The Worst Scenario Method: A Red Thread Running Through Various Approaches to Problems with Uncertain Input Data

J. Chleboun

Abstract Three ingredients constitute mathematical models dependent on parameters whose value is uncertain: a compact set $\mathscr{U}_{\mathsf{ad}}$ of admissible parameters a, a state problem $A(a)u = f(a)$ with an a-dependent state $u \equiv u(a)$, and a continuous quantity of interest $\Psi(a) = \Phi(a, u(a))$. In the worst scenario method (WSM), the maximum of Ψ over $\mathscr{U}_{\mathsf{ad}}$ is identified. By mastering the WSM and if an adequate characterization of input uncertainty is available, the analyst can easily step forward to a more complex uncertainty analysis, namely that based on the Dempster-Shafer theory or fuzzy set theory. Elements of the above non-stochastic approaches to uncertainty modeling are presented with the emphasis on uncertain functions appearing in problems driven by differential equations.

1 Introduction

Since uncertainty in input parameters accompanies most, if not all, mathematical and computational models, its impact on model outputs deserves attention. We will focus on the worst scenario method (WSM) that can be applied as a stand-alone method (Subsection 2.1) or used as a fundamental part of other approaches such as the Dempster-Shafer theory (Subsection 2.2) and fuzzy set theory (Subsection 2.3). That is, by mastering the WSM, the analyst can easily step forward to a more complex uncertainty analysis if an adequate characterization of input uncertainty is available. Attention is paid to uncertain functions appearing in problems driven by differential equations (Section 3). The goal of this paper is two-fold: (A) to provide the reader with an insight into non-stochastic uncertainty modeling, and (B) to show the reader how non-stochastic uncertainty in input functions can be treated.

Jan Chleboun
Institute of Mathematics, Academy of Sciences, Žitná 25, 115 67 Praha 1,
and Faculty of Engineering, Czech Technical University, Karlovo nám. 13, 121 35 Praha 2,
Czech Republic, e-mail: chleb@math.cas.cz

Although other sources aiming at (A) can be found in the literature, (B) seems to be a rather uncommon subject.

The assessment of uncertainty in data is, essentially, equivalent to the weighting of data. Consequently, as uncertainty propagates through a model, the model outputs are also weighted and the determination of these weights counts among the analyst's ultimate goals. Different weighting approaches result in different methods or even theories.

Stochastic methods stem from weighting the values of input parameters by the probability of their occurrence. Stochastic methods can yield strong results but the analyst should be aware of the fact that they also assume rather strong input information such as the probability distribution of uncertain input parameters and a possible correlation between them, for example. Such information is not always available or it is itself highly uncertain. If this is the case, other methods of weighting input data can be more appropriate, reliable, and realistic.

2 Non-Stochastic Methods

Three representatives of non-stochastic methods will be introduced. Let us start with the basic mathematical framework that will be shared by all the presented methods:

(a) Let the state problem be represented by $A(a)u = f(a)$, an a-dependent equation where a is an input parameter. The existence and uniqueness of the state solution $u \equiv u(a)$ is assumed for any a considered.

(b) Let the a-dependent solution $u(a)$ be evaluated by $\Phi(a, u(a))$, a real-valued criterion-functional often called the quantity of interest that can directly depend on a. Owing to the uniqueness of $u(a)$, the criterion-functional Φ gives rise to the criterion-functional $\Psi(a) = \Phi(a, u(a))$. It is assumed that both u and Ψ depend continuously on a.

Both (a) and (b) deserve a few comments. State problems are not limited to equations; variational inequalities, for instance, are also possible; see [16]. The parameter a can be a scalar, a vector, a tensor, a function, or an n-tuple of functions.

The criterion-functional can represent quantities such as local temperature, local stress invariants, potential energy, or the distance between u and an *a priori* given function.

To illustrate (a) and (b), let us consider a steady heat flow problem depending on a thermal conductivity coefficient a; see also (11)-(13). The state equation (together with relevant boundary conditions) determines the temperature field $u(a)$ in the problem domain. Let $\Psi(a)$, the a-dependent quantity of interest, be defined as an average temperature in a small fixed subdomain; see (14). A change in a can cause a change in $u(a)$ and $\Psi(a)$.

2.1 Worst Scenario Method

It happens quite often that the parameter a cannot be uniquely determined and that we only know that a belongs to $\mathscr{U}_{\mathsf{ad}}$, a set of admissible values. These can originate from measurements or expert opinions, for instance. In other words, a is uncertain, so are $u(a)$ and $\Psi(a)$.

In the worst scenario method, the input values are not weighted. The significance of $a_1 \in \mathscr{U}_{\mathsf{ad}}$ is equal to the significance of $a_2 \in \mathscr{U}_{\mathsf{ad}}$. Given $\mathscr{U}_{\mathsf{ad}}$, the goal of the method is to find $a^0 \in \mathscr{U}_{\mathsf{ad}}$ such that

$$a^0 = \underset{a \in \mathscr{U}_{\mathsf{ad}}}{\arg\max}\, \Psi(a). \tag{1}$$

Since large values of quantities commonly used in engineering (such as mechanical stress, displacement, temperature) are usually considered dangerous, the maximum values correspond to the worst scenario that can happen among all $\mathscr{U}_{\mathsf{ad}}$-driven scenarios. Problem (1) is also known as *anti-optimization*; see [8, 9].

A slight modification of (1) leads to the best scenario problem: find $a_0 \in \mathscr{U}_{\mathsf{ad}}$ such that

$$a_0 = \underset{a \in \mathscr{U}_{\mathsf{ad}}}{\arg\min}\, \Psi(a). \tag{2}$$

It is not generally guaranteed that such a_0 and a^0 exist. If $\mathscr{U}_{\mathsf{ad}}$ is a compact subset of a Banach space and Ψ is continuous, then a^0 and a_0 exist and, if $\mathscr{U}_{\mathsf{ad}}$ is connected, determine I_Ψ, the range of $\Psi|_{\mathscr{U}_{\mathsf{ad}}}$:

$$I_\Psi = [\Psi(a_0), \Psi(a^0)]. \tag{3}$$

From the computational standpoint, convex $\mathscr{U}_{\mathsf{ad}}$ are preferred.

The above assumptions are fulfilled in many engineering problems; see [16] for examples from heat transfer, elasticity and plasticity theory as well as other fields. A short survey of mostly PDE-oriented applications of the method appeared in [14].

2.2 Dempster-Shafer Theory

Although the range (3) is useful to know when one analyzes the impact of uncertainty in input parameters on the quantity of interest, the plain range is dissatisfactory in many practical problems where some weights can be attributed to the input values even if these weights are not probabilistic. Then the analyst should strive for determining the weights of model outputs.

In the approach stemming from the works of Dempster and Shafer (see [6, 19]), sets are weighted. Details and examples can also be found in [1, 3], for instance.

Let us confine ourselves to the most essential ideas relevant to our purpose. We assume that U_i, where $i = 1, 2, \ldots, k$, are given convex and compact subsets (called focal elements) of a Banach space. Moreover, let each U_i have an assigned weight $m_U(U_i) > 0$ such that $\sum_{i=1}^{k} m_U(U_i) = 1$. These weights represent the information we have about U_i. Some U_i, for instance, can originate from less reliable measurements than the others. This would be indicated by the lower weights of these U_i.

By solving (1) and (2), where $\mathscr{U}_{\mathsf{ad}} = U_i$, we obtain the respective scenarios a_0^i and a_i^0. Consequently, see (3), we arrive at intervals I_Ψ^i that will constitute a new family of focal elements, now in $\mathbb{R}$, the space of real numbers. If it happens that $I_\Psi^i = I_\Psi^j$ for some $i \neq j$, the interval is considered only once; thus a family of $\hat{k}$ intervals $\hat{I}_\Psi^l$ is established, where $l = 1, 2, \ldots, \hat{k}$ and $1 \leq \hat{k} \leq k$.

The extension principle allows for deriving $m_\Psi(\hat{I}_\Psi^l)$, the weight of $\hat{I}_\Psi^l$:

$$m_\Psi(\hat{I}_\Psi^l) = \sum_{\{j \in \{1,2,\ldots,k\}:\ I_\Psi^j = \hat{I}_\Psi^l\}} m_U(U_j), \quad l = 1, 2, \ldots, \hat{k}. \tag{4}$$

The quantity $m_\Psi(\hat{I}_\Psi^l)$ can be interpreted as a measure of the amount of “likelihood” (the weight) that is assigned to $\hat{I}_\Psi^l$; see [17]. This assignment is determined by the criterion-functional Ψ and by the “likelihood” assigned to the sets U_i.

Once $m_\Psi(\hat{I}_\Psi^l)$ is determined for $l \in K = \{1, 2, \ldots, \hat{k}\}$ and $m_\Psi(\emptyset) = 0$ is defined, two mappings from subsets of $\mathbb{R}$ to the interval $[0,1]$ can be introduced. These are *Bel*, *belief*, and *Pl*, *plausibility*:

$$Bel(S) = \sum_{\{l \in K:\ \hat{I}_\Psi^l \subset S\}} m_\Psi(\hat{I}_\Psi^l), \qquad Pl(S) = \sum_{\{l \in K|\ \hat{I}_\Psi^l \cap S \neq \emptyset\}} m_\Psi(\hat{I}_\Psi^l), \qquad S \subset \mathbb{R}. \tag{5}$$

Referring to [17] again, we can interpret $Bel(S)$ as a lower bound on the likelihood of S and $Pl(S)$ as an upper bound on the likelihood of S. According to [1], $Bel(S)$ (and similarly $Pl(S)$) can also be interpreted as a lower (upper) limit on the strength of evidence at hand.

Example 1. Let us consider a loaded cantilever beam with one end fixed and the other supported by a spring whose stiffness a is uncertain and represented by five different intervals U_i with respective weights 0.1, 0.4, 0.1, 0.25, and 0.15. Let Ψ be defined as the displacement of the supported tip of the cantilever. Let $[72, 82]$, $[68, 74]$, $[73, 79]$, $[71, 83]$, and $[76, 84]$ be the respective displacement intervals I_Ψ^i determined by the worst (best) scenario problems (1)–(2) solved for $a \in U_i$, $i = 1, 2 \ldots, 5$. Then

$$m_\Psi(I_\Psi^1) = m_\Psi([72, 82]) = 0.1, \qquad m_\Psi(I_\Psi^2) = m_\Psi([68, 74]) = 0.4, \tag{6}$$

$$m_\Psi(I_\Psi^3) = m_\Psi([73, 79]) = 0.1, \qquad m_\Psi(I_\Psi^4) = m_\Psi([71, 83]) = 0.25, \tag{7}$$

$$m_\Psi(I_\Psi^5) = m_\Psi([76, 84]) = 0.15. \tag{8}$$

To analyze the uncertainty in Ψ, let us graph $Bel([x, x+d])$ and $Pl([x, x+d])$, where $d \in \{1, 2\}$ is fixed and $x \in [60, 90]$. In other words, the intervals $[x, x+d]$ chosen in the space of output data (that is, displacements) will be assessed through the evidence that we have about the input datasets. Fig. 1 shows the results for

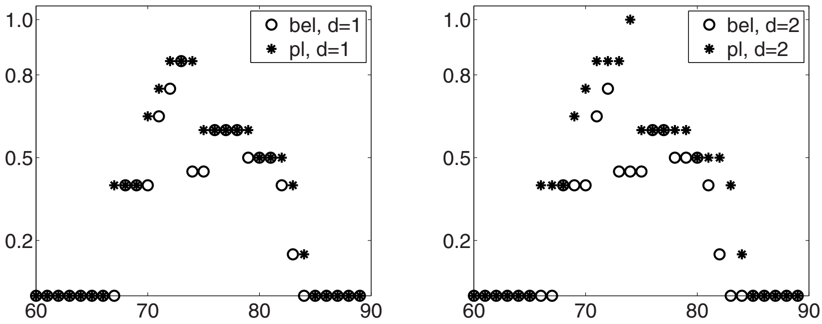

Fig. 1 Example 1; the vertical axis shows $Bel([x,x+d])$ and $Pl([x,x+d])$, the horizontal axis shows x.

$x = 60, 61, \ldots, 90$. Such graphs help the analyst to formulate a conclusion or make a decision. Thinking of the uncertain displacement magnitude in the above example, the analyst would hardly overlook the significance of values around 73, for instance.

Although the sets of scalar values were considered in this example, U_i could be sets of functions as well. Take, for instance, a set of functions representing an uncertain non-constant thickness of the beam.

2.3 *Fuzzy Set Theory*

In fuzzy set theory, points are weighted by a membership function with values in the interval $[0,1]$; see [1, 3, 7, 20, 21, 22]. For our purposes, a zero membership value will not indicate that the point does not belong to the (fuzzy) set. Indeed, we assume that a compact and convex admissible set $\mathscr{U}_{\mathsf{ad}}$ is given together with a membership function $\mu_{\mathscr{U}_{\mathsf{ad}}} : \mathscr{U}_{\mathsf{ad}} \to [0,1]$. A non-constant membership function indicates that not all members of $\mathscr{U}_{\mathsf{ad}}$ are equally possible. The higher $\mu_{\mathscr{U}_{\mathsf{ad}}}(a)$, the higher the possibility of a. We allow for $\mu_{\mathscr{U}_{\mathsf{ad}}}(a)$ to be equal to zero. Typically, $\mu_{\mathscr{U}_{\mathsf{ad}}}(a) > 0$ if a belongs to the interior of $\mathscr{U}_{\mathsf{ad}}$.

Special nested subsets of $\mathscr{U}_{\mathsf{ad}}$, called α-cuts, are defined as follows:

$$\mathscr{U}_{\mathsf{ad}}^{\alpha} = \{a \in \mathscr{U}_{\mathsf{ad}} : \ \mu_{\mathscr{U}_{\mathsf{ad}}}(a) \geq \alpha\}, \qquad \alpha \in [0,1]. \tag{9}$$

For any $\alpha \in [0,1]$, let us assume that the set $\mathscr{U}_{\mathsf{ad}}^{\alpha}$ is a convex and compact subset of $\mathscr{U}_{\mathsf{ad}}$; the compactness is guaranteed if, for instance, $\mu_{\mathscr{U}_{\mathsf{ad}}}$ is a continuous map.

By determining the best and the worst scenarios in $\mathscr{U}_{\mathsf{ad}}^{\alpha}$, we infer I_{Ψ}^{α}, the α-dependent intervals; cf. (3). These intervals are nothing else than the α-cuts of I_{Ψ}, the image of $\mathscr{U}_{\mathsf{ad}}$ under the map Ψ. To characterize the fuzziness of I_{Ψ}, the relevant membership function μ_{Ψ} is inferred (the extension principle):

$$\mu_\Psi(y) = \max\{\alpha :\ y \in I_\Psi^\alpha\}, \qquad y \in I_\Psi. \tag{10}$$

The degree of possibility of $\Psi(a)$, the a-dependent quantity of interest, is given by $\mu_\Psi(\Psi(a))$, $a \in \mathscr{U}_{\mathsf{ad}}$. A computational example will be presented later.

Remark 1. In information-gap decision theory [2], a non-fuzzy approach is introduced (besides other concepts) that also leads to the calculation of α-dependent worst scenarios. It is assumed there that α controls the amount of uncertainty present in an admissible set $\mathscr{U}_{\mathsf{ad}}^\alpha$ (α controls the "size" of the admissible set; the larger the α, the larger the size of $\mathscr{U}_{\mathsf{ad}}^\alpha$). It is also assumed that a value α exists such that $\Psi(a^0)$ determined by the worst scenario in $\mathscr{U}_{\mathsf{ad}}^\alpha$ is less than $q \in \mathbb{R}$, a given maximum acceptable value of the quantity of interest.

The goal is to find the maximum $\alpha_{\max} \in \mathbb{R}$ such that $\Psi(a^0) \le q$, where $a^0 \in \mathscr{U}_{\mathsf{ad}}^{\alpha_{\max}}$ maximizes Ψ over $\mathscr{U}_{\mathsf{ad}}^{\alpha_{\max}}$, that is, the maximum acceptable amount of uncertainty is to be identified.

3 Admissible Sets of Functions

In differential equations and the associated boundary conditions, parameters and right-hand sides often take the form of functions and are burdened with uncertainty. To introduce uncertain functions, we will present an approach stemming from the definition of admissible functions used in shape optimization; see [11].

For illustration, let us consider the following quasilinear PDE defined in Ω, a bounded domain in $\mathbb{R}^2$,

$$-\operatorname{div}(a(u)\operatorname{grad} u) = f(x,u), \tag{11}$$

$$u|_{\partial\Omega} = 0, \tag{12}$$

where a does not directly depend on $x \in \Omega$ but depends on the solution u; the right-hand side f depends both on the spatial variable x and the solution u. This boundary value problem can model a nonlinear thermal conductivity problem; we refer to [15] for a more general setting applied to modeling the temperature field in a transformer.

An admissible set $\mathscr{U}_{\mathsf{ad}}$, typical of many applications, can be defined as follows

$$\mathscr{U}_{\mathsf{ad}} = \left\{a \in \mathscr{U}_{\mathsf{ad}}^0(C_L) :\ a_{\min}(t) \le a(t) \le a_{\max}(t) \quad \forall t \in \mathbb{R}\right\}, \tag{13}$$

$$\mathscr{U}_{\mathsf{ad}}^0(C_L) = \big\{a \in C^{(0),1}(\mathbb{R}) \text{ (i.e., Lipschitz functions on } \mathbb{R}\text{)}: \\ |\mathrm{d}a/\mathrm{d}t| \le C_L \text{ a.e. in } \mathbb{R},\ a(t) = \text{const. for } t \notin [T_0, T_1]\big\},$$

where $a_{\min}, a_{\max} \in \hat{\mathscr{U}}_{\mathsf{ad}} = \left\{a \in \mathscr{U}_{\mathsf{ad}}^0(C_L) :\ 0 < a_1 \le a(t) \le a_2 < +\infty \quad \forall\, t \in \mathbb{R}\right\}$ are given functions and C_L, a_1, a_2, T_0, T_1 are given constants such that $C_L > 0$, $a_1 < a_2$, and $-\infty < T_0 < T_1 < +\infty$; see [16, 13, 4].

The criterion-functional

$$\Psi(a) = (\mathrm{meas}_2 G)^{-1} \int_G u(a)(x)\,\mathrm{d}x \tag{14}$$

represents the a-dependent temperature u averaged over a fixed set $G \subset \Omega$.

It can be proved that the worst scenario problem (1) based on (11)–(14) (where the boundary conditions can be more complex) has at least one solution; see [13, 16]. Two features of the problem are crucial for the proof: (i) $\Psi(a)$ is continuous with respect to $a \in \mathscr{U}_{\mathsf{ad}}$ and the standard norm in $C(\mathbb{R})$, the space of functions continuous on $\mathbb{R}$; and (ii) $\mathscr{U}_{\mathsf{ad}}$ is compact in $C(\mathbb{R})$ (by virtue of the Arzelà-Ascoli theorem).

Generally speaking, variants of both (i) and (ii) appear in the analysis of other worst scenario problems with uncertain functions (13) or similar; see [16] for examples from continuum mechanics (e.g., elasticity or plasticity). In (i), the continuous dependence of $u(a)$ on a is the most substantial but usually also the most demanding part of the proof. The solvability of (2) is also ensured by (i)–(ii).

Remark 2. To ensure the compactness of the admissible set $\mathscr{U}_{\mathsf{ad}}$, rather strict assumptions are employed in (13). These, however, can be too restrictive in problems where other families of input functions have to be considered (discontinuous or oscillating functions, for instance). Consequently, such an admissible set might not be compact in a standard space of functions, and its compactification in a special space is necessary. Such *relaxed* problems appear and are analyzed in optimization-oriented modeling (see [18] and the references therein) and could also be considered in uncertainty modeling.

3.1 Approximation

To solve the state problem $A(a)u = f(a)$ (imagine (11)–(12), for instance, and allow an a-dependent f), one has to resort to a numerical method such as the finite element method (FEM), the finite difference method, the boundary element method, etc. These methods deliver an approximate state solution u_h defined on a mesh characterized by $h > 0$, the discretization parameter. Let us note that the uniqueness of u_h may be an open problem in certain situations even if u is unique; see [13, 15]. Non-unique state solutions u_h can be handled under some assumptions; see [16, Chapter II]. The uniqueness of u_h is assumed henceforth.

The functions from the admissible set $\mathscr{U}_{\mathsf{ad}}$ can be approximated by continuous, piece-wise linear functions controlled by the vertical position of M nodes bound by possible constraints; see C_L, $a_{\min}$, and $a_{\max}$ in (13). These functions constitute the approximate admissible set $\mathscr{U}_{\mathsf{ad}}^M$, which is identifiable with a compact subset of $\mathbb{R}^M$.

The approximate best and worst scenario problems

$$a_{0h}^M = \arg\min_{a \in \mathscr{U}_{\mathsf{ad}}^M} \Phi(a, u_h(a)) \quad \text{and} \quad a_h^{0M} = \arg\max_{a \in \mathscr{U}_{\mathsf{ad}}^M} \Phi(a, u_h(a)) \tag{15}$$

are, in fact, finite dimensional constrained optimization problems.

The typical relationship between a_h^{0M} and a^0 (or a_{0h}^M and a_0) is as follows: If $\{a_h^{0M}\}$ is a sequence of the solutions to (15) controlled by $h \to 0+$ and $M \to \infty$, then a subsequence $\{a_{h_k}^{0M_k}\}$ exists such that, for $k \to \infty$,

$$a_{h_k}^{0M_k} \to a^0, \quad u_{h_k}(a_{h_k}^{0M_k}) \to u(a^0), \quad \text{and} \quad \Phi(a_{h_k}^{0M_k}, u_{h_k}(a_{h_k}^{0M_k}) \to \Phi(a^0, u(a^0)),$$

where the first and second sequences converge in proper spaces and topologies; see [10]. Similar convergence results for various worst scenario problems can be found in [16].

If it happens that more than one admissible set are available for the analyzed problem, say (11)-(12), and that the analyst can assess each $\mathscr{U}_{\mathsf{ad}}^i$ by $m(\mathscr{U}_{\mathsf{ad}}^i)$, the "likelihood" of $\mathscr{U}_{\mathsf{ad}}^i$, then the transition from the WSM to the Dempster-Shafer approach is straightforward.

Indeed, by finding the worst and the best scenarios, one determines the ranges (3) for each $\mathscr{U}_{\mathsf{ad}}^i$. By identifying U^i with $\mathscr{U}_{\mathsf{ad}}^i$ and obtaining m_Ψ (see (4)), the analyst is ready for the assessment of various sets $S \subset \mathbb{R}$ through (5), that is, for the assessment of the bounds of the likelihood that S is related to the uncertain values of Ψ.

Let us pay more attention to the fuzzy set approach.

3.2 Fuzzification of $\mathscr{U}_{\mathsf{ad}}$

Different concepts of fuzziness can be merged with functions see [1, Section 2.4.9]. We will simply retain $\mathscr{U}_{\mathsf{ad}}$ as a set of crisp functions but we will add a membership function to $\mathscr{U}_{\mathsf{ad}}$. In other words, we will weight $a \in \mathscr{U}_{\mathsf{ad}}$. Two forms of weighting will be introduced; see also [5].

The first approach is rather straightforward. It is based on the distance between $a \in \mathscr{U}_{\mathsf{ad}}$ and a given function a_{mid}; the details follow.

For illustration, let us recall (13) and define $a_{\text{mid}}(t) = (a_{\min}(t) + a_{\max}(t))/2$ and $a_{\text{dif}}(t) = (a_{\max}(t) - a_{\min}(t))/2$, where $t \in \mathbb{R}$. It is assumed that a_{dif} is positive on the real axis. For $\alpha \in [0,1]$, we then define

$$\mathscr{U}_{\mathsf{ad}}^\alpha = \left\{a \in \mathscr{U}_{\mathsf{ad}}^0(C_L) : \ |a(t) - a_{\text{mid}}(t)| \le (1-\alpha)a_{\text{dif}}(t) \ \forall t \in \mathbb{R}\right\}, \tag{16}$$

that is, we define the α-cuts of $\mathscr{U}_{\mathsf{ad}}$. This concept is close to fuzzy functions [1] or to controlling the amount of uncertainty through α; see [2]. Nevertheless, in (16), we still consider crisp functions. If $\alpha = 1$, then $\mathscr{U}_{\mathsf{ad}}^\alpha = \{a_{\text{mid}}\}$. If $\alpha = 0$, then $\mathscr{U}_{\mathsf{ad}}^\alpha = \mathscr{U}_{\mathsf{ad}}$.

The membership function value (the weight) of $a \in \mathscr{U}_{\mathsf{ad}}$ is defined as

$$\mu(a) = \max\{\alpha \in [0,1] : \ a \in \mathscr{U}_{\mathsf{ad}}^\alpha\}. \tag{17}$$

With this μ, definition (9) leads to $\mathscr{U}_{\mathsf{ad}}^\alpha$ defined in (16).

If $\mathscr{U}_{\mathsf{ad}}$ is fuzzy, so is $\mathscr{U}_{\mathsf{ad}}^M$. The approximate problems (15) result in optimization problems with simple bounds (determined by $a_{\min}$ and $a_{\max}$) and linear constraints (determined by C_L). The approximate best and worst scenarios in $\mathscr{U}_{\mathrm{ad}}^{M,\alpha}$, an α-cut of $\mathscr{U}_{\mathrm{ad}}^{M}$, are again obtained through solving optimization problems with simple bounds (determined by $a_{\min}$, $a_{\max}$, and α) and linear constraints (determined by C_L). Common optimization software coupled with FEM software can often be applied to solve such problems.

The other approach to weighting $\mathscr{U}_{\mathrm{ad}}$ is motivated by the observation described below. Let $a_1, a_2 \in \mathscr{U}_{\mathrm{ad}}^{\alpha}$ and let the inequality in (16) becomes the equality on the entire set $\mathbb{R}$ if a_1 is considered, and at a single point $t_0 \in \mathbb{R}$ if a_2 is considered. Moreover, let a_2 coincide with a_{mid} except for an interval containing t_0. These a_1 and a_2 share the same α-cuts of $\mathscr{U}_{\mathrm{ad}}$. In many applications, however, the weight of a_2 would be expected greater than the weight of a_1 because a_2 is “closer” to a_{mid}, which has the highest degree of possibility.

We will design a membership function able to separate a_1 from a_2. We first define an auxiliary continuous function $\rho : Q \to [0,1]$, where $Q = \{[t,y] \in \mathbb{R}^2 : \ t \in [T_0,T_1], y \in [a_{\min}(t), a_{\max}(t)]\}$. It is assumed that $\rho(t,\cdot)$ is a concave function for each $t \in [T_0,T_1]$. The functions $\rho(t,\cdot)$ can be viewed as auxiliary membership functions (weights) assessing the degree of possibility of $a(t)$ if $a \in \mathscr{U}_{\mathrm{ad}}$. The graph of $\rho(t,\cdot)$ is shaped accordingly; it is triangular or trapezoidal, which is common in fuzzy set theory. The function ρ can be derived from measurements, estimates, or expert opinions.

We are ready to define $\mu_\rho : \mathscr{U}_{\mathrm{ad}} \to [0,1]$, the membership function associated with $\mathscr{U}_{\mathrm{ad}}$:

$$\mu_\rho(a) = (T_1 - T_0)^{-1} \int_{T_0}^{T_1} \rho(t, a(t))\,\mathrm{d}t. \tag{18}$$

It is evident that we can obtain $\mu_\rho(a_1) < \mu_\rho(a_2)$ if ρ is properly shaped.

Unlike (16), the identification of all the functions a that comprise a particular α-cut is not straightforward. This difficulty also appears in the search for the approximate best and worst scenarios, where, moreover, (18) gives rise to a nonlinear constraint in the definition of $\mathscr{U}_{\mathrm{ad}}^{\alpha}$. If ρ is nonsmooth, μ_ρ is not differentiable at some a. This partial lack of differentiability is also observed in μ_{ρ^M}, a $\mathscr{U}_{\mathrm{ad}}^{M}$-related approximation of μ_ρ based on a piece-wise linear auxiliary function ρ^M that approximates ρ.

Since the use of nonsmooth (triangular, trapezoidal) $\rho(t,\cdot)$ is common and the piece-wise linearity of $\rho(t,\cdot)$ is advantageous in many respects, nonsmooth optimization seems to be unavoidable in solving (15)-like problems on the α-cuts determined by μ_{ρ^M}.

A closer inspection reveals, however, that the approximate (15)-like problems defined on $\mathscr{U}_{\mathrm{ad}}^{M,\alpha}$, the μ_{ρ^M}-based α-cuts of $\mathscr{U}_{\mathrm{ad}}^{M}$, can be decomposed into a finite sequence of smooth optimization subproblems.

Indeed, $a^M \in \mathscr{U}_{\mathrm{ad}}^{M}$ is uniquely determined by the values $a_i \equiv a(t_i)$ at fixed points t_i, where $i = 1,2,\ldots,M$. Let us assume that ρ^M is piece-wise linear and determined by the continuous functions $\rho(t_i,\cdot)$ that are linear on intervals $\theta_{ij} = [y_{i,j}, y^{i,j}]$, where $i = 1,2,\ldots,M$, $j = 1,2,\ldots,N$, and $y^{i,j} = y_{i,j+1}$ if $j = 1,2,\ldots,N-1$. It is $[y_{i,1}, y^{i,N}] =$

$[a_{\min}(t_i), a_{\max}(t_i)]$. Typically, $N = 2$ ($N = 3$) if $\rho(t,\cdot)$ is triangularly (trapezoidally) shaped.

As long as $a_i \in \theta_{ij}$ for $i = 1,2,\dots,M$ and for a *fixed* set $\mathscr{J}$ of indices j, μ_{ρ^M} is differentiable (left- and right-differentiable at the ends of θ_{ij}) and the related optimization subproblem is smooth. The differentiability is lost at one point when a_i passes from the current interval θ_{ij} to its neighbor θ_{ik}, $k \neq j$, but it is again restored if $a_i \in \theta_{ik}$ and $\mathscr{J}$ is updated. The updated set of indices determines a new smooth optimization subproblem.

The partial derivative of μ_{ρ^M} with respect to a_i, where $i = 1,2,\dots,M$, can be obtained in a closed form in each of the subproblems. Consequently, the analytic gradient of μ_{ρ^M} exists except for some points and can be employed in the calculation of $\partial\Psi/\partial a_i$, which is important in a gradient-based search for the best and worst scenarios in $\mathscr{U}_{\mathsf{ad}}^{M,\alpha}$.

Example 2. Let u, the a-dependent solution to the boundary value problem

$$-(a(x)u'(x))' = f \text{ on } \Omega = (0,1), \quad u(0) = 0 = u(1),$$

be evaluated through the criterion-functional (quantity of interest)

$$\Psi(a) = \int_\Omega (u(x) - \sin(2\pi x))^2 \,\mathrm{d}x.$$

In the state problem, f is chosen in such a way that if $a(x) = 1 + x$, then $u(x) = \sin(2\pi x)$ and, consequently, $\Psi(a) = 0$.

The parameter a belongs to the admissible set $\mathscr{U}_{\mathsf{ad}}$ determined by the quadratic function $g(x) = 1.5 + x^2$ and two constants. In detail,

$$\mathscr{U}_{\mathsf{ad}} = \left\{ a \in C^{(0),1}([0,1]) : \; |a(x) - g(x)| \leq 0.5 \text{ and } |a'(x) - g'(x)| \leq 0.8 \right\}.$$

The auxiliary function ρ is “triangular”, that is, $\rho(x,\cdot)$ is determined by the linear interpolation of the points $[x, g(x) - 0.5, 0]$, $[x, g(x), 1]$, and $[x, g(x) + 0.5, 0]$, where $x \in [0,1]$. The membership function μ_ρ is given by (18), where $T_0 = 0$ and $T_1 = 1$.

The goal is to infer μ_Ψ, the membership function of the quantity of interest; see Subsection 2.3 and (10).

To achieve the goal at least approximately, see Fig. 2, the state equation was solved by the finite element method with piece-wise linear basis functions, and $\mathscr{U}_{\mathsf{ad}}$ was approximated by continuous piece-wise linear functions constituting $\mathscr{U}_{\mathsf{ad}}^M$, where $M = 15$. The optimization problems, see (15), were solved on the α-cuts of $\mathscr{U}_{\mathsf{ad}}^M$ for $\alpha = 0, 0.05, 0.1, \dots, 1$.

The gradient of Ψ was calculated via the adjoint equation technique [12]; an explicit formula was obtained for the gradient of μ_{ρ^M} at the points of differentiability. The search for the best and the worst scenarios in the α-cuts was based on the NAG® Foundation (MATLAB®) Toolbox `E04UCF` routine for constrained sequential quadratic programming.

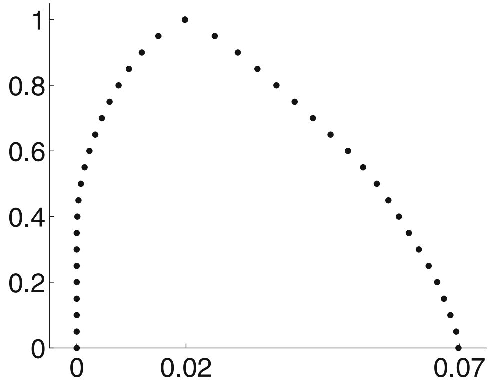

Fig. 2 Example 2. The approximation of μ_Ψ inferred from (10), where $\alpha = 0, 0.05, 0.1, \ldots, 1$. The horizontal axis shows the Ψ values, the vertical axis shows the α values. We observe that $a(x) = 1 + x$ belongs to the α-cuts if $\alpha = 0, 0.05, 0.1, \ldots, 0.35$. Indeed, for these α, the best scenario implies the zero value of Ψ. If $\alpha = 1$, then the α-cut comprises only the function g.

4 Conclusions

The worst scenario method is appropriate if we know only the set of admissible inputs but we do not have information that would enable us to weight the importance (possibility or likelihood) of input data. Since searching for the best scenario is mathematically equivalent to the worst scenario search, the WSM eventually delivers the range of the quantity of interest Ψ induced by the uncertainty in inputs.

If more extensive information on inputs is available (inputs can be weighted in some sense) and if it complies with the Dempster-Shafer or fuzzy set theory assumptions (which are less demanding than the probability theory assumptions), the uncertainty in an output quantity of interest can be weighted too. To achieve this, the WSM has to be repeatedly applied to obtain (3)-like ranges that are pivotal in the other two approaches for obtaining m_Ψ and μ_Ψ; see Subsection 2.2 and (4) as well as Section 2.3 and (10).

From the computational standpoint, solving (15)-like problems is crucial in all the above-mentioned methods. In the case of smooth problems, the gradients of both Ψ and the constraints are available, which can speed up the search for the minimum (maximum) of Ψ.

If (15) leads to a nonsmooth optimization problem, we can (a) try to decompose it to smooth subproblems, (b) use a subgradient-based technique, or (c) apply an evolution strategy that partly or completely avoids the need for the (sub)gradient.

However, it is fair to say that the worst scenario method is computationally challenging because it asks for solving a global optimization problem. Nevertheless, we can benefit from theoretical and software tools that have proved themselves well in optimal design, control theory, parameter identification, and sensitivity analysis.

Acknowledgements The research was supported by the Academy of Sciences of the Czech Republic through Institutional Research Plan No. AV0Z10190503 and grant No. IAA100190803 from the Grant Agency of AS CR.

References

1. Ayyub, B.M., Klir, G.J.: Uncertainty Modeling and Analysis in Engineering and the Sciences. Chapman & Hall/CRC, Taylor & Francis Group, Boca Raton (2006)
2. Ben-Haim, Y.: Information-gap Decision Theory. Academic Press, San Diego (2001)
3. Bernardini, A.: What are the random and fuzzy sets and how to use them for uncertainty modelling in engineering systems? In: I. Elishakoff (ed.) Whys and Hows in Uncertainty Modelling, Probability, Fuzziness and Anti-Optimization, CISM Courses and Lectures No. 388. Springer–Verlag, Wien, New York (1999)
4. Chleboun, J.: On a reliable solution of a quasilinear elliptic equation with uncertain coefficients: sensitivity analysis and numerical examples. Nonlinear Anal. Theory Methods Appl. **44**, 375–388 (2001)
5. Chleboun, J.: On fuzzy input data and the worst scenario method. Appl. Math. (Prague) **48**, 487–496 (2003)
6. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. Ann. Math. Stat. **38**, 325–339 (1967)
7. Dubois, D., Prade, H. (eds.): Fundamentals of Fuzzy Sets. Foreword by Lotfi Zadeh, *The Handbooks of Fuzzy Sets Series*, vol. 7. Kluwer Academic Publishers, Dordrecht (2000)
8. Elishakoff, I.: An idea of the uncertainty triangle. Shock Vib. Dig. **22**, 1 (1990)
9. Elishakoff, I.E., Haftka, R.T., Fang, J.: Structural design under bounded uncertainty – optimization with anti-optimization. Comput. Struct. **53**, 1401–1405 (1994)
10. Harasim, P.: On the worst scenario method: A modified convergence theorem and its application to an uncertain differential equation. Appl. Math. (2008). To appear
11. Haslinger, J., Neittaanmäki, P.: Finite Element Approximation for Optimal Shape, Material and Topology Design. J. Wiley, Chichester (1996)
12. Haug, E.J., Choi, K.K., Komkov, V.: Design Sensitivity Analysis of Structural Systems. Academic Press, Orlando (1986)
13. Hlaváček, I.: Reliable solution of a quasilinear nonpotential elliptic problem of a nonmonotone type with respect to the uncertainty in coefficients. J. Math. Anal. Appl. **212**, 452–466 (1997)
14. Hlaváček, I.: Uncertain input data problems and the worst scenario method. Appl. Math. **52**, 187–196 (2007)
15. Hlaváček, I., Křížek, M., Malý, J.: On Galerkin approximations of a quasilinear nonpotential elliptic problem of a nonmonotone type. J. Math. Anal. Appl. **184**, 168–189 (1994)
16. Hlaváček, I., Chleboun, J., Babuška, I.: Uncertain Input Data Problems and the Worst Scenario Method, *North-Holland Series in Applied Mathematics and Mechanics*, vol. 46. Elsevier, Amsterdam (2004)
17. Oberkampf, W.L., Helton, J.C., Sentz, K.: Mathematical representation of uncertainty. Research Article AIAA 2001-1645, American Institute of Aeronautics and Astronautics, Reston, VA (2001)
18. Roubíček, T.: Relaxation in Optimization Theory and Variational Calculus. De Gruyter Series in Nonlinear Analysis and Applications. Walter de Gruyter, Berlin (1997)
19. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton, NJ (1976)
20. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
21. Zadeh, L.A.: Fuzzy sets as a basis for theory of possibility. Fuzzy Sets Syst. **1**, 3–28 (1978)
22. Zimmermann, H.J.: Fuzzy Set Theory — and its Applications, fourth edn. Kluwer Academic Publishers, Boston (2001)