Robust design with imprecise random variables and its application in hydrokinetic turbine optimization

Zhen Hu, Xiaoping Du, Nitin S. Kolekar & Arindam Banerjee

Department of Mechanical and Aerospace Engineering, Missouri University of Science and Technology, Rolla, Missouri, USA
Department of Mechanical Engineering and Mechanics, Lehigh University, Bethlehem, Pennsylvania, USA

Published online: 10 May 2013.

To cite this article: Zhen Hu, Xiaoping Du, Nitin S. Kolekar & Arindam Banerjee (2013): Robust design with imprecise random variables and its application in hydrokinetic turbine optimization, Engineering Optimization, DOI: 10.1080/0305215X.2013.772603

To link to this article: http://dx.doi.org/10.1080/0305215X.2013.772603
Robust design with imprecise random variables and its application in hydrokinetic turbine optimization

Zhen Hu\textsuperscript{a}, Xiaoping Du\textsuperscript{a\ast}, Nitin S. Kolekar\textsuperscript{b} and Arindam Banerjee\textsuperscript{b}

\textsuperscript{a}Department of Mechanical and Aerospace Engineering, Missouri University of Science and Technology, Rolla, Missouri, USA; \textsuperscript{b}Department of Mechanical Engineering and Mechanics, Lehigh University, Bethlehem, Pennsylvania, USA

(Received 3 October 2012; final version received 14 January 2013)

In robust design, uncertainty is commonly modelled with precise probability distributions. In reality, the distribution types and distribution parameters may not always be available owing to limited data. This research develops a robust design methodology to accommodate the mixture of both precise and imprecise random variables. By incorporating the Taguchi quality loss function and the minimax regret criterion, the methodology mitigates the effects of not only uncertain parameters but also uncertainties in the models of the uncertain parameters. Hydrokinetic turbine systems are a relatively new alternative energy technology, and both precise and imprecise random variables exist in the design of such systems. The developed methodology is applied to the robust design optimization of a hydrokinetic turbine system. The results demonstrate the effectiveness of the proposed methodology.

Keywords: robust design; uncertainty; optimization; hydrokinetic energy

1. Introduction

Robust design (Taguchi 1993; Taguchi, Chowdhury, and Taguchi 2000) is a design methodology that determines optimal design variables so that the effects of noises (uncertainties) are minimized. If robustness is achieved, the performance of a product will not be sensitive to variations within the product or in its operating environment. As a result, a robust product can perform its intended function properly in the presence of uncertainties.

The key to robust design is the management of uncertainty. Uncertainty exists in any engineering systems. Uncertainty may come from stochastic physical nature; for example, the ultimate stress of composite materials, the temperature of an engine, and the river flow velocity of a hydrokinetic turbine are all random. Uncertainty may also come from the lack of knowledge or scarcity of data. The ignorance and mistreatment of uncertainty may lead to quality losses (Dubey and Yadava 2008) or even catastrophe (Radaev 2000). As a major approach to uncertainty initiated by Taguchi (Taguchi 1993; Taguchi, Chowdhury, and Taguchi 2000), robust design has been widely investigated and applied (Youn \textit{et al.} 2007; Choi \textit{et al.} 2008; Lu and Li 2009; Lu, Li, and Chan 2010; Ruderman \textit{et al.} 2010; Saha and Ray 2011). For instance, Ramakrishnan and Rao (1996) proposed a quality loss function (QLF) for the robust design of general design situations. Karpel, Moulin, and

\textsuperscript{\ast}Corresponding author. Email: dux@mst.edu

© 2013 Taylor & Francis
Idan (2003) applied the robust design method to the optimization of aeroservoelastic design. Du and Chen (2000, 2002) developed an efficient robust design method for multidisciplinary design optimization problems, which are subjected to uncertain parameters and uncertain models. Using the computationally expensive finite element simulations, Wiebenga, Van Den Boogaard, and Klaseboer (2012) developed a generally applicable strategy for modelling and efficiently solving robust optimization problems. Similarly, Papadimitriou and Giannakoglou (2012) applied the third sensitivity derivatives for robust aerodynamic design to account for environmental uncertainties.

In the above robust design methods, uncertain variables are usually treated as random variables, whose distributions are assumed precisely known. This kind of uncertainty is usually referred to as aleatory uncertainty (Dolšek 2012; Du 2012). Aleatory uncertainty is also called objective uncertainty, which arises from natural variability and is irreducible. Representative examples include variations in material properties and variation of product dimensions due to manufacturing imprecision. In many engineering applications, however, precise probability distributions may not always be available owing to the lack of sufficient data (Utkin 2004). As a result, there is also uncertainty in the models of the above-mentioned aleatory uncertainties. This additional or secondary uncertainty is called epistemic uncertainty (Baillie, Azzopardi, and Ruthven 2008; Du 2008; Huang and Zhang 2009; Veneziano, Agarwal, and Karaca 2009). For instance, the river velocity of the hydrokinetic turbine is associated with aleatory uncertainty because the natural variability is inherent in the river flow climate (Hu and Du 2012). But most of the time, the recorded historical river velocity data are too limited to precisely describe the river velocity with a specified probabilistic model. Instead, several hypothetical probabilistic models may be applicable for modelling the river velocity. In this situation, the model of the river velocity is also associated with epistemic uncertainty in the design of a hydrokinetic turbine system. Widely used approaches to epistemic uncertainty or imprecise random variables include interval arithmetic, fuzzy sets, imprecise probability theory, and modelling the imprecise random variables with several candidate distributions. However, in most traditional robust design methodologies, only the precise probability theory is employed.

In recent decades, much progress was made in related areas. For instance, for uncertainty analysis with imprecise random variables, Han and Jiang proposed a hybrid reliability analysis approach for problems with limited information (Jiang et al. 2012). Du (2008) developed a unified uncertainty analysis method based on the evidence theory. In the area of reliability-based design (RBD) with imprecise random variables (Walley 1991; Weichselberger 2000), Aughenbaugh and Paredis (2006) discussed the significance of using imprecise probabilities in engineering design. Based on the imprecise probability theory, Utkin (Utkin 2004; Ahmad and Kamaruddin 2012) developed a method for estimating the bounds for the structural reliability. Nikolaidis and Mourelatos (2011) applied the polynomial chaos expansion (PCE) method to the approximation of reliability upper and lower bounds for systems with imprecise random variables. Herrmann (2009) presented an approach for solving imprecise probability design optimization problems.

The imprecise probability for RBD could be extended to robust design. When the imprecise probability is applied to the robust design, however, it is difficult for decision makers to determine which probability level or hypothetical model should be employed for the optimization even if probability bounds can be provided by the p-box (Aughenbaugh and Paredis 2006). Moreover, no matter which bound or model is used for optimization, it always brings regrets to the decision maker. The reason is that another bound or model may be better than the one used (Li and Huang 2006; Conde and Candia 2007). It is therefore worthwhile to quantify the regret and ultimately minimize it during the robust optimization process. Using the minimax regret (MMR) criterion (Zhang 2011; Stoye 2012) can be a solution to the aforementioned problem.

The MMR criterion has been widely applied in policy management, optimization and decision making under uncertainty (Renou and Schlag 2010, 2011; Stoye 2011). It is known as an effective method to identify the strategy, which minimizes the maximum regret or loss when the
information on system variables is insufficient. For example, a method was proposed for a linear programming problem with interval objective function using the MMR criterion (Inuiguchi and Kume 1991; Inuiguchi and Sakawa 1995). After incorporating the MMR analysis framework into the interval-parameter programming, Li and Huang (Li and Huang 2009; Li, Huang, and Nie 2009) developed an interval MMR programming method. Chang and Davila (2007) applied the MMR analysis method to address uncertainties in waste streams in major cities and improve solid waste management strategies. From the literature reviews it can be found that most of the applications of the MMR decision criterion are limited to the policy management. The purpose of this work is to integrate the MMR criterion and robust design so that the optimal design variables can be determined with both precise and imprecise random variables in engineering applications. The Taguchi QLF is employed to evaluate the robustness of product performance, and the MMR criterion is used to minimize the effect of uncertainty or maximize the robustness.

As a new technology for alternative energy, hydrokinetic turbines have attracted much attention recently. During the development of a new hydrokinetic turbine system (Hu et al. 2012), many challenges exist, including the treatment of imprecise random river flow velocity. The proposed method was then applied to the robust design optimization of the hydrokinetic turbine system. Promising results have been obtained.

The remainder of this article is organized as follows. Section 2 reviews the traditional robust design methodology for random variables with precise probability distributions. Section 3 discusses the types of imprecise random variables. The effects of random variables are then investigated in Section 4. Section 5 presents the robust design method with precise and imprecise random variables. The developed method is then applied to the design optimization of a hydrokinetic turbine in Section 6. Conclusions are drawn in Section 7.

2. Robust design with aleatory uncertainty

In this section, the traditional robust design method is reviewed. The method is based on the Taguchi QLF for problems with precisely known random variables.

2.1. Taguchi quality loss function

The Taguchi QLF is widely used as a robustness metric in robust design optimization (Chen, Wiecek, and Zhang 1999). There are three types of performance variables, which are referred to as quality characteristics (QCs), in a QLF. They are nominal-the-best QCs, smaller-the-better QCs and larger-the-better QCs. The QLF can explicitly represent the effect of a deviation from the target on the quality loss. Minimizing the expected quality loss can bring the mean value of a nominal-the-best QC to its target and reduce the variability of the QC simultaneously.

Define $Z$ as a QC. The QLF of a nominal-the-best QC is given by (Tsui 1992)

$$L(Z, T) = k(Z - T)^2$$

where $k$ is the quality loss coefficient, and $T$ is the target or desired value of $Z$.

The QLFs for a smaller-the-better and a larger-the-better QC are given by

$$L(Z) = kZ^2$$

and

$$L(Z) = \frac{k}{Z^2}$$

respectively.
In this work, the nominal-the-best QC is employed. The expected quality loss is

\[ C(\mu_Z, \sigma_Z) = E[L(Z, T)] = k \left[ \sigma_Z^2 + (\mu_Z - T)^2 \right] \]

and \( \mu_Z \) and \( \sigma_Z \) are the mean and standard deviation of \( Z \), respectively.

### 2.2. Robust design based on the Taguchi quality loss function

Let \( Z \) be given by

\[ Z = f(d, X) \]

where \( d = (d_1, d_2, \ldots, d_n) \) is the vector of design variables, \( f(\cdot) \) is the response function, and \( X = (X_1, X_2, \ldots, X_m) \) is a vector of random variables whose distributions are precisely known.

With the Taguchi QLF, a robust design model for problems with precisely known random variables is given as follows:

\[
\begin{aligned}
& \min_d C(\mu_Z(d, X), \sigma_Z(d, X)) \\
\text{subject to} \\
& g_i(d, \mu_X) \leq 0, \quad i = 1, 2, \ldots, n_i \\
& h_j(d, \mu_X) = 0, \quad j = 1, 2, \ldots, n_e \\
& \Pr\{m_k(d, X) < 0\} \geq P_k, \quad k = 1, 2, \ldots, n_u \\
& d^L \leq d \leq d^U
\end{aligned}
\]

where \( \mu_X \) are the mean values of \( X \); \( g_i(d, \mu_X) \leq 0, i = 1, 2, \ldots, n_i \) are deterministic inequality constraint functions; \( h_j(d, \mu_X) = 0, j = 1, 2, \ldots, n_e \) are deterministic equality constraint functions; \( \Pr\{m_k(d, X) < 0\} \geq P_k, k = 1, 2, \ldots, n_u \) are the probabilistic constraint functions; \( d^L \) and \( d^U \) are lower and upper bounds of \( d \), respectively; and \( \mu_Z(d, X) \) and \( \sigma_Z(d, X) \) are the mean and standard deviation of \( Z \), respectively.

\( \mu_Z(d, X) \) and \( \sigma_Z(d, X) \) can be calculated either by analytical methods or by Monte Carlo simulation (MCS).

With analytical methods,

\[
\begin{aligned}
& \mu_Z(d, X) = \int_{-\infty}^{\infty} f(d, X)p_X(X) dX \\
& \sigma_Z(d, X) = \sqrt{\int_{-\infty}^{\infty} [f(d, X) - \mu_Z(d, X)]^2 p_X(X) dX}
\end{aligned}
\]

where \( p_X(X) \) is the joint probability density function (PDF) of \( X \).

With MCS,

\[
\begin{aligned}
& \mu_Z(d, X) = \frac{1}{N} \sum_{i=1}^{N} f(d, x_i) \\
& \sigma_Z(d, X) \approx \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} [f(d, x_i) - \mu_Z(d, X)]^2}
\end{aligned}
\]

where \( N \) is the number of samples, and \( x_i \) is the \( i \)th sample of \( X \).
MCS generally requires hundreds of calls of function $Z = f(d, X)$. If the function is computationally cheap, MCS can be employed. Otherwise, an analytical method should be used. However, when the dimension of $X$ is large, Equations (9) and (10) for the analytical method may also be computationally expensive because of the high-dimensional integrals. In this case, other approximation methods may be used (Madsen et al. 1986; Du, Sudjianto, and Huang 2005; Huang and Du 2008; Zhang and Du 2010; Banerjee and Smith 2011; Kim et al. 2011; Millwater and Feng 2011).

3. Imprecise random variables

Probability distributions are usually obtained from statistical data. When the data are limited, obtaining a precise distribution is difficult, and the following types of imprecise random variables may be encountered:

- type I: random variables with multiple candidate distributions
- type II: probability distributions whose parameters, such as means and standard deviations, are also uncertain.

In this article, an imprecise random variable is denoted by $\tilde{Y}$.

3.1. Random variables with multiple candidate distributions

When statistical data are not sufficient to produce a precise distribution, several candidate distributions, which can fit the data well, may be obtained. Without any prior knowledge about the distribution of a random variable, a standard procedure of hypothesis testing (Mcgill, Maurer, and Weiser 2006) can be followed by making a hypothesis that the random variable follows distribution A. The hypothesis test can be performed for distributions B, C, D, and so on. If the sample size of the data is not large enough, it is common that several distributions, such as distributions A, B, C and D, can pass the hypothesis test. When this happens, multiple candidate distributions can be used to describe a single random variable, and the distribution is therefore not precisely known. Figure 1 shows the cumulative probability density function (CDF) of an imprecise random variable with three possible distributions.

This situation is commonly encountered in practical applications. For example, the wind speed of a wind turbine at different locations may be described by a Weibull, generalized Rayleigh, lognormal or kappa distribution (He 2009; Morgan et al. 2011). The initial crack size of materials is suggested to be a lognormal or Weibull distribution (Zhang and Mahadevan 2000). Similarly, the equivalent initial flaw size may be modelled as a lognormal, Weibull or three-parameter Weibull distribution (Maymon 2005; Cross, Makeev, and Armanios 2007; Makeev, Nikishkov, and Armanios 2007). When data are too limited, all these empirical models can be used as candidate distributions for the imprecise random variable.

3.2. Random variables with imprecise distribution parameters

For this case, the distribution type of a random variable may be known, but the distribution parameters may not be precise. A distribution parameter such as the mean is typically estimated from available data. The confidence of the estimation largely depends on the sample size: the larger the sample size, the higher the confidence. The estimate may be given within an interval, which is usually called the confidence interval. At a certain confidence level, the width of the confidence interval depends on the sample size. When the sample size is large, the width of the
confidence interval is small, and the midpoint of the interval can then be used as the associated
distribution parameter. If the sample size is too small, however, the interval will be wide. In this
case, the midpoint cannot be simply used, and then the random variable has imprecise distribution
parameters.

Let $\tilde{Y}$ be a random variable with imprecise distribution parameters and $p_{\tilde{Y}} = f_p(y, \tilde{r}_Y)$ be its
probability density function (PDF). Suppose the distribution parameters are

$$\tilde{r}_Y \in [r_Y, \bar{r}_Y]$$  \hspace{1cm} (11)

in which $\tilde{r}_Y$ is the vector of distribution parameters, and $r_Y$ and $\bar{r}_Y$ are the lower and upper bounds,
respectively.

As an example, Figure 2 shows the CDF curves of an imprecise random variable $\tilde{Y}$ with a
known normal distribution, a known standard deviation, but an imprecise mean. Figures 3 and 4
present the CDF curves of a normal random variable with an imprecise standard deviation, and
with both imprecise mean and standard deviation, respectively.

To account for the random variables with imprecise distribution parameters, the p-box method
(Aughenbaugh and Paredis 2006) can be used, where a p-box expresses the CDF of an impre-
cise random variable by interval bounds as indicated in Figures 2–4. However, no matter which
bounds are used, they may always result in regrets. A new method is needed to accommo-
date imprecise random variables for robust design. In the next section, the effects of imprecise
random variables on the robustness of a product are further explained, and a new robust
design method is then developed using the MMR criterion to minimize the regrets due to the
impreciseness.

4. General effects of random variables

Robust design optimization is used to minimize the effects of uncertainties during the design
process. The effects of both aleatory and epistemic uncertainties are explained below.
4.1. Effects of aleatory uncertainty (precise random variables)

Owing to the involvement of aleatory uncertainty, the performance variable $Z$ is also a random variable. To make the performance insensitive to aleatory uncertainty, robust design optimization presented in Section 2 brings the mean of $Z$ to its target and at the same time reduces the standard deviation of $Z$. Figure 5 illustrates the PDF of $Z$ before and after robust design optimization. After the optimization, the distribution of $Z$ is shifted to the target, and the distribution is also shrunk.
4.2. Effects of epistemic uncertainties

When there are also imprecise random variables $\tilde{Y}$, the performance function given in Equation (5) becomes

$$Z = f(d, X, \tilde{Y})$$  \hspace{1cm} (12)

As discussed in Section 3, there may be several models for the distribution of an imprecise random variable. If only one of them is used in the robust design optimization described in Section 2, the obtained solution may not be optimal for other models.
In fact, the performance variable $Z$ is a random variable because it is a function of random variables $X$, and it is also an imprecise random variable because it is also a function of imprecise random variables $\tilde{Y}$. As a result, the mean and standard deviation of $Z$ will not be single-valued quantities. The task of this work is to develop a new robust design optimization approach to deal with the combined aleatory and epistemic uncertainty.

5. Robust design with the mixture of aleatory and epistemic uncertainties

The new robust design optimization method is built upon the Taguchi QLF and the MMR criterion, which will be reviewed first. Then, the new robust design optimization method is discussed.

5.1. The minimax regret criterion

The MMR criterion has been adopted for decision making when the information on input variables is incomplete (Manski 2007; Tetenov 2012). Let $U = [u^{(1)}, u^{(2)}, \ldots, u^{(n)}]$ be a vector of the candidate models for the imprecise random variables $\tilde{U}$, and $s^*_i$ be the optimal strategy for the design obtained from the assumption that only the $i$th model $u^{(i)}$ is correct.

The regret of strategy $s$ with respect to the $i$th model $u^{(i)}$ is given by

$$ R(s, u^{(i)}) = \max_{u} \{ R(s, u^{(j)}) \} $$

where $V(s, u^{(i)})$ is the quality loss of strategy $s$ if the QLF is used. Equation (13) indicates that $R(s, u^{(i)})$ is always positive.

The maximum regret of strategy $s$ is then computed by

$$ R_{\text{max}}(s, U) = \max_{j=1, 2, \ldots, n} \{ R(s, u^{(j)}) \} $$

According to the MMR criterion, the optimal strategy for the vector of imprecise information $U$ is defined as

$$ s^* = \arg \min_d R_{\text{max}}(s, U) $$

The minimax analysis is formulated as an optimization problem as follows:

$$ \text{MMR} = \text{Min} \{ \max_u \{ R(s, u^{(i)}) \} \} $$

subject to

$$ g_i(s) \leq 0, \quad i = 1, 2, \ldots, n_i $$

$$ h_j(s) = 0, \quad j = 1, 2, \ldots, n_e $$

where $s = [s_1, s_2, \ldots, s_n]$ is the space of strategies.

5.2. Robust design based on the minimax regret criterion

As discussed in Section 5.1, the MMR criterion can minimize the regret for decision making under uncertainty. In this section, robust design methods with precise and imprecise random variables are introduced based on the MMR criterion. For different types of imprecise random variables, three models are proposed.
5.2.1. Model 1: Robust design with type I imprecise random variables

Let $\tilde{Y}$ be the vector of imprecise random variables. For type I imprecise random variables, $\tilde{Y}$ is given by

$$\tilde{Y} = [Y^{(1)}, Y^{(2)}, \ldots, Y^{(n)}]$$

(17)

where $Y^{(i)}$ stands for the $i$th hypothetical distributions for $\tilde{Y}$.

With the MMR criterion, the following robust design optimization model is proposed:

$$\min_d R_{\max}(d)$$

subject to

$$R_{\max}(d) = \max\{r(d,X,Y^{(1)}), r(d,X,Y^{(2)}), \ldots, r(d,X,Y^{(n)})\}$$

$$r(d,X,Y^{(i)}) = C(\mu_Z(d,X,Y^{(i)}), \sigma_Z(d,X,Y^{(i)}))$$

$$- C(\mu_Z(d^{(i)}_{opt},X,Y^{(i)}), \sigma_Z(d^{(i)}_{opt},X,Y^{(i)})), \quad i = 1, 2, \ldots n_h$$

$$g_j(d, \mu_X, \mu_{Y^{(i)}}) \leq 0, \quad j = 1, 2, \ldots, n_j$$

$$h_k(d, \mu_X, \mu_{Y^{(i)}}) = 0, \quad k = 1, 2, \ldots, n_e$$

$$\Pr\{m_l(d,X,Y^{(i)}) < 0\} \geq P_l, \quad l = 1, 2, \ldots, n_u$$

$$d^L \leq d \leq d^U$$

(18)

where $n_h$ is the number of hypothetical distributions, $r(d,X,Y^{(i)})$ is the regret of the design obtained from the $i$th hypothetical distribution, $R_{\max}(d)$ is the maximum regret, and $d^{(i)}_{opt}, i = 1, 2, \ldots, n_h$ is the optimal design using the $i$th hypothetical distribution from the following optimization model:

$$\min_d C(\mu_Z(d^{(i)}_{opt},X,Y^{(i)}), \sigma_Z(d^{(i)}_{opt},X,Y^{(i)}))$$

subject to

$$g_j(d^{(i)}_{opt}, \mu_X, \mu_{Y^{(i)}}) \leq 0, \quad j = 1, 2, \ldots, n_j$$

$$h_k(d^{(i)}_{opt}, \mu_X, \mu_{Y^{(i)}}) = 0, \quad k = 1, 2, \ldots, n_e$$

$$\Pr\{m_l(d,X,Y^{(i)}) < 0\} \geq P_l, \quad l = 1, 2, \ldots, n_u$$

$$d^L \leq d^{(i)}_{opt} \leq d^U$$

(19)

$$\mu_Z(d^{(i)}_{opt},X,Y^{(i)}) \text{ and } \sigma_Z(d^{(i)}_{opt},X,Y^{(i)})$$

are solved using Equations (7) and (8) or Equations (9) and (10).

5.2.2. Model 2: Robust design with type II imprecise random variables

Let $\tilde{Y}$ be the vector of imprecise random variables. For type II imprecise random variables, $\tilde{Y}$ is presented by

$$\tilde{Y} = Y(\bar{r}_Y), \bar{r}_Y \in [\underline{r}_Y, \bar{r}_Y]$$

(20)

in which $Y(\bar{r}_Y)$ are the vectors of random variables $\tilde{Y}$ given the imprecise distribution parameters $\bar{r}_Y$. 
Then, the robust design optimization model is given by

\[
\begin{align*}
\min_{\mathbf{d}} & \quad R_{\text{max}}(\mathbf{d}) \\
\text{Subject to} & \quad R_{\text{max}}(\mathbf{d}) = \max_{\mathbf{\bar{r}}} r(\mathbf{d}, \mathbf{\bar{r}}) \\
\text{Subject to} & \quad r(\mathbf{d}, \mathbf{\bar{r}}) = C(\mu_Z(\mathbf{d}, X, Y(\mathbf{\bar{r}})), \sigma_Z(\mathbf{d}, X, Y(\mathbf{\bar{r}}))) \\
& \quad -C(\mu_Z(\mathbf{d}_{\text{opt}}, X, Y(\mathbf{\bar{r}})), \sigma_Z(\mathbf{d}_{\text{opt}}, X, Y(\mathbf{\bar{r}}))) \\
& \quad \mathbf{d}_{\text{opt}} = \min_{\mathbf{d}} C(\mu_Z(\mathbf{d}, X, Y(\mathbf{\bar{r}})), \sigma_Z(\mathbf{d}, X, Y(\mathbf{\bar{r}}))) \\
& \quad \mathbf{r}_Y < \mathbf{\bar{r}} < \mathbf{\bar{r}}_Y \\
& \quad g_j(\mathbf{d}, \mu_X, \mu_Y(\mathbf{\bar{r}})) \leq 0, \quad j = 1, 2, \ldots, n_j \\
& \quad h_k(\mathbf{d}, \mu_X, \mu_Y(\mathbf{\bar{r}})) = 0, \quad k = 1, 2, \ldots, n_e \\
& \quad \Pr\{m_l(\mathbf{d}, X, Y(\mathbf{\bar{r}})) < 0\} \geq P_l, \quad l = 1, 2, \ldots, n_u \\
& \quad \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U
\end{align*}
\]

The robust design optimization model involves a triple-loop procedure.

5.2.3. Model 3: Robust design with both type I and II imprecise random variables

When both type I and II imprecise random variables are involved, \( \mathbf{\bar{Y}} \) is given by

\[ \mathbf{\bar{Y}} = [Y(\mathbf{\bar{r}}), Y^{(1)}, Y^{(2)}, \ldots, Y^{(\nu)}], \mathbf{\bar{r}}_Y \in [\mathbf{r}_Y, \mathbf{\bar{r}}_Y] \]

After combining Models 1 and 2, the new robust design optimization model is developed as follows:

\[
\begin{align*}
\min_{\mathbf{d}} & \quad R_{\text{max}}(\mathbf{d}) \\
\text{Subject to} & \quad R_{\text{max}}(\mathbf{d}) = \max_{\mathbf{\bar{r}}} r(\mathbf{d}, \mathbf{\bar{r}}) \\
\text{Subject to} & \quad r(\mathbf{d}, \mathbf{\bar{r}}) = \max\{r_l(\mathbf{d}, \mathbf{\bar{r}}, Y^{(1)}), r_l(\mathbf{d}, \mathbf{\bar{r}}, Y^{(2)}), \ldots, r_l(\mathbf{d}, \mathbf{\bar{r}}, Y^{(\nu)})\} \\
& \quad r_l(\mathbf{d}, \mathbf{\bar{r}}, Y^{(i)}) = C(\mu_Z(\mathbf{d}, X, Y(\mathbf{\bar{r}}), Y^{(i)}), \sigma_Z(\mathbf{d}, X, Y(\mathbf{\bar{r}}), Y^{(i)})) \\
& \quad -C(\mu_Z(\mathbf{d}_{\text{opt}}, X, Y(\mathbf{\bar{r}}), Y^{(i)}), \sigma_Z(\mathbf{d}_{\text{opt}}, X, Y(\mathbf{\bar{r}}), Y^{(i)})), \forall \ i = 1, 2, \ldots, n_h \\
& \quad \mathbf{d}_{\text{opt}} = \min_{\mathbf{d}} C(\mu_Z(\mathbf{d}, X, Y(\mathbf{\bar{r}}), Y^{(i)}), \sigma_Z(\mathbf{d}, X, Y(\mathbf{\bar{r}}), Y^{(i)})) \\
& \quad \text{subject to} \\
& \quad g_j(\mathbf{d}, \mu_X, \mu_Y(\mathbf{\bar{r}}), \mu_Y^{(i)}) \leq 0, \quad j = 1, 2, \ldots, n_j \\
& \quad h_k(\mathbf{d}, \mu_X, \mu_Y(\mathbf{\bar{r}}), \mu_Y^{(i)}) = 0, \quad k = 1, 2, \ldots, n_e \\
& \quad \Pr\{m_l(\mathbf{d}, X, Y(\mathbf{\bar{r}}), Y^{(i)}) < 0\} \geq P_l, \quad l = 1, 2, \ldots, n_u \\
& \quad \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U
\end{align*}
\]

It is a quadruple-loop problem.
5.3. Numerical procedure

The numerical procedure includes three steps.

Step 1: Perform robust design optimization for every hypothetical distribution and obtain the minimum quality loss.
Step 2: Compute the regrets for design strategies.
Step 3: Identify the maximum regret for design strategies and obtain the optimal design by minimizing the maximum regret.

In the subsequent subsections, the numerical procedures are summarized for Models 1, 2 and 3 in detail.

5.3.1. Numerical procedure of Model 1

The flowchart for the first robust design optimization model is given in Figure 6.

The main steps are summarized below.

Step 1: Perform robust design optimization for \( Z = f(d, X, Y^{(i)}) \) using Equation (19).
After the optimization, the minimum quality loss \( C_{\min} = [C_{\min}(1), C_{\min}(2), \ldots, C_{\min}(n_h)] \) corresponding to each hypothetical distribution is obtained.
Step 2: Use Equation (18) to calculate the regret \( r(d, X, Y^{(i)}) \) for each hypothetical distribution.
Step 3: Identify \( R_{\max}(d) \) and minimize it by changing the design variables.

5.3.2. Numerical procedure of Model 2

Figure 7 shows the flowchart of the numerical procedure.

![Flowchart of the first robust design optimization model.](image)
Figure 7. Flowchart of the second robust design optimization model.

The main steps are described below.

Step 1: Initialize $d$ and $\bar{r}_Y$.

Step 2: Perform the inner loop optimization for the response $Z = f(d, X, Y(\bar{r}_Y))$ under distribution parameters $\bar{r}_Y$.

Step 3: Compute regret $r(d, \bar{r}_Y)$ of design $d$ with distribution parameters $\bar{r}_Y$.

Step 4: Check the convergence. If the regret $r(d, \bar{r}_Y)$ is the maximum, go to next step; otherwise, generate new point for $\bar{r}_Y$ and go to Step 2.

Step 5: Check convergence. If the maximum regret $R_{\text{max}}$ reaches the minimum, stop; otherwise, generate a new point for $d$ and go to Step 1.

5.3.3. Numerical procedure of Model 3

Figure 8 depicts the flowchart of the numerical procedure.

Similar to Model 2, there are five steps as follows:

Step 1: Initialize $d$ and $\bar{r}_Y$.

Step 2: Solve submodels $i, i = 1, 2, \ldots, n_h$, and obtain the regrets $r_{\text{temp}}(d, \bar{r}_Y, Y^{(i)})$ with parameters of $d$ and $\bar{r}_Y$. 
Step 3: Identify the maximal regret $r(d, \tilde{r}_Y)$ from regrets $r_i(d, \tilde{r}_Y, Y^{(i)}), i = 1, 2, \ldots, n$.

Step 4: Check convergence. If the regret $r(d, \tilde{r}_Y)$ reaches the maximum, go to the next step; otherwise, generate a new point for $\tilde{r}_Y$ and go to Step 2.

Step 5: Check convergence. If the maximum regret $R_{\text{max}}$ is obtained, stop; otherwise, generate a new point for $d$ and go to Step 1.

It is more computationally expensive to perform robust design with imprecise random variables than a design with only precise random variables. Since there are multiple loops involved in the proposed models, solving the models efficiently is a challenging task. For practical applications, the solution is to create meta-models to replace the expensive models that are needed to obtain the objective and constraint functions. The drawback of doing so is the sacrifice of accuracy. Hence, there is a trade-off between efficiency and accuracy. For problems with costly simulation models, constructing meta-models takes most of the time for the robust design. Carefully
constructing meta-models is the key to maintaining satisfactory efficiency with acceptable accuracy. Developing more efficient algorithms for robustness assessment will also help improve the efficiency.

6. Robust design optimization of hydrokinetic turbine

In this section, the application of the proposed method in the design of a hydrokinetic turbine system is discussed.

6.1. Problem statement

As one of the most sustainable, clean and carbon-free energy sources, hydropower has drawn the attention of many engineers and researchers (Nitin, Suchi, and Arindam 2011). The most conventional and commonly used way is to construct water dams, which extract energy from running water flow. The construction of water dams, however, has many disadvantages, including the expensive initial construction cost, special requirements for natural sites and hazards brought to the environment. Hydrokinetic turbine systems and tidal turbine systems have many advantages over water dams (Anyi, Kirke, and Ali 2010; Lago, Ponta, and Chen 2010; Hu and Du 2012). Hydrokinetic turbines have the same working principle as that of wind turbines. They are portable; they have a cheap initial construction cost and no special requirements for application sites. Although much progress has been made in the area of the design of hydrokinetic turbines, the commercialization of hydrokinetic turbines is still limited by their power efficiency.

One factor that affects the robustness of the power output is the natural variability of river velocity, which is inherent in the water climate. For instance, a hydrokinetic turbine with a constant rotational speed may have maximum power efficiency under a certain river velocity. But when the river velocity varies to another level, the efficiency may become low. To maximize the power productivity of a hydrokinetic turbine, the probabilistic characteristics of river velocity should be considered during the design process. One approach to this is through robust design optimization.

The uncertainties involved in the design of hydrokinetic turbines can be classified into two groups: the uncertainty of geometric dimensions due to manufacturing imprecision and the uncertainty in the river velocity. The uncertainty in dimensions is commonly modelled by normal distributions. The distribution parameters can be easily obtained from the nominal dimension variables for the mean values and from the tolerances for standard deviations. It is, however, difficult to obtain the distribution of the river velocity. The data available are not sufficient enough to determine a precise distribution. For the reasons mentioned previously, several candidate distributions for the river velocity may be obtained. There are, therefore, both precise and imprecise random variables in the design optimization. The proposed new robust design optimization methods can be applied to the optimization of hydrokinetic turbines.

The blades of a hydrokinetic turbine with a constant chord length are shown in Figure 9. The three-blade system was developed for potential operation in the Missouri River. The diameter of the hydrokinetic turbine rotor is 1 m. The task of the robust design is to optimize the chord length and rotational speed, which in turn maximize the energy productivity of the turbine.

The design variables are, therefore, the rotational speed \( \omega \) and the mean of the chord length \( \mu_c \). The dimension variables \( c \) and \( r_{\text{root}} \), which are the chord length and radius of turbine blades, are precise random variables. The imprecise random variable is the river flow velocity \( v \).
6.2. Robust design optimization model for hydrokinetic turbines

6.2.1. Robust design optimization model

Since the river velocity $v$ is an imprecise random variable with multiple hypothetical distributions, Model 1 is applicable to this design problem. The optimization model is given by

$$
\begin{align*}
\min_{\mathbf{d}=\{\mu_c, \omega\}} & \quad R_{\text{max}}(\mu_c, \omega) \\
\text{subject to} & \quad R_{\text{max}}(\mu_c, \omega) = \max \{ r(\mu_c, \omega, r_{\text{root}}, v^{(1)}), r(\mu_c, \omega, r_{\text{root}}, v^{(2)}), \ldots, r(\mu_c, \omega, r_{\text{root}}, v^{(n_h)}) \} \\
& \quad r(\mu_c, \omega, r_{\text{root}}, v^{(i)}) = C(\mu_{P_0}(c, \omega, r_{\text{root}}, v^{(i)}), \sigma_{P_0}(c, \omega, r_{\text{root}}, v^{(i)})) \\
& \quad \quad \quad - C(\mu_{P_0}(c_{\text{opt}}, \omega_{\text{opt}}, r_{\text{root}}, v^{(i)}), \sigma_{P_0}(c_{\text{opt}}, \omega_{\text{opt}}, r_{\text{root}}, v^{(i)})), \quad i = 1, 2, \ldots n_h \\
& \quad C(\mu_{P_0}, \sigma_{P_0}) = k_p \left[ \sigma_{P_0}^2 + (\mu_{P_0} - T_p)^2 \right] \\
& \quad 0.167 \text{ m} \leq \mu_c \leq 0.25 \text{ m} \\
& \quad 2 \text{ rad/s} \leq \omega \leq 10 \text{ rad/s}
\end{align*}
$$

where $n$ is the number of hypothetical distributions of the river velocity; $r_{\text{root}}$ is the radius of the hydrokinetic turbine blade; $v^{(i)}$ is the $i$th hypothetical distribution of the river velocity; $k_p$ is the QLF coefficient of the power output; $T_p$ is the target of the power output; $\mu_{P_0}(c, \omega, r_{\text{root}}, v^{(i)})$ and $\sigma_{P_0}(c, \omega, r_{\text{root}}, v^{(i)})$ are the mean and standard deviation of the power output, respectively, for the $i$th hypothetical distribution of the river velocity; $c_{\text{opt}}$ and $\omega_{\text{opt}}$ are the optimal chord length and rotational speed, respectively, with respect to the $i$th hypothetical distribution, which can be obtained by solving the following optimization model:

$$
\begin{align*}
\min_{\mathbf{d}=\{c, \omega\}} & \quad k_p \left[ \sigma_{P_0}^2(c, \omega, r_{\text{root}}, v^{(i)}) + (\mu_{P_0}(c, \omega, r_{\text{root}}, v^{(i)}) - T_p)^2 \right]
\end{align*}
$$
The power output is given by
\[
Po = 0.5 \rho v^3 \pi r_{\text{root}}^2 C_p(\lambda, c)
\]  
where \( \rho \) is the river water density (kg/m\(^3\)), \( v \) is the river velocity (m/s), \( C_p(\lambda, c) \) is the power coefficient, and \( \lambda \) is the tip speed ratio (TSR) of the turbine.

The TSR \( \lambda \) is given by
\[
\lambda = \frac{r_{\text{root}} \omega}{v}
\]
Substituting Equation (27) into Equation (26) yields
\[
Po(c, \omega, r_{\text{root}}, v) = 0.5 \rho v^3 \pi r_{\text{root}}^2 C_p(r_{\text{root}} \omega/v, c)
\]
Equations (26)–(28) show that the key to the design is to compute the power coefficient \( C_p(\lambda, c) \).

### 6.2.2. Construction of surrogate model for \( C_p(\lambda, c) \)

The computational fluid dynamics (CFD) simulation is needed to compute the power coefficient. The CFD simulation, however, is very computationally expensive. It cannot be directly applied to the robust design. Based on the CFD simulations at specified points from design of experiment (DOE), surrogate models were constructed using the PCE (Xiu and Karniadakis 2003; Xiu and Shen 2009) method for \( C_p(\lambda, c) \).

#### 6.2.2.1 Computational fluid dynamics simulation

A CFD analysis was performed to study the effect of operating parameters on the performance of the hydrokinetic turbine. The turbine model used for this study was a 1 m radius–constant chord–no twist turbine made from single aerofoil SG6043. ANSYS Fluent (ANSYS 2009) was used on an eight-core, 25 GB RAM computer. A multiple reference frame technique was employed to model the flow over the turbine, wherein the turbine was placed within an inner domain, which rotates inside the outer stationary domain (Figure 10a). The fluid-flow governing equations were solved in the rotating domain for the inner domain and stationary frame for the outer domain. The transformation of the flow variables took place at the interface between the two domains. Figure 10(a) shows various boundary conditions imposed on the CFD model. The velocity inlet boundary condition was specified at the inlet of the outer flow domain while the pressure outlet boundary condition was specified at the exit. The outer cylinder surface was modelled as a symmetry boundary so that there was no flux and flow across the boundary. The turbine was treated as a solid wall, and the common region between the two domains was defined as the interior through which flow transfer takes place.

The flow domain consisted of about 2.8 million tetrahedral/hybrid elements. The mesh for the boundary layer near the turbine blade surface was locally refined to accurately simulate the near-wall boundary layer flow field (Figure 10b).

A second order discretization scheme was used to solve the convective terms in the fluid-flow governing equations. The pressure velocity coupling was solved using the semi-implicit method for pressure linked equation (SIMPLE) algorithm. The pressure staggering options (PRESTO) scheme was adopted owing to its superiority for flows with steep pressure gradient such as the present case (ANSYS). A \( k-\omega \) shear stress transport (SST) turbulence model of Menter (1994) was used to characterize the turbulent flow around the turbine. The convergence criteria for residuals...
of continuity, momentum, turbulent kinetic energy ($k$) and specific dissipation ($\omega$) equations were set to $10^{-4}$ RMS. The fluid domain mesh (Figure 10) was generated using the ANSYS meshing tool, and the grid resolution requirements were well established by keeping $y^+ \sim 120$, so that the wall boundary layer was adequately resolved with good accuracy. The procedure and results of the CFD analysis are discussed in detail in Mukherji et al. (2011).

Table 1 presents various parameters chosen for the CFD analysis while Table 2 summarizes the results of the CFD analysis for the constant chord turbines used for the robust design study.

**Table 1. Parameters for computational fluid dynamics (CFD) analysis.**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrofoil</td>
<td>SG-6043</td>
</tr>
<tr>
<td>Density ($\rho$)</td>
<td>998.2 kg/m$^3$</td>
</tr>
<tr>
<td>Pressure ($p$)</td>
<td>101.3 kPa</td>
</tr>
<tr>
<td>Rotor radius ($R$)</td>
<td>1 m</td>
</tr>
<tr>
<td>Chord length ($c$)</td>
<td>0.167–0.3 m</td>
</tr>
<tr>
<td>Number of blades ($N$)</td>
<td>2–4</td>
</tr>
<tr>
<td>Blade pitch ($\theta_p$)</td>
<td>$10^\circ$</td>
</tr>
<tr>
<td>Rotor speed ($\Omega$)</td>
<td>3–8 rad/s</td>
</tr>
<tr>
<td>Fluid speed ($U_\infty$)</td>
<td>2 m/s</td>
</tr>
<tr>
<td>Turbulence model</td>
<td>$k-\omega$ SST</td>
</tr>
<tr>
<td>Interpolating scheme</td>
<td>2nd order upwind</td>
</tr>
<tr>
<td>Pressure scheme</td>
<td>PRESTO</td>
</tr>
<tr>
<td>Residual error</td>
<td>$1 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

**Table 2. Results of computational fluid dynamics (CFD) simulations.**

<table>
<thead>
<tr>
<th>$c$ (m)</th>
<th>0.167</th>
<th>0.2</th>
<th>0.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSR</td>
<td>$C_p$</td>
<td>$C_p$</td>
<td>$C_p$</td>
</tr>
<tr>
<td>1.5</td>
<td>0.06</td>
<td>0.06</td>
<td>0.08</td>
</tr>
<tr>
<td>2</td>
<td>0.09</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>2.5</td>
<td>0.12</td>
<td>0.12</td>
<td>0.11</td>
</tr>
<tr>
<td>3</td>
<td>0.13</td>
<td>0.16</td>
<td>0.10</td>
</tr>
<tr>
<td>3.5</td>
<td>0.14</td>
<td>0.12</td>
<td>0.08</td>
</tr>
<tr>
<td>4</td>
<td>0.10</td>
<td>0.08</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Note: TSR = tip speed ratio.
Various parameters, such as the TSR and chord length \( (c) \), were tested to quantify their effect on the performance of the hydrokinetic turbine. The CFD simulations were carried out for turbines with chord length varying from 0.167 to 0.3 m and TSR varying from 1.5 to 4.

6.2.2.2 Polynomial chaos expansion method The expansion order for the PCE is three, the expansion interval of the TSR is \([1.5, 4.0]\) and the interval of the chord length is \([0.167, 0.25]\) m. The constructed surrogate model is given by

\[
\hat{C}_p(\xi_1, \xi_2) = \sum_{k=0}^{9} \chi_k \Gamma_k(\xi)
\]

\[
= \chi_0 + \sum_{i=1}^{2} \chi_i L_1(\xi_i) + \chi_3 L_1(\xi_1)L_1(\xi_2) + \sum_{i=1}^{2} \chi_{3+i} L_2(\xi_i)
\]

\[
+ \chi_6 L_2(\xi_1)L_1(\xi_2) + \chi_7 L_1(\xi_1)L_2(\xi_2) + \sum_{i=1}^{2} \chi_{7+i} L_3(\xi_i)
\]

(29)

where

\[
\xi_1 = \frac{2\lambda - L_\lambda - U_\lambda}{U_\lambda - L_\lambda}
\]

(30)

\[
\xi_2 = \frac{2c - L_c - U_c}{U_c - L_c}
\]

(31)

in which \( U_\lambda \) and \( L_\lambda \) are the upper and lower bounds of the TSR expansion interval, respectively; \( U_c \) and \( L_c \) are the upper and lower bounds of the chord length expansion interval, respectively; and \( L_i(\cdot) \) is the \( i \)th order Legendre polynomial basis. For \( i = 1, L_1(x) = x \); for \( i = 2, L_2(x) = 1/2(3x^2 - 1) \); and for \( i = 3, L_3(x) = 1/2(5x^3 - 3x) \). The Legendre polynomial bases were selected to perform the chaos expansion because the design variables can be treated as generalized variables with uniform distributions and they have equal weights over the expansion intervals.

To compute coefficients \( \chi_k, k = 0, 1, 2, \ldots, 9 \), or the PCE, the point collocation method was employed (Wei, Cui, and Chen 2008; Eldred and Burkar 2009; Hu and Youn 2011). Three-dimensional CFD simulations were performed first at the sample points generated from DOE. Based on the results of the CFD simulations, the power coefficients of the turbine blades were obtained. With \( N_p \) CFD simulations, the coefficients \( \chi_k \) were solved with the following equation:

\[
\begin{pmatrix}
\Gamma_0(\xi_1^1) & \Gamma_1(\xi_1^1) & \cdots & \Gamma_9(\xi_1^1) \\
\Gamma_0(\xi_1^2) & \Gamma_1(\xi_1^2) & \cdots & \Gamma_9(\xi_1^2) \\
\vdots & \vdots & \ddots & \vdots \\
\Gamma_0(\xi_1^{N_p}) & \Gamma_1(\xi_1^{N_p}) & \cdots & \Gamma_9(\xi_1^{N_p})
\end{pmatrix}
\begin{pmatrix}
\chi_0^1 \\
\chi_1^1 \\
\vdots \\
\chi_0^{N_p}
\end{pmatrix}
=
\begin{pmatrix}
C_p(\xi_1^1) \\
C_p(\xi_1^2) \\
\vdots \\
C_p(\xi_1^{N_p})
\end{pmatrix}
\]

(32)

where \( \xi_i^k = [\xi_1^k, \xi_2^k], i = 1, \ldots, N_p \), is the \( i \)th group of sample points, and \( C_p(\xi_i^k) \) is the power coefficient with the \( i \)th group of sample points obtained from the CFD simulations. The coefficients are given by

\[
\chi_0 = 0.1048, \quad \chi_1 = 0.0091, \quad \chi_2 = 0, \quad \chi_3 = -0.0171, \quad \chi_4 = -0.0262, \quad \chi_5 = -0.0035, \quad \chi_6 = 0.0059, \quad \chi_7 = 2.659 \times 10^{-4}, \quad \chi_8 = -0.0019, \quad \chi_9 = -0.0090
\]

(33)
After the construction of the surrogate model, the goodness of fit was evaluated by calculating the coefficients of determination. The coefficients of determination $R^2$ are given by

$$R^2 = 1 - \frac{\sum (C_p - \hat{C}_p)^2}{\sum (C_p - \bar{C}_p)^2} \quad (34)$$

where $C_p$ is the data obtained from the CFD simulations, $\bar{C}_p$ is the mean of all data from the CFD simulations, and $\hat{C}_p(\xi)$ is the data computed from the surrogate model. $R^2$ is 0.9370, which indicates that the fitted model is accurate. (A coefficient of determination of 0.9 or higher is considered a good fit.)

Figure 11 plots the scatter diagram of the results obtained from the CFD simulations and the surrogate model. The diagram shows that the points are evenly distributed at the two sides of $C_p = \bar{C}_p$. It indicates that the surrogate model obtained in Equation (29) can accurately describe the relationship between power coefficient, TSR and chord length.

Figure 12 presents the surrogate model of the power coefficient and data points obtained from the CFD simulations. The shaded surface is the surrogate model, while the star points represent the data from the CFD simulations.

All the equations for calculating the power output $P_0(c, \omega, r_{\text{root}}, v)$ are now available. With the surrogate model for the power coefficient $C_p$, it is computationally cheap to compute $P_0(c, \omega, r_{\text{root}}, v)$. The MCS method can therefore be applied to the computation of the mean and standard deviation of $P_0(c, \omega, r_{\text{root}}, v)$, which are needed by Equations (24) and (25). With MCS, the mean and standard deviation are computed by

$$\mu_{P_0}(c, \omega, r_{\text{root}}, v^{(j)}) = \frac{1}{N} \sum_{i=1}^{N} P_0(c_i, \omega, r_{\text{root},i}, v^{(j)}_i) \quad (35)$$

$$\sigma_{P_0}(c, \omega, r_{\text{root}}, v^{(j)}) = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} [P_0(c_i, \omega, r_{\text{root},i}, v^{(j)}_i) - \mu_{P_0}(c, \omega, r_{\text{root}}, v^{(j)})]^2} \quad (36)$$

where $c_i, r_{\text{root},i}$ and $v^{(j)}_i$ are the $i$th sample of $c$, $r_{\text{root}}$ and $v^{(j)}$, respectively.
6.3. Data

Table 3 presents the precise parameters and random variables for the design optimization of the hydrokinetic turbine. Among these parameters, \( r_{\text{root}} \) and \( c \) are truncated at three sigma level owing to the manufacturing tolerance, and the river velocity \( v \) is truncated at 0.8 m/s and 4.5 m/s because of the cut-in and cut-out river velocities.

The river velocity \( v \) can be obtained from the historical river velocity data (as shown in Figure 13) of the Missouri River from the year 1898 to 1989.

<table>
<thead>
<tr>
<th>Variable</th>
<th>( \rho )</th>
<th>( r_{\text{root}} )</th>
<th>( c )</th>
<th>( k_p )</th>
<th>( T_p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Constant</td>
<td>Truncated normal</td>
<td>Truncated normal</td>
<td>Constant</td>
<td>Constant</td>
</tr>
<tr>
<td>Mean</td>
<td>( 1 \times 10^3 \text{ kg/m}^3 )</td>
<td>1 m</td>
<td>( \mu_c )</td>
<td>2</td>
<td>6000</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>N/A</td>
<td>( 1 \times 10^{-2} \text{ m} )</td>
<td>( 1 \times 10^{-3} \text{ m} )</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Figure 12. Surrogate model of the power coefficient and data points from simulations.

Figure 13. Historical river velocity data of the Missouri River at Hermann, Missouri station.
Before fitting distributions for the historical data, the Lilliefors test is performed, which is a special case of the Kolmogorov–Smirnov goodness-of-fit test. The Lilliefors test tests raw data against normal, lognormal, extreme value, Weibull and exponential distributions without specifying the distribution parameters. Under 95% confidence, the normal, lognormal and Weibull distributions were not rejected, and the extreme value distribution was rejected. The normal, lognormal and Weibull distributions were then fitted with the data. Figure 14 shows the velocity data and the three distributions. The associated distribution parameters are given in Table 4.

To verify the three distributions, after their fittings, the Kolmogorov–Smirnov tests were performed using the raw data in Figure 13. All the distribution passed the testing under 95% confidence. This means that the three distributions could serve as candidate distributions for the river velocity.

6.4. **Numerical procedure for the robust design optimization of hydrokinetic turbines**

Figure 15 shows the main steps for the robust design of hydrokinetic turbines.

6.5. **Results and discussion**

Table 5 shows the results from both the robust design optimization with only precise random variables (the traditional method) and the robust design optimization based on the MMR criterion (the new method). In this table, Design $i$, where $i = 1, 2, 3$, means the optimal design considering only the $i$th candidate distribution. The three designs are from the traditional robust design method.

<table>
<thead>
<tr>
<th>Distribution type</th>
<th>Weibull (distribution 1)</th>
<th>Lognormal (distribution 2)</th>
<th>Normal (distribution 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution parameters</td>
<td>1.85 (scale)</td>
<td>1.77 m/s (mean)</td>
<td>1.82 m/s (mean)</td>
</tr>
<tr>
<td></td>
<td>14.17 (shape)</td>
<td>0.13 m/s (Std)</td>
<td>0.14 m/s (Std)</td>
</tr>
</tbody>
</table>

Note: Std = standard deviation.
The ‘optimal design’ in the table stands for the design obtained from the new method with the MMR criterion.

Table 6 gives the means and standard deviations of power outputs of the four designs, and Table 7 provides the quality losses of the four designs.

The two tables clearly show that Design $i$ is the optimal design for the $i$th candidate distribution because of the least quality loss, but is not optimal for the other two candidate distributions. For example, when Design 1 is adopted, for the Weibull distribution, the loss is $4.72 \times 10^7$, which
is the minimum among the four designs. If the distribution is indeed Weibull, Design 1 is then the true optimal design. However, if the true distribution is lognormal, the quality loss will be \(2.93 \times 10^7\), and the minimal loss will be \(2.6685 \times 10^7\) from Design 2; if the true distribution is normal, the quality loss will be \(4.5891 \times 10^7\), and the minimal loss will be \(4.5875 \times 10^7\) from Design 3. Then Design 1 will no longer be optimal for these two situations, for which non-zero regret values will be generated. The regret values for Design 1 are calculated as follows.

For the Weibull distribution,

\[
\text{Quality loss of Design 1} - \text{Minimal quality loss under this distribution} = 4.72 \times 10^7 - 4.72 \times 10^7 = 0
\]

Similarly, for the lognormal distribution,

\[
2.93 \times 10^7 - 2.67 \times 10^7 = 2.59 \times 10^6
\]

For the normal distribution,

\[
4.59 \times 10^7 - 4.58 \times 10^7 = 0.016 \times 10^6
\]

Thus, Design 1 has a maximum regret of \(2.59 \times 10^6\). Similarly, the regret values for Designs 2, 3 and the optimal design can be calculated. All the regret values of the four designs are shown in Table 8 and plotted in Figure 16. In Table 8, the maximum regrets of the four designs are summarized. The maximum regrets for Designs 1, 2, 3 and the optimal design are \(2.5904 \times 10^6\), \(3.1243 \times 10^6\), \(2.2281 \times 10^6\) and \(0.6948 \times 10^6\), respectively.

The results show that the maximum regret of the three designs obtained from the traditional robust design method is about three times that of the optimal design. This indicates that the robust design method based on the MMR criterion has improved the robustness of the design with respect to the uncertainties in the random variables.

<table>
<thead>
<tr>
<th>Hypothetical distributions</th>
<th>Design 1</th>
<th>Design 2</th>
<th>Design 3</th>
<th>Optimal design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull</td>
<td>0</td>
<td>3.1243</td>
<td>0.0156</td>
<td>0.6948</td>
</tr>
<tr>
<td>Lognormal</td>
<td>2.5904</td>
<td>0</td>
<td>2.2281</td>
<td>0.6948</td>
</tr>
<tr>
<td>Normal</td>
<td>0.0155</td>
<td>2.6701</td>
<td>0</td>
<td>0.4995</td>
</tr>
<tr>
<td>Maximum regret</td>
<td>2.5904</td>
<td>3.1243</td>
<td>2.2281</td>
<td>0.6948</td>
</tr>
</tbody>
</table>

Figure 16. Regret values of four different designs under different candidate distributions.
7. Conclusions

In many engineering applications, some random variables are precisely known and others may be imprecisely known. There are uncertainties in both the probabilistic models and probabilistic parameters of random variables owing to incomplete information or limited data for the modelling of random variables. The results of design optimizations will be affected by the model used for the imprecise random variable. To make the design more robust against the uncertainties in random variables, a robust design method based on the MMR criterion is developed. The method is able to minimize the maximum regret of the design with respect to the quality loss. The robust design method with only precise random variables, the Taguchi QLF and the MMR criterion are considered together to obtain an optimal design.

The new method has been applied to the robust design optimization of a hydrokinetic turbine. The results demonstrate that the traditional robust design can introduce large regrets in the quality loss if imprecise random variables exist. The results also show that the new method can reduce the regret significantly.

The application was based on the MCS, which needs to call the objective and constraint functions many times. For applications with expensive functions, solving the robust design optimization will be less efficient. Improving the efficiency will be a direction for future research.

Acknowledgements

The authors gratefully acknowledge the support from the Office of Naval Research through contract ONR N000141010923 (Program Manager Dr. Michele Anderson) and the Intelligent Systems Center at the Missouri University of Science and Technology.

References


