \rho: r_c \leq r(x, \hat{s}) \quad \forall \hat{s} \in \mathcal{X}(\rho, \hat{s}) \}
\]  

where \(r_c\) represents a critical performance level and \(r(x, \hat{s})\) denotes the performance level of decision \(x\) at state \(s\). Furthermore, info-gap decision theory prescribes that the more robust a decision, the better. Hence, for a given value of \(r_c\), the optimal (best) decision, as far as...
robustness is concerned, is a decision whose robustness is the largest. Consequently, info-gap’s decision model for determining the most robust decision is as follows:

\[ \hat{\delta}(\delta) := \max_{x \in X} \hat{\rho}(x, \delta) \]  
\[ = \max_{x \in X, \rho > 0} \{ \rho: r_c \leq r(x, s) \} \quad \forall s \in \mathcal{N}(\rho, \delta). \]  

This is illustrated in Fig. 4b, where the info-gap robustness of two decisions is compared. The neighborhoods \( \mathcal{N}(\rho, \delta) \) are represented by circles centered at \( \delta \), and the values of \( \delta \) satisfying the performance constraint \( r_c \leq r(x, s) \) are represented by the shaded areas. This figure makes vivid that the info-gap robustness of a decision is the radius of the largest circle centered at \( \delta \) such that the decision satisfies the performance constraint \( r_c \leq r(x, s) \) at all the values \( s \) in the circle. But more than this, it is clear by inspection that the info-gap robustness of decision \( x' \) is greater than the info-gap robustness of decision \( x'' \). This means that, according to the precepts of info-gap decision theory, decision \( x'' \) is more robust than decision \( x' \).

But in the applied ecology literature, info-gap decision theory is portrayed as a theory of global robustness. This is clearly brought out by Halpern et al.’s (2006: Fig. 1) comparison of seven approaches to uncertainty, where info-gap decision theory is associated with the most severe uncertainty. This is a faithful reflection of Ben-Haim’s position (2001, 2006, 2010) on the uncertainty that info-gap decision theory is designed for. Indeed, in these texts info-gap decision theory is claimed to be able to deal with a severe uncertainty that is characterized by the following features: (1) the uncertainty space \( \delta \) can be vast (it is often unbounded); (2) the point estimate \( \delta \) is poor, unreliable, and can be substantially wrong; (3) the uncertainty is non-probabilistic and likelihood free.

As a matter of fact, Wintle et al. (2010) go so far as to argue that info-gap decision theory can handle Black Swans (see Taleb 2005, 2007) and even Unknown Unknowns! According to Taleb (2007), a Black Swan is an event with the following three attributes: rarity, extreme impact, and retrospective (though not prospective) predictability.

**Discussion**

Details on the relation between info-gap’s robustness model, the radius of stability model, and Wald’s maximin model, can be found in Sniedovich (2010, 2012a, b). For the purposes of this discussion it suffices to note the following:

**THEOREM 2:** The info-gap robustness model (Eq. 6) is a radius of stability model (Eq. 5).

**Proof.** By inspection, the info-gap robustness model (Eq. 6) is that instance of the radius of stability model (Eq. 5) that is specified by \( A(x) := \{ s \in S(x): r_c \leq r(x, s) \} \) (see Fig. 4).

**THEOREM 3:** The info-gap’s robustness model (Eq. 6) and the info-gap’s decision model for robustness (Eq. 8) are simple instances of Wald’s generic maximin model. More specifically, both are instances of the MP model (Eq. 4).

**Proof.** By inspection, the info-gap’s robustness model (Eq. 6) is that instance of the MP model (Eq. 4) that is specified, for any given \( x \in X \), by \( y = \rho \), \( Y = [0, \infty) \), \( f(y) = y \), and where \( C(y, s) \) is the single constraint \( r_c \leq r(x, s) \), observing that in this context decision \( x \) is fixed and given, and \( \rho \) has the role of the “alternative” in this optimization model (Eq. 6). Similarly, info-gap’s decision model for robustness (Eq. 8) is that instance of the MP model (Eq. 4) that is specified by \( y = (x, \rho) \), \( Y = X \times [0, \infty) \), \( f(x, \rho) = \rho \) and where \( C(x, \rho, s) \) is the single constraint \( r_c \leq r(x, s) \).

For an illuminating analysis of the distinction between local and global robustness, consider the following statement (Hafner et al. 2009:2; emphasis added):

“Robustness is an intrinsic property of many biological systems. To quantify the robustness of a model that represents such a system, two approaches exist: global methods assess the volume in parameter space that is compliant with the proper functioning of the system; and local methods, in contrast, study the model for a given parameter set and determine its robustness. Local methods are fundamentally biased due to the a priori choice of a particular parameter set.”

In the case of info-gap decision theory, this fundamental bias can be described in detail as follows:

**THEOREM 4:** The info-gap robustness of decision \( x \in X \) is determined in total disregard of the performance levels \( r(x, s) \) associated with states \( s \in S(x) \) that are outside the neighborhood \( \mathcal{N}(p^*, \delta) \), where \( p^* = \hat{\rho}(x, \delta) + \varepsilon \) and \( \varepsilon \) can be arbitrarily small but positive.

**Proof.** This follows immediately from the fact that the neighborhoods \( \mathcal{N}(\rho, \delta), \rho \geq 0 \) are nested, namely, that \( \rho' < p^* \) implies that \( \mathcal{N}(\rho', \delta) \subseteq \mathcal{N}(\rho^*, \delta) \).

This means that the evaluation of the info-gap robustness of decisions may completely ignore the decisions’ performance over a significant part of the uncertainty space. The effects of applying radius of stability models (hence, info-gap’s robustness model) in vast uncertainty spaces (which give expression to severe uncertainty) are referred to in Sniedovich (2010, 2012a, b) as the No Man’s Land effect. This issue is illustrated in Fig. 5 where the white circle represents the neighborhood \( \mathcal{N}(p^*, \delta) \) where \( p^* \) is slightly larger than \( \hat{\rho}(x, \delta) \) and the
The dark area represents the No Man’s Land associated with decision $x \in X$.

Another aspect of the fundamental flaw afflicting info-gap’s prescription for the management of severe uncertainty is a result of the combined effect of the “fundamentally biased a priori choice of a particular parameter set” (neighborhood), and the uncertainty being likelihood-free. In greater detail, because the info-gap robustness of a decision is contingent on the value of the point estimate $\hat{\delta}$, the ranking of decisions, hence the optimal decision, is also contingent on the value of the point estimate $\hat{\delta}$. This means that an info-gap robustness analysis may yield different solutions for different values of $\hat{\delta}$.

Considering, however, that the true value of $s$ is severely uncertain, to deal with this matter would take us straight back to square one, to grapple with the quality of the estimate (nominal state), namely with the severe uncertainty in the true value of the state (See Hayes [2011] for an extensive analysis of this issue).

In a word, info-gap’s prescription for the management of severe uncertainty does not deal with the difficulties presented by severe uncertainty: it ignores them. The effects of the No Man’s Land phenomenon seem to be alluded to in the observation by Ben-Tal et al. (2009a:926) that a model that ignores the full range of the parameter values and considers only its “normal” range, “... represents a somewhat ‘irresponsible’ decision maker. ...” Sniedovich (2010, 2012a, b) refers to this as “voodoo decision-making.”

CONCLUSIONS

The narrative in the applied ecology literature, describing info-gap decision theory and arguing for its application in the management of ecosystems subject to severe uncertainty, gives a grossly misleading picture of what this theory is and what it is capable of. Not only that this theory does not offer a new method for quantifying severe uncertainty and for seeking robustness against it, its robustness model is a re-invention of a staple model of local robustness: the radius of stability model (circa 1960). And what is more, this model and the decision model based on it, are simple instances of Wald’s famous maximin model (circa 1940). Thus, as a theory of local robustness, it proves utterly unsuitable for the treatment of severe uncertainty of the type it claims to address. The rich literature on robust decision-making, notably the robust optimization literature, provide tools for dealing properly with a severe uncertainty of the type stipulated, but mishandled, by info-gap decision theory.

LITERATURE CITED


Sniedovich, M. 2012b. Fooled by local robustness. Risk Analysis, in press.


