

Robustness, Fidelity and Prediction-Looseness of Models

Yakov Ben-Haim
Yitzhak Moda'i Chair in
Technology and Economics
Faculty of Mechanical Engineering
Technion—Israel Institute of Technology
Haifa 32000 Israel
yakov@technion.ac.il

François M. Hemez
X-Theoretical Design Division
P.O. Box 1663, M/S T087
Los Alamos National Laboratory
Los Alamos, NM 87545 USA
hemez@lanl.gov

Contents

1	Introduction	2
2	Notation	3
3	Robustness of the Fidelity between Model and Measurement	4
4	'Looseness' of Model Prediction	6
4.1	The Definition	6
4.2	The Theorem	7
4.3	The Dilemma	8
5	Example: A Non-Linear System	9
5.1	Formulation	9
5.2	Prediction Looseness	10
5.3	Sympathetic Robustness and Looseness	11
6	Conclusion: Methodological Issues	12
6.1	Positivism, Prediction and Uncertainty	12
6.2	Induction, Warrant and Prediction	14
7	References	15

Keywords: Modelling, uncertainty, info-gaps, robustness, fidelity to data, prediction.

Abstract

Assessment of the credibility of a mathematical or numerical model of a complex system must combine three components: (1) The fidelity of the model to test data, e.g. as quantified by a mean squared error. (2) The robustness, of model fidelity, to lack of understanding of the underlying processes. (3) The prediction looseness of the model. 'Prediction looseness' is the range of predictions of models that are equivalent in terms of fidelity. The main result of this paper asserts that fidelity, robustness, and prediction looseness are mutually antagonistic. A change in the model that enhances one of these attributes will cause deterioration of another. In particular, increasing the *fidelity* to test data will decrease the *robustness* to imperfect understanding of the process. Likewise, increasing the *robustness* will increase the *predictive looseness*. The conclusion is that focusing only on fidelity-to-data is not a sound decision-making strategy for model building and

validation. A better strategy is to explore the trade-offs between robustness-to-uncertainty, fidelity to data, and tightness of predictions. Our analysis is based on info-gap models of uncertainty, which can be applied to cases of severe uncertainty and lack of knowledge.

1 Introduction

In computational physics and engineering, biological conservation, economics, homeland security, and other fields, numerical models are developed to predict the behavior of a system whose response cannot be measured experimentally and whose behavior is incompletely understood. A key aspect of science-based predictive modelling is to assess the credibility of predictions. Credibility expresses the extent to which the results of simulation reliably represent the phenomenon of interest with a degree of accuracy consistent with the intended use of the model (Doebling, 2002). The literature is rich in examples of the need for credible system modelling (Banks et al, 2003; Bårdsen et al, 2005; Burgman, 2005; Friswell and Mottershead, 1995; Natke and Cempel, 1997).

The study of the credibility of models has a long history. Perhaps the earliest contribution was Laplace’s proof of the central limit theorem and its use to provide a maximum likelihood motivation for least-squares estimation. Hypothesis tests such as the chi-squared test or the Kolmogorov-Smirnov test provide statistical concepts and tools for assessing model credibility. Tests such as these assume asymptotic data sets, statistical independence of observations, and so on. Bayesian methods of model up-dating have also been extensively used. Kennedy and O’Hagan (2001) develop a Bayesian approach to model calibration which derives the posterior distribution of estimated model parameters and of subsequent model predictions, by assuming that functions of the model error are described by a Gaussian process. The choice of an approach to assessing model credibility depends on the analyst’s judgment and understanding of the problem. As Kennedy and O’Hagan note regarding their assumption of normality: “It is, of course, important that normality . . . is a reasonable representation of prior knowledge or beliefs” (2001, p.432).

The main contributions of this paper employ info-gap decision theory, which is particularly suited to situations in which probabilistic information is deficient or lacking. Nonetheless, one element of our analysis can employ probabilistic or statistical tools.

This paper argues that assessment of the credibility of a mathematical or numerical model must combine three components: (1) The fidelity of the model to test data. Fidelity can be quantified in different ways as discussed in section 3, for instance with mean squared deviation between test data and model predictions, or other methods. (2) The robustness, of model fidelity, to lack of understanding of the underlying processes. (3) The prediction looseness of the model. ‘Prediction looseness’ refers to the magnitude of the range of predictions expected from a family of models all of which have fidelity no worse than a specified value. Predictive focus is the complement of predictive looseness. The importance of prediction looseness stems from the fact that, to predict with confidence, there should be little difference (or small looseness) between the predictions of models with the same fidelity. We stress that the credibility of a model depends on all three attributes: fidelity to data, robustness to ignorance, and prediction looseness. No one attribute alone can establish the credibility of a model.

The main results of this paper, theorems 1 and 2, assert that fidelity, robustness, and prediction looseness are mutually antagonistic. A change in the model which enhances one of these attributes will cause deterioration of another. In particular, increasing the *fidelity* to test data will decrease the *robustness* to imperfect understanding of the process. Fidelity to data adds warrant to the predictions by anchoring them in experience. But without robustness to imperfect understanding of the underlying process, the fidelity is not credible. Likewise, increasing the *robustness* will decrease the *predictive focus*. Robustness is needed in order to add warrant to the fidelity, but lack

of predictive focus vitiates the main purpose of the model. A model can be highly warranted by virtue of high fidelity to data and high robustness to epistemic uncertainty, only by increasing its predictive looseness.

It is important to stress that our definition of ‘model’ is not restricted to first-principle partial differential equations. ‘Models’ may be physics-based models, phenomenological equations, back-of-the-envelope calculations, statistical regression models obtained by fitting test data, as well as expert judgment. These are all models in the broad sense that we use these various representations of information and knowledge to make predictions, and they are all based on assumptions of various sorts. Similarly, a ‘family of models’ is not restricted to a model whose coefficients can be changed. The family of models could include all of the above classes as alternatives.

In section 2 we introduce our notation. In section 3 we define the info-gap robustness function and re-iterate the well-known result that robustness and fidelity are antagonistic. The main contribution of the paper appears in section 4 which discusses predictive looseness and its antagonistic relation to robust fidelity. Section 5 presents an heuristic example that illustrates the interaction between model structure and measurement error for the common scientific activity of identifying a mathematical law governing the behavior of a system. We conclude the paper in section 6 with a discussion of some philosophical questions. The scientific activities of validating models and warranting our understanding of the underlying complexity are questioned in light of the trade-offs exposed by the theorems.

2 Notation

Our basic notation is:

y is an observable real-valued vector which is predicted by a model. Examples include material deflections, natural vibration frequencies, population densities, economic variables, etc. The dimension of y is J .

p is a vector of control parameters which characterize the configuration of the system and which may, in some situations, be chosen by the analyst. Examples include material dimensions and properties, loading forces, temperatures, time, location, habitat, economic regime, etc. Specification of the configuration will be important when we wish to distinguish between the configuration p at which a model is up-dated and the configuration p' (which may or may not differ from p) at which forecast or prediction is made.

q is a vector of parameters which specify the structure and coefficients of the model for predicting y . These parameters are not necessarily real-valued coefficients. The elements of q can represent discrete or linguistic variables that select a model structure, a functional form expressing the relation between variables, etc.

The vectors p and q differ from one another. p specifies the regime or configuration of a physical system, either when it was measured or a regime in which its behavior will be predicted. q specifies a mathematical model for describing the physical system. The model depends on the configuration of the system, p , and the model is specified by q which, for instance, may be the values of coefficients.

Our analysis entails two types of mathematical models. The *physical models* $\tilde{M}(p, q)$ and $M(p, q)$ predict values of y . The *uncertainty model* $\mathcal{U}(\alpha, \tilde{M})$ represents the epistemic uncertainty in the physical models.

$\tilde{M}(p, q)$ is a physical model which will be up-dated and used for predicting the value of y in configuration p . Different realizations of $\tilde{M}(p, q)$ are specified by the analyst’s choice of q . $\tilde{M}(p, q)$ is a real J -vector-valued function. We will sometimes refer to $\tilde{M}(p, q)$ by its specifying vector q .

$M(p, q)$ is an alternative possible physical model for predicting the value of y in configuration p , for instance, a model proposed by a competing theory or a different expert. In the absence of

epistemic uncertainty, no alternatives to $\widetilde{M}(p, q)$ would be considered. As the horizon of uncertainty increases, more and more alternative models become candidates. Like $\widetilde{M}(p, q)$, $M(p, q) \in \mathfrak{R}^J$.

$\mathcal{U}(\alpha, \widetilde{M})$ is an info-gap model for the uncertainty in $M(p, q)$. The info-gap model is an unbounded family of nested sets of physical models $M(p, q)$, centered on $\widetilde{M}(p, q)$. We will encounter an example of an info-gap model in section 5. The info-gap parameter for uncertainty in the physical model is α . We will specify generic properties of info-gap models of uncertainty in section 3. See Ben-Haim (2006) for further discussion.

N measurements of y , denoted $\mathcal{Y} = \{y_1, \dots, y_N\}$, have been made in configuration p .

3 Robustness of the Fidelity between Model and Measurement

Let $M(p, q)$ be any physical model in the info-gap model $\mathcal{U}[\alpha, \widetilde{M}(p, q)]$ at horizon of uncertainty α . Let $R[\mathcal{Y}, M(p, q)]$ represent the error of model M with respect to test data \mathcal{Y} . We impose no restrictions of the function $R[\mathcal{Y}, M(p, q)]$ other than it be real-valued, though typically it will be based on a vector norm. For instance:

$$R[\mathcal{Y}, M(p, q)] = \frac{1}{N} \sum_{i=1}^N \|y_i - M(p, q)\|^2 \quad (1)$$

where $\|\cdot\|$ is a vector norm. $R(\mathcal{Y}, M)$ is the weighted least-squared error if $\|\cdot\|$ is the weighted Euclidean norm:

$$\|y_i - M(p, q)\|_{2,W}^2 = [y_i - M(p, q)]^T W [y_i - M(p, q)] \quad (2)$$

where W is a real, symmetric, positive definite matrix. Alternatively, one might use the absolute norm:

$$\|y_i - M(p, q)\|_{\infty,w}^2 = \max_{1 \leq j \leq J} \left| \frac{y_{ij} - M_j(p, q)}{w_j} \right| \quad (3)$$

where $w = (w_1, \dots, w_J)^T$ is a real vector and y_{ij} is the j th element of y_i . In this case $R(\mathcal{Y}, M)$ is the mean, weighted, absolute error.

Various statistical tools for assessing observational error and model fidelity can be employed, such as the level of significance of goodness-of-fit, or a Bayesian measure of risk, etc. The analyst chooses the error function $R[\mathcal{Y}, M(p, q)]$, perhaps in order to filter out or otherwise manage observation error. In our analysis, the only constraint is that $R[\mathcal{Y}, M(p, q)]$ be real scalar-valued.

‘Robustness’ has many meanings. Berger (1980), for instance, has developed Bayesian concepts of robustness. As we will use it, the concept of robustness derives from a prior concept of non-probabilistic uncertainty. Knight (1921) distinguished between ‘risk’ based on known probability distributions and ‘true uncertainty’ for which probability distributions are not known. Wald (1945) studied the problem of statistical hypothesis testing based on a random sample whose probability distribution is not known, but whose distribution is known to belong to a given class of distribution functions. Similarly, Ben-Tal and Nemirovski (1999) are concerned with uncertain data within a prescribed uncertainty set, without any probabilistic information. Likewise Hites *et al.* (2006, p.323) view “robustness as an aptitude to resist to ‘approximations’ or ‘zones of ignorance’”. We are concerned with robustness against Knightian uncertainty.

$\widetilde{M}(p, q)$ is the physical model selected for predicting y . However, the model is surely imperfect. It entails approximations, both acknowledged and unknown. $\widetilde{M}(p, q)$ is robust to these uncertainties if $\widetilde{M}(p, q)$ can err greatly and still reproduce the observed data \mathcal{Y} with acceptable fidelity. Because $\widetilde{M}(p, q)$ is imperfect, the analyst cannot realistically aspire to perfect fidelity to the test data. Rather, $\widetilde{M}(p, q)$ should **satisfice** the fidelity: achieve an acceptable level of fidelity. Let r_c denote

the greatest acceptable error between model and data. The **robustness** of physical model $\widetilde{M}(p, q)$, at configuration p , is the greatest horizon of possible alternative models up to which the error between the model and the data is no greater than r_c :

$$\widehat{\alpha}(q, r_c; p) = \max \left\{ \alpha : \left(\max_{M(p, q) \in \mathcal{U}[\alpha, \widetilde{M}(p, q)]} R[\mathcal{Y}, M(p, q)] \right) \leq r_c \right\} \quad (4)$$

Model q is preferred to model q' if q is more robust than q' , at the same level of satisfied fidelity:

$$q \succ q' \quad \text{if} \quad \widehat{\alpha}(q, r_c; p) > \widehat{\alpha}(q', r_c; p) \quad (5)$$

where ‘ \succ ’ means ‘is preferred to’. The robust-optimal model, $\widehat{q}(r_c, p)$, maximizes the robustness and satisfies the mean-squared error at the level r_c :

$$\widehat{q}(r_c, p) = \arg \max_q \widehat{\alpha}(q, r_c; p) \quad (6)$$

We will denote the maximal robustness, $\widehat{\alpha}[\widehat{q}(r_c, p), r_c; p]$, by $\widehat{\alpha}(r_c; p)$.

The following theorem establishes a basic trade-off between robustness-to-uncertainty, $\widehat{\alpha}(q, r_c; p)$, and fidelity to test data r_c : robustness diminishes ($\widehat{\alpha}$ gets smaller) as greater fidelity is demanded (r_c is reduced). The same trade-off holds also for the maximal robustness $\widehat{\alpha}(r_c; p)$. For proof see Ben-Haim (2000).

For the formulation of this theorem we need to specify two axioms of info-gap models. An info-gap model is a family of nested sets, $\mathcal{U}(\alpha, M)$, $\alpha \geq 0$. These set-valued functions have the following properties:

$$\text{Nesting: } \alpha < \alpha' \text{ implies } \mathcal{U}(\alpha, M) \subseteq \mathcal{U}(\alpha', M) \quad (7)$$

$$\text{Contraction: } M \in \mathcal{U}(0, M) \quad (8)$$

We will employ an additional axiom later on:

$$\text{Translation: } \mathcal{U}(\alpha, M) = \mathcal{U}(\alpha, M') + \{-M'\} + \{M\} \quad (9)$$

The addition operator is the Minkowski sum: adding every element of one set to every element of the other set.

These info-gap axioms are exceedingly unrestrictive and can be used to represent model uncertainties of many sorts. For instance, info-gap models can represent uncertain quadratic terms in nominally linear models, or uncertain slopes of monotonic functional forms. Additionally, info-gap models can represent uncertainty in probabilistic models. For instance, the hyper-parameters of probability densities are often chosen by fitting low-order moments, which determines the unobserved tails of the distribution. This is problematic unless the shape of the distribution is known from first principles. An info-gap model can represent the uncertain shape of the unobserved tails. Numerous examples of info-gap models can be found in (Ben-Haim, 1996, 2006).

Theorem 1 *Robustness improves as fidelity deteriorates.*

Let $\mathcal{U}(\alpha, M)$ be an info-gap model which obeys the axioms of nesting and contraction. Let $\widehat{\alpha}(q, r_c; p)$ and $\widehat{\alpha}(r_c; p)$ be its robustness function and maximal robustness function. $r_c < r'_c$ implies:

$$\widehat{\alpha}(q, r_c; p) \leq \widehat{\alpha}(q, r'_c; p) \quad (10)$$

and

$$\widehat{\alpha}(r_c; p) \leq \widehat{\alpha}(r'_c; p) \quad (11)$$

Eq.(10) asserts that, for any given model, specified by q , the robustness to model-error increases as the required fidelity is relaxed. Eq.(11) asserts that the most robust model at one level of fidelity, r'_c , is more robust than the most robust model at better fidelity, r_c .

Theorem 1 does not imply that it is impossible to find a model that is simultaneously true to the test data (small r_c) and robust to the uncertainty (large $\hat{\alpha}$). For instance, if the noise in the data were very low then the fidelity could be high, regardless of our uncertainty about the model structure. Likewise, if there were little uncertainty about the model structure—if our understanding of the process were nearly complete—then our robustness would be large. What the theorem is stating is that, in any given epistemic state, the robustness to model-error trades off against fidelity to the data.

4 ‘Looseness’ of Model Prediction

In this section we explore the “looseness” of model prediction: the magnitude of the range of predicted values deriving from models which all satisfy a specified fidelity requirement. We prove a theorem whose meaning is that a change in the model which enhances fidelity-robustness to modelling error, $\hat{\alpha}(q, r_c; p)$, also increases the looseness of the model prediction. In other words, under the conditions of the theorem, fidelity-robustness and prediction-looseness are antagonistic attributes of any model. An earlier and less general version of this theorem appears in (Ben-Haim and Hemez, 2004).

4.1 The Definition

We first need some definitions. Let T denote the space in which the elements of the info-gap model $\mathcal{U}(\alpha, M)$ are defined. For any element $\mu \in T$ and any sets U and V in T , define $\rho(\cdot)$ as a “size” function with the following two properties:

$$\text{Nesting:} \quad U \subseteq V \text{ implies } \rho(U) \leq \rho(V) \quad (12)$$

$$\text{Translation invariance:} \quad \rho(U + \mu) = \rho(U) \quad (13)$$

‘Nesting’ means that if U is contained in V then the size of U is no larger than the size of V . ‘Translation invariance’ means that the size of a set does not change as the set is translated in the space.

As an example, let $f(u)$ be a real-valued affine scalar function. The following function satisfies eqs.(12) and (13):

$$\rho(U) = \max_{u \in U} f(u) - \min_{u \in U} f(u) \quad (14)$$

We are now able to define prediction looseness.

The robustness of model $\widetilde{M}(p, q)$ at configuration p , is defined in eq.(4). The robustness of fidelity-to-data can only be evaluated at a configuration, p , for which test data exist. However, $\widetilde{M}(p, q)$ can be used to predict behavior at any configuration, p' ; the prediction is $\widetilde{M}(p', q)$. Of course p and p' can be one and the same: predicting behavior at a configuration for which data exist.

Let $\Lambda(p, p', q)$ denote a set of model predictions at configuration p' , based on information obtained at configuration p . Specifically, based on the info-gap model $\mathcal{U}(\alpha, \widetilde{M})$, $\Lambda(p, p', q)$ is the set of predictions at p' of all models which, at configuration p , have fidelity no worse than the fidelity of $\widetilde{M}(p, q)$. We have no reason to reject any model, M in $\Lambda(p, p', q)$, if fidelity-to-data is used as a measure of merit. The formal definition of $\Lambda(p, p', q)$ is:

$$\Lambda(p, p', q) = \mathcal{U} \left[\hat{\alpha}(q, r_c; p), \widetilde{M}(p', q) \right] \quad (15)$$

$\Lambda(p, p', q)$ is one particular set in the family of nested sets which constitute the info-gap model of uncertainty, $\mathcal{U}(\alpha, \widetilde{M})$, $\alpha \geq 0$. Namely, $\Lambda(p, p', q)$ is the uncertainty set which is centered on $\widetilde{M}(p', q)$ and whose horizon of uncertainty equals the robustness of $\widetilde{M}(p, q)$. If $\Lambda(p, p', q)$ is a large set then the range of predictions of fidelity-equivalent models is large: the prediction looseness is great. If $\Lambda(p, p', q)$ is a small set then the prediction looseness is small.

Using a “size” function, $\rho(\cdot)$, with the properties of eqs.(12) and (13), define $\lambda(p, p', q)$ as:

$$\lambda(p, p', q) = \rho[\Lambda(p, p', q)] \quad (16)$$

We refer to $\lambda(p, p', q)$ as the **predictive looseness** of the model specified by q , $\widetilde{M}(p, q)$.

When $p' = p$, then $\lambda(p, p', q)$ is the looseness in predicting the outcomes, in response to inputs q , in the configuration at which the model was up-dated. This is the most common situation in system modelling. When $p' \neq p$ then $\lambda(p, p', q)$ is the looseness of what is properly called a forecast from one configuration to another.

4.2 The Theorem

Large fidelity-robustness, $\widehat{\alpha}(q, r_c; p)$, and small prediction-looseness, $\lambda(p, p', q)$, are both desirable. We will say that robustness and looseness are **sympathetic** if a change in p or q improves them both; otherwise they are **antagonistic**. The following theorem shows that, under fairly weak conditions, robustness and looseness are always antagonistic. Examples are presented in section 5.

Theorem 2 *Fidelity-robustness and prediction-looseness are antagonistic.*

Let $\mathcal{U}(\alpha, M)$ be an info-gap model which obeys the axioms of nesting, contraction and translation, and let $\widehat{\alpha}(q, r_c; p)$ be its robustness function. For any models q^ and q' and any configurations p^* , p' and p :*

$$\text{If } \widehat{\alpha}(q^*, r_c; p^*) > \widehat{\alpha}(q', r_c; p') \quad (17)$$

$$\text{Then } \Lambda(p^*, p, q^*) \supseteq \Lambda(p', p, q') + \mu \quad (18)$$

$$\text{and } \lambda(p^*, p, q^*) \geq \lambda(p', p, q') \quad (19)$$

where λ and Λ are related by eq.(16) by the $\rho(\cdot)$ function, and $\mu = \widetilde{M}(p, q^*) - \widetilde{M}(p, q')$.

The supposition in relation (17) asserts that model q^* is more robust (at configuration p^*) than is model q' (at configuration p' , which may be the same as p^*). Relations (18) and (19), which are implied by relation (17), assert that q^* makes looser predictions than q' at any configuration p (which may be the same as either p^* or p' , or may differ from both). In other words, the change from (q', p') to (q^*, p^*) enhanced the robustness but decreased the predictive focus at p .

Proof. Define the following concise notation. $\alpha^* = \widehat{\alpha}(q^*, r_c; p^*)$, $\alpha' = \widehat{\alpha}(q', r_c; p')$, $M^* = \widetilde{M}(p, q^*)$, $M' = \widetilde{M}(p, q')$.

Supposition (17) and the axiom of nesting, relation (7), imply:

$$\mathcal{U}(\alpha^*, M^*) \supseteq \mathcal{U}(\alpha', M^*) \quad (20)$$

The axiom of translation, relation (9), implies:

$$\mathcal{U}(\alpha', M^*) = \mathcal{U}(\alpha', M') + \{-M'\} + \{M^*\} \quad (21)$$

Combining relations (20) and (21) we obtain:

$$\mathcal{U}(\alpha^*, M^*) \supseteq \mathcal{U}(\alpha', M') + \{-M'\} + \{M^*\} \quad (22)$$

From eq.(15) we see that $\Lambda(p^*, p, q^*) = \mathcal{U}(\alpha^*, M^*)$ and $\Lambda(p', p, q') = \mathcal{U}(\alpha', M')$. Thus (22) is precisely (18).

Eq.(22), together with the properties of translation invariance and nesting, eqs.(12) and (13), imply:

$$\rho[\mathcal{U}(\alpha^*, M^*)] \geq \rho[\mathcal{U}(\alpha', M')] \quad (23)$$

With the definition of predictive looseness in eq.(16), this implies eq.(19). This completes the proof. ■

4.3 The Dilemma

Three quantities are central to the info-gap analysis of modelling and forecasting: fidelity of the model to the test data, r_c ; robustness, to model-uncertainty, of the fidelity, $\hat{\alpha}(q, r_c; p)$; and prediction-looseness $\lambda(p, p', q)$. Two trade-offs relate these quantities.

- Robustness decreases as fidelity improves: $\hat{\alpha}(q, r_c; p)$ gets smaller as r_c gets smaller (theorem 1).
- Robustness decreases as looseness improves: $\hat{\alpha}(q, r_c; p)$ gets smaller as $\lambda(p, p', q)$ gets smaller (theorem 2).

These trade-offs imply that it is not possible to simultaneously increase the robustness, increase the fidelity, and decrease the prediction looseness. This is illustrated schematically in fig. 1. The right quadrant shows the trade-off between robustness to model uncertainty $\hat{\alpha}$ and fidelity to test data r_c , as asserted by theorem 1: poor fidelity (large r_c) implies good robustness (large $\hat{\alpha}$). The left quadrant portrays theorem 2. The horizontal axis to the left portrays the predictive looseness λ and the curve shows that the predictive looseness and the robustness increase together. Poor fidelity r_1 has high robustness $\hat{\alpha}_1$ and large predictive looseness λ_1 . Good fidelity r_2 has low robustness $\hat{\alpha}_2$ and small predictive looseness λ_2 . Fidelity r_c and predictive looseness λ are sympathetic: they improve or deteriorate together. But they are both antagonistic to robustness $\hat{\alpha}$, so that good (small) values of r_c and λ are unreliable when associated with poor (small) values of $\hat{\alpha}$.

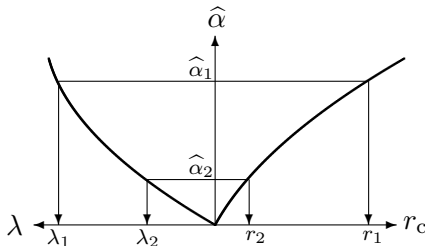


Figure 1: Schematic illustration of antagonism between fidelity r_c , robustness $\hat{\alpha}$ and predictive looseness λ .

High fidelity (small r_c) of model $\tilde{M}(p, q)$ implies that the model is true to the measurements, which adds warrant to the model. Low fidelity means that the model is not responsive to the measurements.

Large robustness (large $\hat{\alpha}(q, r_c; p)$) of model $\tilde{M}(p, q)$ means that a wide selection of models $M(p, q)$ around the model $\tilde{M}(p, q)$ have fidelity to the test data no worse than r_c . That is, $\tilde{M}(p, q)$ could be modified greatly without diminishing the fidelity. Hence, if r_c represents high fidelity, then large robustness strengthens belief in the validity of the model, $\tilde{M}(p, q)$, because the fidelity of $\tilde{M}(p, q)$ is immune to errors and imperfections in its formulation. A small value of robustness impugns the model because, even if r_c corresponds to high fidelity, this trueness to the data may be the accidental result of the specific erroneous structure of the model. In a perfect world, free of

info-gaps, robustness would not be important. The importance of robustness, for warrant of model $\widetilde{M}(p, q)$, derives from the analyst’s info-gaps which induce a lack of confidence that the structure of $\widetilde{M}(p, q)$ is fundamentally correct. If $\widetilde{M}(p, q)$ is robust, then its imperfections, whatever they may be, are only marginally important since they do not seriously diminish the model’s fidelity to the data.

Small predictive looseness (small $\lambda(p, p', q)$) implies that all the models which are equivalent to $\widetilde{M}(p, q)$ in terms of satisficing the fidelity to the test data, also agree in their predictions of the system behavior in configuration p' . A large value of looseness means that fidelity-equivalent models strongly disagree in their predictions of the system behavior.

The dilemma. ‘Truth’ is a difficult philosophical concept. Nonetheless, at least from a pragmatic point of view, FIDELITY to data is necessary (though not sufficient) for trueness of the model. ROBUSTNESS to model uncertainty indicates trueness of fidelity. LOOSENESS of model prediction increases as fidelity-robustness to model-uncertainty improves.

The dilemma results from the conflict between two uncertainties: spread of the data (calibration and measurement errors, and the experimental variability due to lack of control of the experiment) and epistemic limitation (imperfect understanding of the process). The need for fidelity arises from spread of the data, while the need for robustness arises from epistemic uncertainty about the structure of the model. We explore this dilemma further in section 6.

The dilemma may be resolved, in practical applications, by making value judgments of various sorts. The question ‘How robust is robust enough?’ calls for judgments like those elicited by the question ‘How safe is safe enough?’. Those value judgments can be based on formal tools such as analogical reasoning (Ben-Haim, 2006, chapter 4), by appeal to guidelines, or with informal reasoning. Determination of acceptable fidelity may be assisted by probabilistic considerations if the fidelity relates to aleatory uncertainty of the measurements. Determining acceptable prediction looseness will often be supported by the implications of prediction error. Finally, judgments must balance the three attributes. A vast array of tools are available for these judgments, including formal reasoning (Fagin *et al.* 1995), artificial intelligence (Lawry, 2006), and expert elicitation (Meyer and Booker, 2001). Further exploration of these issues is beyond the scope of this paper.

5 Example: A Non-Linear System

In this section we will consider a system which is modelled as a linear input-output relation in which uncertain quadratic terms are ignored. The model may represent a mechanical force-displacement relation, or an economic Phillips curve relating inflation to unemployment, or the biological relation between the rate of flow of river water and the spawning success of fish, and so on. The example will demonstrate the trade off asserted by theorem 2, and the violation of the trade off when the conditions of the theorem are not satisfied.

5.1 Formulation

Consider a multi-input single-output system subject to a vector $p = (p_1, \dots, p_K)^T$ of inputs. The scalar output is y . Our current best estimate of the input-output relation is a linear model $\widetilde{M}(p, q)$ where q represents system properties. Specifically, $\widetilde{M}(p, q) = p^T q$ where $q = (q_1, \dots, q_K)^T$ is a vector of model coefficients.

Let us consider unmodelled quadratic terms which may in fact be present in the input-output relation:

$$M(p, q) = \widetilde{M}(p, q) + \sum_{i=1}^K Q_i p_i^2 = p^T q + \pi^T Q \quad (24)$$

where $\pi = (p_1^2, \dots, p_K^2)^T$ and $Q = (Q_1, \dots, Q_K)^T$.

We have no direct data on Q for the system of interest, but evidence from other somewhat similar systems indicates that Q is typically small on average but can vary substantially. Let us define S^{-1} as a real, symmetric, positive definite matrix specifying a K -dimensional ellipsoid which approximates the cluster of the observed or suspected Q vectors of the related systems. For instance, S^{-1} may be the inverse of the population covariance matrix S of observed Q vectors. Alternatively, S^{-1} may be chosen to approximate experts' opinions of the dispersion of the Q vectors. In any case, since the data are very scarce, and not derived from the system which is actually of interest, it is possible that values of Q well outside this ellipsoid can occur for the system being analyzed. Furthermore, the evidence is insufficient to verify the choice of a probability distribution for Q .

In light of these considerations we represent the uncertainty in the input-output relation, arising from the uncertain quadratic terms, by the following ellipsoid-bound info-gap model:

$$\mathcal{U}(\alpha, \widetilde{M}) = \left\{ M(p, q) = p^T q + \pi^T Q : Q^T S^{-1} Q \leq \alpha^2 \right\}, \quad \alpha \geq 0 \quad (25)$$

This info-gap model is an unbounded family of nested sets $\mathcal{U}(\alpha, \widetilde{M})$, $\alpha \geq 0$, where each $\mathcal{U}(\alpha, \widetilde{M})$ is a set of input-output models $M(p, q)$. Two levels of uncertainty are entailed in this info-gap model. At any horizon of uncertainty, α , the specific realization Q is unknown. In addition, the horizon of uncertainty, α , is unknown.

The info-gap model $\mathcal{U}(\alpha, \widetilde{M})$ in eq.(25) obeys the axioms of nesting and contraction. Note that π depends on p . Thus $\mathcal{U}(\alpha, \widetilde{M})$ obeys the axiom of translation only if the configuration p is fixed. Consequently we can apply theorem 2 only if we hold p constant. That is, for the current example, theorem 2 implies that, *if* model q^* is more robust than model q' in the same configuration p , *then* the prediction looseness of model q^* exceeds the prediction looseness of model q' . In this example we will simplify our notation and define $\lambda(p, q) = \lambda(p, p, q)$ and $\Lambda(p, q) = \Lambda(p, p, q)$.

5.2 Prediction Looseness

In the present example the set $\Lambda(p, q)$, defined in eq.(15), is the interval from the least to the greatest predicted values, as the model $M(p, q)$ varies continuously on the uncertainty set whose horizon of uncertainty equals the robustness, $\mathcal{U}[\widehat{\alpha}(q, r_c; p), \widetilde{M}]$. As in eq.(14), the following function satisfies eqs.(12) and (13):

$$\lambda(p, q) = \max_{M \in \Lambda(p, q)} M - \min_{M \in \Lambda(p, q)} M \quad (26)$$

We are thus entitled to adopt this function as the predictive looseness in this example. It is readily shown, using Lagrange optimization, that the looseness is:

$$\lambda(p, q) = 2\widehat{\alpha}(q, r_c; p)\sqrt{\pi^T S \pi} \quad (27)$$

As anticipated by theorem 2, the looseness and robustness are antagonistic (at fixed p): any change in the model coefficients q which improves the robustness (makes $\widehat{\alpha}$ larger) causes the prediction looseness to deteriorate (makes λ larger as well).

Eq.(27) indicates that a plot of λ vs. $\widehat{\alpha}$ is a line of positive slope through the origin, generated by varying the model characteristics q , as shown in fig. 2. The slope, $2\sqrt{\pi^T S \pi}$, can be thought of as the cost (in terms of lost predictive focus) of a unit increase in robustness. A large slope means that a small improvement in robustness to uncertainty entails a large increase in the range of predictions of models whose fidelity to the test data is no worse than r_c .

The slope in fig. 2 is large if π , the squared inputs, are large, or if the population covariance matrix S , is large. 'Large S ' means, roughly, that the measurements are widely dispersed. Large

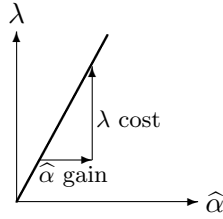


Figure 2: Prediction looseness λ vs. robustness $\hat{\alpha}$.

measurement variability makes the attainment of robustness to epistemic uncertainty very costly in terms of predictive focus.

Also interesting is the interaction between π and S and its influence on the trade-off between looseness and robustness. Let s_1 and s_K be the least and greatest eigenvalues of S . ($s_1 > 0$ since S is positive definite.) Then one finds that:

$$s_1 \leq \frac{\pi^T S \pi}{\|\pi\|^2} \leq s_K \quad (28)$$

This demonstrates how the inputs π and model-uncertainty S combine to determine the trade-off between prediction-looseness and robustness. Squared-input vectors π which are oriented along eigenvectors of S with large eigenvalues (near s_K) will entail large slope of λ vs. $\hat{\alpha}$. On the other hand, input vectors which “activate” relatively certain modes of S (those with small eigenvalues near s_1) will induce lower looseness-costs for model-robustness.

Relations (28) have practical implications for modelling the behavior of the system. Let V_1 be the sub-space spanned by eigenvectors of S having small eigenvalues, and let V_2 be the complementary sub-space. If the relevant inputs lie in V_1 , then the model can be robustified more than if the relevant inputs lie in V_2 , at the same level of prediction looseness. This implies that the analyst’s judgments regarding achievable robustness and its cost in prediction-looseness, will be influenced by prior judgments about the range of anticipated inputs. This in turn may influence the choice of the model, q .

5.3 Sympathetic Robustness and Looseness

As noted above, the info-gap model in eq.(25) does not obey the axiom of translation if the input vector p varies. Consequently, theorem 2 does not apply to changes in the input vector. In fact, we will show that, in a simple realization of the current example, prediction-looseness and robustness can be sympathetic with respect to changes in the input: a change in p which improves the robustness (causes $\hat{\alpha}$ to increase), may also improve the prediction looseness (causes λ to decrease).

For simplicity, let p and q be scalars, so a single input p is applied to a system with coefficient q (which we assume is positive). In the info-gap model of eq.(25) Q is a scalar and $S^{-1} = 1/s^2$ where s is an estimate of the variation of the Q values. Define $\tilde{R} = \frac{1}{N} \sum_{i=1}^N (y_i - pq)^2$, which is the mean squared error of the nominal model, $\tilde{M} = pq$. Define $F = \frac{1}{N} \sum_{i=1}^N (y_i - pq)$, which is the average error of \tilde{M} .

The robustness, defined in eq.(4), is found to be:

$$\hat{\alpha}(q, r_c; p) = \begin{cases} \frac{1}{sp^2} \left(-|F| + \sqrt{F^2 - \tilde{R} + r_c} \right) & \text{if } \tilde{R} \leq r_c \\ 0 & \text{else} \end{cases} \quad (29)$$

In the particularly simple case that only a single measurement, y , is available, so that $N = 1$ and $F^2 = \tilde{R}$, the positive part of eq.(29) becomes:

$$\hat{\alpha}(q, r_c; p) = \frac{1}{sp^2} (-|pq - y| + \sqrt{r_c}) \quad (30)$$

Combining this with eq.(27) we find that the prediction looseness is:

$$\lambda(p, q) = 2(-|pq - y| + \sqrt{r_c}) \quad (31)$$

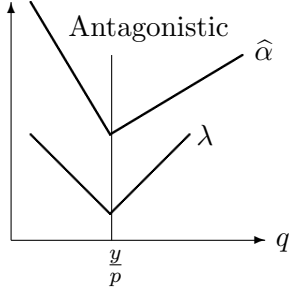


Figure 3: Robustness $\hat{\alpha}$ and prediction looseness λ vs. flexibility q .

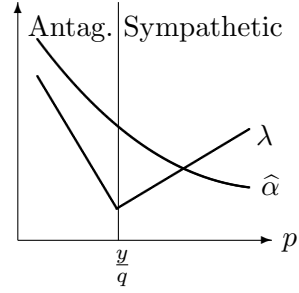


Figure 4: Robustness $\hat{\alpha}$ and prediction looseness λ vs. input p .

Comparing eqs.(30) and (31) we see that robustness $\hat{\alpha}$ and looseness λ are *antagonistic* with respect to the model parameter q : a change in q which augments robustness also increases the looseness, as anticipated by theorem 2. This is illustrated in fig. 3, where we see that the slopes of $\hat{\alpha}$ and λ always have the same sign. The difference in the magnitudes of these slopes determines the trade-off cost: how much robustness must be relinquished in order to reduce the prediction looseness, or vice versa.

If the conditions of theorem 2 are not satisfied, then robustness and looseness can be *sympathetic* with respect to change in the input p . This means that there is a range of p values for which a change in p improves both the robustness and the looseness: augments $\hat{\alpha}$ and diminishes λ . Fig. 4 shows that the slopes of $\hat{\alpha}$ and λ have opposite signs for $p > y/q$. However, if $p < y/q$ then robustness and looseness are antagonistic with respect to p ; the slopes of $\hat{\alpha}$ and λ have the same sign. Once again the difference between the slopes determines the trade-off cost.

6 Conclusion: Methodological Issues

In this section we will raise some methodological issues related to the problems of measurement, modelling and prediction which are dealt with in this paper. In section 6.1 we discuss the relation between our results and the legacy of scientific positivism. In section 6.2 we identify some questions which our results motivate regarding induction, warrant and prediction.

6.1 Positivism, Prediction and Uncertainty

In this paper we have dealt with severe uncertainty and its implications for modelling and predicting complex behavior. We have stressed that models err not only because the test data with which

they are calibrated are noisy, but because our choice of model structure is based on incomplete or even fallacious understanding of the processes being modelled.

Knight recognized uncertainty of this sort in economics, calling it “true uncertainty” as opposed to probabilistic risk for which test data and understanding are more extensive (Knight, 1921). Knightian uncertainty is characteristic of highly complex and dynamic systems.

The sort of complex dynamism which induces Knightian uncertainty was identified by Popper as causing ‘indeterminism’. For Popper, indeterminism arises in intelligent learning systems (Popper, 1982, pp.62–63). By ‘intelligence’ is meant the ability to adapt behavior to experience, regardless of whether this is done willfully or not. By ‘learning’ is meant the discovery of facts or situations which previously were unknown. A discovery tomorrow cannot be known today. Consequently, the behavior tomorrow of an intelligent learning system cannot be entirely predicted today because of the behavioral impact of tomorrow’s discoveries. Hence, all intelligent learning systems are accompanied by an inherent indeterminism, an inherent recalcitrance to law-like or predictive modelling. Shackle writes similarly of the “dissolution of rational determinacy” (Shackle, 1972, chap. 22).

Hayek also focussed on the limitations of human understanding of complex systems, especially highly dispersed systems such as, but not limited to, economic markets. It is possible, Hayek contended, to predict tendencies or trends, and to identify principles governing these patterns. However, it is impossible to consistently and reliably forecast or explain the detailed evolution of such systems (Caldwell, 2004, pp. 339–340, 346). Poole concurs with Hayek’s scepticism and with Shackle-Popper indeterminism:

[A]nyone interested in monetary policy should spend less time on economic forecasts and more time on implications of forecast surprises. . . . [P]olicy needs to be informed by the best guesses incorporated in forecasts and by knowledge of forecast errors. . . . And the surprises that create forecast errors also create the need for policy changes that cannot be anticipated in advance because the surprises cannot be anticipated. (Poole, 2004, pp. 1, 4)

Our example in section 5 illustrates these limitations in forecasting power for a class of models used widely in engineering, economics and elsewhere.

The upshot of Hayek’s position is a rejection of the strong scientific optimism which characterized much activity in social and technological disciplines in the 19th and 20th centuries. This optimism, sometimes referred to as positivism, is the conviction that thorough application of scientific understanding will lead to better control and utilization of any system—technical or human—whatever the goals may be. Positivism underlies activities such as central economic planning over the full gamut from socialism to the modern welfare state, and deferential acceptance of technological and medical prognostication. Knight’s uncertainty, Popper’s indeterminism and Hayek’s scepticism all stand in stark challenge to the positivists’ optimism about the predictive power of models.

There is no doubt that quantitative models sometimes do make impressively accurate and useful predictions concerning highly irregular occurrences. For instance, Bayesian models have been used to anticipate severe forest fires (Beckage and Platt, 2004). Severe events are, by definition, rare and unusual, and the ability to use historical data to predict seasons or years in which severe fires will occur is a welcome success of advanced modelling technique. By identifying patterns in related meteorological events, the model predicts with admirable accuracy the propensity for greatly abnormal fires in subsequent periods of 3 to 12 months. However, in contrast to this successful prediction of a tendency, forecasting the precise acreage of fire devastation is far from reliable. As Hayek anticipated, identification of the main factors governing a complex phenomenon (e.g. forest fires) leads to useful predictions of trends, but not to successful detailed quantitative forecasts.

The theorems presented in this paper have a decidedly Hayekian flavor, and are based on info-gap quantification of Knightian uncertainty. Models can be faithful to experience and can indicate future trends. However, complex processes may be incompletely comprehended. When this is the case, theorem 1 states that model-fidelity is lost as we increase the robustness to our ignorance of that complexity. Furthermore, theorem 2 asserts that, as we enhance our robustness to ignorance, we lose predictive focus; models whose fidelity is reliable will tend to be poor predictors if info-gaps accompany our understanding. Combining these constraints we find that, for complex and imperfectly understood processes, we will face challenges when models are expected to do more than to predict general trends. Hayekian scepticism about optimistic positivism seems well founded in light of our theoretical results. On the other hand, the robustness and looseness functions developed in this paper provide quantitative tools for assisting the analyst in identifying models which have adequate fidelity, robustness and predictive focus.

6.2 Induction, Warrant and Prediction

Hume and the problem of induction. The conflict between robustness, fidelity and prediction looseness is reminiscent of Hume's critique of empirical induction (Hume, 1777). Our analysis shows that past measurements, accompanied by incomplete understanding of the measured process, cannot unequivocally establish true predictions of the behavior of the system. However, the issue is not inherently temporal. The problem is not that *past* measurements do not logically bind *future* behavior. The problem is epistemic. We use test data to select a model of system behavior. But data alone are insufficient for the selection process, which depends also on the analyst's understanding of the measured process. Experience shows that this understanding is incomplete. Also, like Hume, we claim that no amount of data can confirm the contention that no aspect of the process has been overlooked.

Hume's critique is clearly related to the antagonism between robustness and prediction looseness. However, can this antagonism be reduced to the classical problem of scientific induction? To explore this question further we consider the issue of warrant for a proposition.

Haack, warrant and prediction. Our main result, theorem 2, establishes general conditions under which a physical model can be highly warranted by virtue of high fidelity to test data and high robustness to epistemic uncertainty, only by limiting its predictive power. The questions we wish to raise (but not answer) are: Is 'warrant' overrated as a tool for model selection? Philosophers have been greatly interested in model selection in natural science. Is model selection in natural science fundamentally different from model selection in utilitarian disciplines such as engineering, economics, medicine, or law? Robust fidelity is anti-correlated with predictive power. Can concepts of warrant for a model be modified to correlate better with predictive power?

Our only contribution to these questions is to explain that robust fidelity can be understood at least partially as a type of warrant within the context of one of the most attractive of contemporary epistemological theories of warrant.

Susan Haack, in expounding her epistemological theory of 'foundherentism' (Haack, 1993), explains that

how warranted an empirical claim is depends on how well it is supported by experiential evidence and background beliefs, how reasonable those background beliefs are, independent of the belief in question, and how much of the relevant evidence the evidence includes.

... Briefly and very roughly, how well evidence supports a claim depends on how well the claim is explanatorily integrated with the evidence. (Haack 2001, p.11)

The relation between foundherentism as a holistic theory of warrant, and info-gap robust-satisficing as a methodology for severely testing hypotheses, is not simple or unambiguous (Ben-Haim, 2006), pp.332–333]. However, the following claims seem reasonable:

- If physical model \tilde{M} has high fidelity to the test data, then it is “supported by experiential evidence”. All “relevant evidence” has been included if any outliers in the data have been excluded as irrelevant only on substantive evidential grounds.

- The concept of “background beliefs” can be interpreted in the present context to refer to the array of competing models which collectively represent the universe of potential and actual understandings of the processes involved. Consequently, if the robustness of \tilde{M} to epistemic uncertainty is large, meaning that \tilde{M} could be modified in many ways without significantly diminishing its fidelity, then \tilde{M} is “supported by . . . background beliefs”.

- The physical models M contained in the info-gap model $\mathcal{U}(\alpha, \tilde{M})$ at low α are quite similar to, and thus “dependent” on, \tilde{M} . However, at large horizon of uncertainty, the set $\mathcal{U}(\alpha, \tilde{M})$ contains an infinity of models which differ greatly from the “centerpoint” model \tilde{M} . Thus, at high robustness, many models lie in the universe of discourse that are highly unrelated to, and thus “independent” of, \tilde{M} .

When modelling imperfectly understood processes, the upshot of all this is that a model that is warranted by large robustness at high fidelity, seems to satisfy Haack’s criteria for epistemological warrant. But high fidelity at high robustness is anti-correlated with high predictive power (under the stipulations of theorem 2). What good is robust fidelity in particular, and epistemological warrant in general, if predictive power is not only not entailed, but in fact counter-indicated?

Acknowledgement

The authors are indebted to Prof. Susan Haack for her stimulating correspondence.

7 References

1. Banks, H.T. and C. Castillo-Chavez, eds., 2003, *Bioterrorism: Mathematical Modeling Applications in Homeland Security*, Frontiers in Applied Mathematics 28, SIAM, Philadelphia.
2. Bårdsen, Gunnar, Øyvind Eitrheim, Eilev S. Jansen, and Ragnar Nymoen, 2005, *The Econometrics of Macroeconomic Modelling*, Oxford University Press.
3. Beckage, Brian and William J. Platt, 2004, Predicting severe wildfire years in the Florida Everglades, *Frontiers in Ecology and the Environment*, vol. 1, #5, pp.235–239.
4. Ben-Haim, Yakov, 1996, *Robust Reliability in the Mechanical Sciences*, Springer-Verlag, Berlin.
5. Ben-Haim, Yakov, 2000, Robust rationality and decisions under severe uncertainty, *Journal of the Franklin Institute*, 337: 171–199.
6. Ben-Haim, Yakov, 2006, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty*, 2nd edition, Academic Press, London.
7. Ben-Haim, Yakov and Francois Hemez, 2004, Robustness, fidelity, and prediction-looseness of models, ASME Engineering Systems Design and Analysis Conference, Manchester, UK, 19–22 July 2004.
8. Ben-Tal, A. and A. Nemirovski, 1999. Robust solutions of uncertain linear programs, *Oper. Res. Lett.* 25, 1–13.

9. Berger, James O., 1980, *Statistical Decision Theory: Foundations, Concepts, and Methods*, Springer Verlag, New York.
10. Burgman, Mark, 2005, *Risks and Decisions for Conservation and Environmental Management*, Cambridge University Press, Cambridge.
11. Caldwell, Bruce, 2004, *Hayek's Challenge: An Intellectual Biography of F.A. Hayek*, University of Chicago Press, Chicago.
12. Doebling, S.W., 2002, Structural Dynamics Model Validation: Pushing the Envelope, Intl. Conf. on Structural Dynamics Modeling: Test, Analysis, Correlation and Validation, Madeira Island, Portugal, June 3–5, 2002.
13. Fagin, Ronald and Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi, 1995, *Reasoning About Knowledge*, MIT Press.
14. Friswell, M.I. and J.E. Mottershead, 1995, *Finite Element Model Updating in Structural Dynamics*, Kluwer Academic Publishers, Dordrecht.
15. Haack, Susan, 1993, *Evidence and Inquiry: Towards Reconstruction in Epistemology*. Blackwell.
16. Haack, Susan, 2001, An Epistemologist in the Bramble-Bush: At the Supreme Court with Mr. Joiner. *Journal of Health Politics, Policy and Law*, 26: 217–248.
17. Hites, R., De Smet, Y., Risse, N., Salazar-Neumann, M., Vincke, P., 2006, About the applicability of MCDA to some robustness problems, *Eur. J. Oper. Res.* 174, 322–332.
18. Hume, David, 1777, *An Inquiry Concerning Human Understanding*.
19. Kennedy, Marc C. and Anthony O'Hagan, 2001, Bayesian calibration of computer models, *J. Royal Statistical Soc.*, B, 63 (Part 3): 425–464.
20. Knight, Frank H., 1921, *Risk, Uncertainty and Profit*, Houghton Mifflin Co., re-issued by University of Chicago Press, 1971.
21. Lawry, Jonathan, 2006, *Modelling and Reasoning with Vague Concepts*, Springer, New York.
22. Meyer, Mary A. and Jane M. Booker, 2001, *Eliciting and Analyzing Expert Judgment: A Practical Guide*, ASA-SIAM, New York.
23. Natke, H.G. and C. Cempel, 1997, *Model-Aided Diagnosis of Mechanical Systems*, Springer, Berlin, New York.
24. Poole, William, 2004, Best guesses and surprises, *Federal Reserve Bank of St. Louis Review*, Jan/Feb 2004, 86(3), pp.1–7.
25. Popper, Karl, 1982, *The Open Universe: An Argument for Indeterminism*. From the Postscript to *The Logic of Scientific Discovery*. Routledge.
26. Shackle, G.L.S., 1972, *Epistemics and Economics: A Critique of Economic Doctrines*, Transaction Publishers 1992, originally published by Cambridge University Press, 1972.
27. Wald, A., 1945, Statistical decision functions which minimize the maximum risk, *Annals of Mathematics*, 46(2), 265–280.