# Advanced Statistical Tools for Modelling of Composition and Processing Parameters for Alloy Development

**Greg Zrazhevsky, Alex Golodnikov, Stan Uryasev, and Alex Zrazhevsky**

**Abstract** The paper presents new statistical approaches for modeling highly variable mechanical properties and screening specimens in development of new materials. Particularly, for steels, Charpy V-Notch (CVN) exhibits substantial scatter which complicates prediction of impact toughness. The paper proposes to use Conditional Value-at-Risk (CVaR) for screening specimens with respect to CVN. Two approaches to estimation of CVaR are discussed. The first approach is based on linear regression coming from the Mixed-Quantile Quadrangle, and the second approach builds CVN distribution with percentile regression, and then directly calculates CVaR. The accuracy of estimated CVaR is assessed with some variant of the coefficient of multiple determination. We estimated discrepancy between estimates derived by two approaches with the Mean Absolute Percentage error. We compared VaR and CVaR risk measures in the screening process. We proposed a modified procedure for ranking specimens, which takes into account the uncertainty in estimates of CVaR.

**Keywords** Steel • Toughness • Statistical Modeling • CVaR • Screening • Samples

G. Zrazhevsky (✉)
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
e-mail: zgrig@univ.kiev.ua

A. Golodnikov
Institute of Cybernetics of NAS of Ukraine, Kyiv, Ukraine

S. Uryasev
University of Florida, Gainesville, FL, USA

A. Zrazhevsky
American Optimal Decisions, Inc., Gainesville, FL, USA



A. Migdalas, A. Karakitsiou (eds.), *Optimization, Control, and Applications in the Information Age*, Springer Proceedings in Mathematics & Statistics 130,
DOI 10.1007/978-3-319-18567-5_21

## 1 Introduction

Development of new steels is an extremely costly and time-consuming process. The process involves two main stages: (1) deciding on chemical composition and processing parameters of the steel, based on previous experiences; (2) for the suggested composition of steel, production and testing of trial commercial specimens (see some description of the test of steel specimen at this link[1]). This testing process generates data on important mechanical characteristics of new steels such as yield tensile strength, elongation, and impact toughness (Charpy V-notch—CVN). The CVN impact test is designed to provide a measure of metal resistance to fast fracture in the presence of a flaw or notch. It has been used extensively in mechanical testing of steel products, in research, and in procurement specifications for over three decades. These mechanical characteristics are the basis for evaluation of obtained experimental specimens and selection of steels with the best properties for further development and more extensive testing. The selection of experimental specimens is not a trivial task. The experimental values of the mechanical characteristics in the selection process are not reliable, since they are random by its nature and may significantly depend on non-controlled conditions of physical experiments. The first stage in the development process can be done with statistical models such as ordinary linear regression model predicting underlying mean values of corresponding mechanical characteristics. While the tensile strength can be predicted with reasonable accuracy [1], the prediction of impact toughness is a much more difficult problem because experimental CVN data exhibit substantial scatter. The Charpy test does not provide a measure of an invariant material property, and CVN values depend on many parameters, including specimen geometry, stress distribution around the notch, and microstructural inhomogeneities around the notch tip. More on the CVN test, including the reasons behind the scatter and statistical aspects of this type of data analysis, can be found in [2–6]. Creating alloys with the best CVN values, therefore, results in multiple specimens for each experimental condition, leading to complex and expensive experimental programs.

To overcome this difficulty of predicting CVN, paper [1] suggested to use quantile regression, a nonparametric generalization of the ordinary least square regression introduced by Koenker and Bassett [7]. This technique predicts any given quantile (or percentile) of the distribution of CVN, rather than a single mean value (as in standard mean square regression). The quantile regression imposes minimal distributional assumptions on the data (the response is not required to be normal; data may be heteroscedastic, that is the variability in the response can vary for different values of the explanatory variables). Quantile regression combines results of measurement of the same dependent variable (CVN) that were collected from different specimens. Paper [1] used the quantile regression for predicting 20 % of CVN in the screening phase to assure that a specimen satisfies the toughness requirement (i.e., model-predicted 20 % value is higher than specified threshold).

[1] http://theconstructor.org/structural-engg/tensile-test-on-mild-steel-specimen/3514/.

In financial risk management, quantile, called the Value-at-Risk (VaR), is used to estimate tails of distributions. However, in recent years, Conditional Value-at-Risk (CVaR) is frequently used instead of VaR. Risk measure VaR provides a lower bound for the right tail of the distribution. Therefore, VaR does not measure the outcomes which of the most concern. In contrast to VaR, risk measure CVaR is an average of values upper VaR when right tail is estimated. More on the VaR, CVaR, and quantile regression can be found in [8–14].

This paper expands the approach described in [1] and considers CVaR instead of VaR for the estimating the right tail of CVN distribution. Simulation results presented in the paper are based on real-life data set described in [1]. This data set includes alloy chemical composition, plate thickness, and processing parameters (treating and ageing temperatures) for 34 specimens. Apart from these parameters, there correspond also to each specimen three values of CVN at $-84\,^{\circ}\mathrm{C}$.

Section 2 outlines the Risk Quadrangle [8] theory. Within this theory quantile (VaR) regression, which was pioneered in statistics by Koenker and Bassett [7], can be estimated using linear regression by minimizing Koenker–Bassett error (see [8]). CVaR is presented as mixed quantile (mixed VaR), and can be estimated using Linear Regression by minimizing Rockafellar error (see [8]). Thus, the first approach to estimation of CVaR is based on a Mixed-Quantile-Based Quadrangle.

Section 3 proposes the second approach to estimation of CVaR, which first generates sample of large size and builds CVN distribution, and then directly calculates CVaR of CVN. Accuracy of generated CVN distribution is assessed by using quantile regression equivalent of the coefficient of multiple determination. This section compares numerical performance of two approaches to CVaR estimations: (1) mixed percentile regression based on Rockafellar error; (2) distribution built with percentile regression.

Section 4 analyzes probabilistic models of CVN for different specimens, and compares three rules of specimens screening: (1) rule suggested in [3], which is based on the average and the 20th VaR of the CVN distribution; (2) rule suggested in [1] which is based only on the 20th VaR of the CVN distribution; (3) rule which is based only on the 20th CVaR of the CVN distribution. This section compares also performance of two approaches to estimation of CVaR in the process of screening specimens with respect to CVN.

Section 5 investigates precision of CVaR estimation based on distribution built with percentile regression. This section proposes a modified procedure for screening specimens, which takes into consideration uncertainty in estimates of CVaR.

## 2 Percentile and Mixed Percentile Regression

The Risk Quadrangle [8] theory defines groups of stochastic functionals called Quadrangles. Every quadrangle contains so-called Risk, Deviation, Error and Regret (negative utility). These elements of quadrangle are linked by so-called Statistics functional.

CVaR and VaR are elements of *Percentile Quadrangle*, in particular, CVaR is Risk and VaR is Statistics in this quadrangle (see [8]). Quadrangle is named after its Statistics, in this case Statistics is percentile (VaR). The Koenker–Bassett error is the Error in Percentile Quadrangle. Therefore, percentile (VaR) can be estimated using Linear Regression by minimizing Koenker–Bassett error (see [8]).

This section considers also *Mixed Percentile Quadrangle*. Mixed Percentile is Statistics and Rockafellar error is Error in *Mixed Percentile Quadrangle*. Mixed percentile (mixed VaR) can be estimated using Linear Regression by minimizing Rockafellar error (see [8]). CVaR for discrete distribution can be presented as Mixed VaR. Therefore, CVaR is Statistics in Mixed Percentile Quadrangle. It is interesting to observe that CVaR is Risk in Percentile Quadrangle and Statistics in Mixed Percentile Quadrangle.

Let us explain described concepts with exact mathematical terms. Let $X$ be a random cost; $X_i(x),\ i = 0,1,\ldots,m$, is a family of random costs depending on a decision vector $x = (x_1,\ldots,x_n)$ belonging to a subset $S$. Measure of risk $\mathbf{R}$ aggregates the overall uncertain cost in $X$ into a single numerical value $\mathbf{R}(X)$. This measure is used to model the statement "$X$ adequately $\leq C$" by the inequality $\mathbf{R}(X) \leq C$.

Consider a family of random costs $X_i(x)$, $i = 0,1,\ldots,m$, depending on a decision vector $x = (x_1,\ldots,x_n)$ belonging to a subset $S$. A potential aim in choosing $x$ from $S$ would be to keep the random variable $X_i(x)$ adequately $\leq c_i$ for $i = 1,\ldots,m$, while achieving the lowest $c_0$ such that $X_0(x)$ is adequately $\leq c_0$. The way "adequately" could have different meaning for different $i$, and the notion of a risk measure addresses this issue. A selection of risk measure $\mathbf{R}_i$ that pins down the intended sense of "adequately" in each case leads to a optimization problem having the form

$$\text{choose } x \in S \text{ to minimize } \mathbf{R}_0(X_0(x)) \text{ subject to } \mathbf{R}_i(X_i(x)) \leq c_i \text{ for } i = 1,\ldots,m.$$

A measure of deviation $\mathbf{D}$ deals with uncertainty in a random variable $X$ quantifying its nonconstancy. Thus $\mathbf{D}(X)$ is a generalization of the standard deviation $\sigma(X)$. Consideration of nonstandard measures of deviation in place of standard deviation is motivated by their ability to capture "heavy tail behavior" in probability distributions.

A measure of regret, $\mathbf{v}$, is introduced to quantify the net displeasure $\mathbf{v}(X)$ perceived in the potential mix of outcomes of random "costs" $X$. Regret comes up in penalty approaches to constraints in stochastic optimization and, in mirror image, corresponds to measure of "utility" $\mathbf{U}$ in a context of gains $Y$ instead of losses $X$ (which is typical in economics: $\mathbf{v}(X) = -\mathbf{U}(-X)$, $\mathbf{U}(Y) = -\mathbf{v}(-Y)$). In applying $U$ to $Y$ the last is considered not as absolute gain but gain relative to some threshold, e.g., $Y = Y_0 - B$ where $Y_0$ is absolute gain and $B$ is a benchmark.

A measure of error, $\varepsilon$, assigns to a random variable $X$ a value $\varepsilon(X)$ that quantifies the nonzeroness in $X$. Classical examples are the norms

$$\|X\|_1 = E\,|X|\,, \quad \|X\|_p = [E\,|X|^p]^{1/p} \text{ for } p \in (1,\infty), \quad \|X\|_\infty = \sup|X|\,.$$

Given an error measure $\varepsilon$ and a random variable $X$, one can look for a constant $C$ nearest to $X$ in the sense of minimizing $\varepsilon(X-C)$. The resulting minimum "$\varepsilon$-distance," denoted by $\boldsymbol{D}(X)$, is a deviation measure (see [8]). The $C$ value in the minimum, denoted by $\mathbf{S}(X)$, can be called the "statistic" associated with $X$ by $\varepsilon$. The case $\varepsilon(X) = \|X\|_2$ produces $\mathbf{S}(X) = EX$ and $\mathbf{D}(X) = \sigma(X)$. The generation of a particular deviation measure $\mathbf{D}$ and statistic $\mathbf{S}$ from an error measure $\varepsilon$ has implications for statistical estimation in the sense of generalized regression.

Regression is a way of approximating a random variable $Y$ by a function $f(X_1,\ldots,X_n)$ of one or more random variables $X_j$. The regression evaluates with error measure $\varepsilon$ how far the random difference $Z_f = Y - f(X_1,\ldots,X_n)$ is from 0. For an error $\varepsilon$ and a collection $\mathbf{C}$ of regression functions $f$, the basic problem of regression for $Y$ with respect to $X_1,\ldots,X_n$ is to

$$\text{minimize } \varepsilon(Z_f) \text{ over } f \in \mathbf{C}, \text{ where } Z_f = Y - f(X_1,\ldots,X_n). \tag{1}$$

To illustrate richness of the quadrangle scheme and the interrelationships between quadrangle objects, consider the *Quantile Quadrangle*, and a *Mixed Quantile Quadrangle* [8].

The *Quantile Quadrangle* combines quantile statistics with concepts from risk. By tying "Conditional Value-at-Risk", on the optimization side, to Quantile Regression as pioneered in statistics by Koenker and Bassett [7], it underscores a relationship that might go unrecognized without the risk quadrangle scheme.

Let us consider the (cumulative) distribution function $F_X(x) = P\{X \le x\}$ of a random variable $X$ and the quantile values associated with it. If, for a probability level $\alpha \in (0,1)$, there is a unique $x$ such that $F_X(x) = \alpha$, then $x$, by definition, is the $\alpha$-quantile $q_\alpha(X)$. In general, however, there are two values to consider as extremes:

$$q_\alpha^+(X) = \inf\{x | F_X(x) > \alpha\}, \qquad q_\alpha^-(X) = \sup\{x | F_X(x) < \alpha\}.$$

It is customary, when these differ, to take the lower value as the $\alpha$-quantile, noting that, because $F_X$ is right-continuous, this is the lowest $x$ such that $F_X(x) = \alpha$. Consider the entire interval between the two competing values as the quantile,

$$q_\alpha(X) = \left[q_\alpha^-(X), q_\alpha^+(X)\right].$$

In finance, the Value-at-Risk term is used for quantile, and upper VaR $\mathrm{VaR}_\alpha^+(X) = q_\alpha^+(X)$ along with a lower VaR $\mathrm{VaR}_\alpha^-(X) = q_\alpha^-(X)$, and the VaR interval $\mathrm{VaR}_\alpha(X) = \left[\mathrm{VaR}_\alpha^-(X), \mathrm{VaR}_\alpha^+(X)\right]$ is identical to the quantile interval $q_\alpha(X)$.

Besides VaR, the example coming under consideration involves the CVaR of $X$ at level $\alpha \in (0,1)$, defined by

$$\mathrm{CVaR}_\alpha(X) = \text{ expectation of } X \text{ in its } \alpha\text{-tail},$$

which is also expressed by

$$\mathrm{CVaR}_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_\tau(X)\,d\tau. \tag{2}$$

Conditional Value-at-Risk $\mathrm{CVaR}_\alpha(X)$ is also called in [8] by superquantile $\overline{q}_\alpha(X)$.

Let $X_+ = \max\{0, X\}$, $X_- = \max\{0, -X\}$, $X = X_+ - X_-$.

A *Quantile Quadrangle* has the following elements:

- statistic $\mathbf{S}(X) = \mathrm{VaR}_\alpha(X) = q_\alpha(X)$ = quantile;
- risk $\mathbf{R}(X) = \mathrm{CVaR}_\alpha(X) = \overline{q}_\alpha(X)$= superquantile;
- deviation $\mathbf{D}(X) = \mathrm{CVaR}_\alpha(X - EX) = \overline{q}_\alpha(X - EX)$ = superquantile-deviation;
- regret $\mathbf{v}(X) = \dfrac{1}{1-\alpha} EX_+$ = average absolute loss, scaled;
- error $\varepsilon(X) = E\left[\dfrac{\alpha}{1-\alpha} X_+ + X_-\right]$ = normalized Koenker–Basset error.

The original Koenker–Bassett Error expression differs from the normalized Koenker-Basset error in *Quantile Quadrangle* by a positive factor. In order to build regression function $f(X_1, \ldots, X_n)$ which approximates percentile of random variable $Y$ one should solve the optimization problem (1) with normalized Koenker–Basset error.

Consider the case when random variable $Y$ is approximated by the linear function of a vector of $K+1$ explanatory variables $\mathbf{x}' = [1, x_1, \ldots, x_K]$, $Y = \mathbf{x}'\beta + \delta$, where $\beta' = [\beta_0, \beta_1, \ldots, \beta_K]$. The $\delta$ is zero-mean random term that accounts for the surplus variability or scatter in $Y$ that cannot be explained by explanatory variables $x_1, \ldots, x_K$. The cumulative effects of unmeasured and/or unforeseen variables are usually lumped into the stochastic $\delta$ term.

Let $Y$ denote the logarithm of CVN, $\ln(\mathrm{CVN})$, and $Y_1, \ldots, Y_n$ are observations of the random variable $Y$ at points $\mathbf{x}_i' = \left[1, x_1^i, \ldots, x_K^i\right]$, $i = 1, \ldots, n$. Then estimates of coefficients of the $\alpha$-th quantile regression function can be found by minimizing the normalized Koenker–Basset error

$$\frac{1}{n}\left\{\sum_{i:Y_i \geq \mathbf{x}_i'\beta_\alpha} \frac{\alpha}{(1-\alpha)} \left|Y_i - \mathbf{x}_i'\beta_\alpha\right| + \sum_{i:Y_i \leq \mathbf{x}_i'\beta_\alpha} \left|Y_i - \mathbf{x}_i'\beta_\alpha\right|\right\}. \tag{3}$$

A *Mixed-Quantile Quadrangle* has the following elements for confidence levels $\alpha_k \in (0,1)$ and weights $\lambda_k > 0$, $\sum_{k=1}^{r} \lambda_k = 1$:

- statistic $\mathbf{S}(X) = \sum_{k=1}^{r} \lambda_k q_{\alpha_k}(X) = \sum_{k=1}^{r} \lambda_k \mathrm{VaR}_{\alpha_k}(X)$ = mixed quantile;
- risk $\mathbf{R}(X) = \sum_{k=1}^{r} \lambda_k \overline{q}_{\alpha_k}(X) = \sum_{k=1}^{r} \lambda_k \mathrm{CVaR}_{\alpha_k}(X)$ = mixed superquantile;
- deviation $\mathbf{D}(X) = \sum_{k=1}^{r} \lambda_k \overline{q}_{\alpha_k}(X - EX) = \sum_{k=1}^{r} \lambda_k \mathrm{CVaR}_{\alpha_k}(X - EX)$ = corresponding mixture of superquantile deviations;

- regret $\mathbf{v}(X) = \min\limits_{B_1,\ldots,B_r} \left\{ \sum\limits_{k=1}^{r} \lambda_k \mathbf{v}_{\alpha_k}(X - B_k) \,\middle|\, \sum\limits_{k=1}^{r} \lambda_k B_k = 0 \right\}$ = derived balance of the regrets $\mathbf{v}_{\alpha_k}(X) = \dfrac{1}{1-\alpha_k} EX_+$;
- error $\varepsilon(X) = \min\limits_{B_1,\ldots,B_r} \left\{ \sum\limits_{k=1}^{r} \lambda_k \varepsilon_{\alpha_k}(X - B_k) \,\middle|\, \sum\limits_{k=1}^{r} \lambda_k B_k = 0 \right\}$ = Rockafellar error function,

where $\varepsilon_{\alpha_k}(X) = E\left[\dfrac{\alpha_k}{1-\alpha_k} X_+ + X_-\right]$ = normalized Koenker–Basset error with $\alpha_k$.

Relationship between $\mathrm{CVaR}_\alpha(Y)$ and *Mixed-Quantile Quadrangle* is established by the formula (2). Classical numerical integration uses a finite subdivision of the interval $[\alpha, 1]$ and replaces the integrand in (2) by a nearby step function or piecewise linear function based on the quantiles marking that subdivision. It is easy to see that the value of the integral for that approximated integrand is actually a mixed quantile expression. Thus for confidence levels $\alpha_r \in (\alpha, 1)$, $r = 1, \ldots, R$, and weights $\lambda_r > 0$, $\sum\limits_{r=1}^{R} \lambda_r = 1$, $\mathrm{CVaR}_\alpha(Y)$ in (2) can be approximated by the mixed quantile

$$\mathrm{CVaR}_\alpha(Y) \approx \sum_{r=1}^{R} \lambda_r \mathrm{VaR}_{\alpha_r}(Y). \tag{4}$$

CVaR regression function is a generalization of the mixed quantile (4) to the case when $Y$ is a linear function of a vector of $K+1$ explanatory variables $\mathbf{x}' = [1, x_1, \ldots, x_K]$ plus random error, $Y = \mathbf{x}'\beta + \varepsilon$, where $\beta' = [\beta_0, \beta_1, \ldots, \beta_K]$.

It is estimated by minimizing Rockafellar error function with

$$\varepsilon_{\alpha_k}(X) = E\left[\frac{\alpha_k}{1-\alpha_k} X_+ + X_-\right] = \\ = \frac{1}{n} \left\{ \sum_{i:Y_i \geq \mathbf{x}_i'\beta_\alpha} \frac{\alpha_k}{(1-\alpha_k)} \left| Y_i - \mathbf{x}_i'\beta_\alpha \right| + \sum_{i:Y_i \leq \mathbf{x}_i'\beta_\alpha} \left| Y_i - \mathbf{x}_i'\beta_\alpha \right| \right\}.$$

An important issue in regression with many independent variables is how to choose a subset of variables so that a large portion of variability of $Y$ is explained by these few $x$ variables. In the ordinary least squares model, one can measure the proportion of variability through a quantity known as $R^2$; e.g., $R^2 = 90\,\%$ means that only $10\,\%$ of the variation in $Y$ cannot be accounted for by the $x$ variables. In the quantile regression, criteria $R^2$ is not applicable. To assess the goodness of fit, Koenker and Machado [11] introduced a quantile regression equivalent of the coefficient of multiple determination, $R^1(\alpha)$, which measures the proportion of the variability in the response that is accounted for by the fitted $\alpha$-th quantile surface.

In order to exclude variables that contribute little to the explanation of $Y$, we applied the stepwise variable selection method. As in its standard least squares regression counterpart, this procedure adds and removes explanatory variables alternately from the model using significance tests, eventually converging to a final subset of variables.

## 3 Building CVN Distributions and Estimation of CVaR for Specimens

In Sect. 2 we considered methods for estimating VaR and CVaR using the *Quantile Quadrangle*, and a *Mixed Quantile Quadrangle*. This section suggests an alternative approach to estimation of VaR and CVaR for CVN distributions. The idea of the approach is to build CVN distribution for each specimen. Then with these distributions we estimate VaR and CVaR.

With quantile regression we can estimate any quantile of the CVN distribution. Moreover, quantile regression can estimate the whole distribution by estimating all quantiles (or at least estimating quantiles with some fine grid). The idea here is to use information over a large number of quantiles to reduce the approximation error and enhance the accuracy of the estimated whole distribution for each specimen. By definition, quantile is the inverse function of the corresponding cumulative probability distribution function.

We constructed distribution of CVN for each specimen using the following procedure, repeated 10,000 times:

1. Draw a random value from the Uniform $(0, 1)$ distribution and treat it as a probability level $\alpha$;
2. Build quantile regression model for this probability level;
3. For each specimen calculate quantile by substituting specimen-specific composition and processing parameters into quantile regression model with parameter $\alpha$ and treat it as a random realization of specimen-specific CVN random value.

Thus, for each specimen we generated specimen-specific large sample of CVN values, which is used for building empirical specimen-specific CVN distribution.

This procedure together with quantile regression techniques is the tool, which transforms available information about all produced and tested trial commercial specimens (chemical composition, processing parameters, and CVN) in the distribution of CVN for each specimen. The more values of probability levels $\alpha$ are involved in this procedure, the more accurate is the transformation. Thus our current-state of knowledge about interrelation between chemical composition, processing parameters, and CVN values is completely presented in form of the CVN distribution. In this sense for large sample size (10,000 or more) we can consider such distributions as a good probabilistic models of CVN for corresponding specimen.

**Table 1** Accuracy of different regions of the generated CVN distribution

| Range of $\alpha$ | Range of coefficients $R^1(\alpha)$ |
|---|---|
| 0.050–0.250 | 0.4219–0.4424 |
| 0.251–0.350 | 0.3488–0.4400 |
| 0.351–0.533 | 0.4400–0.4871 |
| 0.534–0.665 | 0.3400–0.3796 |
| 0.666–0.750 | 0.3015–0.3399 |
| 0.751–0.832 | 0.2500–0.3015 |
| 0.833–0.928 | 0.2001–0.2500 |
| 0.929–0.980 | 0.1954–0.3765 |

Accuracy of generated CVN distribution may be assessed by using quantile regression equivalent of the coefficient of multiple determination, $R^1(\alpha)$, which measures the proportion of the "variability" in the response that is accounted for by the fitted $\alpha$-th quantile surface.

Table 1 shows that for the wide range of probability levels ($0.05 \leq \alpha \leq 0.832$) the quantile regression functions capture more than 25 % of the variability in the response that is accounted for by the fitted $\alpha$-th quantile surface. The most accurate portion of the generated CVN distribution is in the range of probability levels $0.351 \leq \alpha \leq 0.533$. Quantile regression functions corresponding to this range of $\alpha$ account for 44–48.71 % of the response variability.

For constructed CVN distributions we can determine for each specimen the following characteristics: average, $\alpha$-th quantile, and $\alpha$-th CVaR. We used these characteristics in the process of screening specimens with respect to CVN.

Let us compare numerical performance of two approaches to CVaR estimations: (1) mixed percentile regression based on Rockafellar error; (2) distribution built with percentile regression. We transformed specimen-specific CVN distributions into corresponding distributions of ln(CVN) and calculated $\alpha$-th CVaR for each distribution. Let $J$ denote the total number of specimens, and let $\mathrm{CVaR}_\alpha(\ln(\mathrm{CVN}_j))$ be $\alpha$-th CVaR of ln(CVN) for $j$-th specimen, $j = 1,\ldots,J$. Suppose that $\mathrm{CVaR}_\alpha(\ln(\mathrm{CVN})) = \mathbf{x}'\beta$ is CVaR regression function found by minimizing Rockafellar error function, $\mathbf{x}' = [1, x_1, \ldots, x_K]$ are explanatory variables, $\beta' = [\beta_0, \beta_1, \ldots, \beta_K]$ are coefficients of CVaR regression function. Here $\mathbf{x}'_j = \left[1, x_1^j, \ldots, x_K^j\right]$ are values of explanatory variables for $j$-th specimen, $j = 1,\ldots,J$. Then $\mathbf{x}'_j\beta$ is estimate of $\alpha$-th CVaR of ln(CVN) derived from CVaR regression.

Discrepancy between estimates based on Rockafellar error and on distribution built with percentile regression can be measured by Mean Absolute Error (MAE),

$$\mathrm{MAE} = \frac{1}{J}\sum_{j=1}^{J}\left|\mathrm{CVaR}_\alpha(\ln(\mathrm{CVN}_j)) - \mathbf{x}'_j\beta\right|, \tag{5}$$

**Table 2** Average discrepancy between estimates based on Rockafellar error and on distribution built with percentile regression for equal weights $\lambda_1, \ldots, \lambda_R$

| Number of VaRs in Mixed Quantile, R | MAE | MAPE (%) |
|---|---|---|
| 5 | 0.1519 | 4.04 |
| 10 | 0.1643 | 4.31 |
| 15 | 0.1631 | 4.27 |
| 20 | 0.167 | 4.36 |
| 25 | 0.1680 | 4.38 |
| 30 | 0.1754 | 4.55 |

or Mean Absolute Percentage Error (MAPE),

$$\text{MAPE} = \frac{1}{J}\sum_{j=1}^{J}\left|\frac{\text{CVaR}_\alpha(\ln(\text{CVN}_j)) - \mathbf{x}_j'\beta}{\text{CVaR}_\alpha(\ln(\text{CVN}_j))}\right|. \tag{6}$$

Using original data with three observed values of CVN for each specimen, we built 20 %-CVaR regression functions for different values of VaRs in the mixed quantile (4), with equal distance between adjacent values of confidence levels $\alpha_r \in (0.2, 1)$, $r = 1, \ldots, R$, and with equal values of weights $\lambda_1, \ldots, \lambda_R$. Average discrepancies between these estimates are presented in Table 2.

Analyzing Table 2 we draw the following conclusions:

1. Two approaches to CVaR estimations (based on Rockafellar error and on distribution built with percentile regression) provide on average similar estimates of $\text{CVaR}_{0.2}(\ln(\text{CVN}))$ for all specimens. The average discrepancy between these estimates does not exceed 4.55 % in the conducted numerical experiments.
2. This result only slightly depends on the number of VaRs in the mixed quantile. The average discrepancy between these estimates increases from 4.04 to 4.55 % with increase of number of VaRs in mixed quantile from 5 to 30.

We investigated also the case when for a given set of confidence levels $\alpha_r \in (0.2, 1)$, $r = 1, \ldots, R$, corresponding weights $\lambda_1, \ldots, \lambda_R$ are chosen to get the best approximation of $\text{CVaR}_\alpha(\ln(\text{CVN}_j))$ by the sum $\sum_{r=1}^{R} \lambda_r \text{VaR}_{\alpha_r}(\ln(\text{CVN}_j))$ in (4). For this purpose we minimized MAE of such approximation for all specimens

$$\min_{\lambda_1,\ldots,\lambda_R} \frac{1}{J}\sum_{j=1}^{J}\left|\text{CVaR}_\alpha(\ln(\text{CVN}_j)) - \sum_{r=1}^{R}\lambda_r \text{VaR}_{\alpha_r}(\ln(\text{CVN}_j))\right| \tag{7}$$

subject to

$$\sum_{r=1}^{R}\lambda_r = 1, \tag{8}$$

$$\lambda_r \geq 0, \quad r = 1, \ldots, R. \tag{9}$$

**Table 3** Average discrepancy between estimates based on Rockafellar error and on distribution built with percentile regression with weights $\lambda_r^*, r = 1, \ldots, R$

| Number of VaRs in Mixed Quantile, R | MAE | MAPE (%) |
|---|---|---|
| 5 | 0.1422 | 3.74 |
| 10 | 0.1906 | 4.93 |
| 15 | 0.1763 | 4.54 |
| 20 | 0.1802 | 4.63 |
| 25 | 0.1798 | 4.63 |
| 30 | 0.1816 | 4.67 |

We solved optimization problem (7)–(9) for different number of VaRs in mixed quantile, $R = 5, 10, 15, 20, 25, 30$. For all values of $r$ optimal solutions, $\lambda_r^*$, $r = 1, \ldots, R$, are not equal. For example, for $R = 5$, $\lambda_1^* = 0.390$, $\lambda_2^* = 0$, $\lambda_3^* = 0.198$, $\lambda_4^* = 0$, $\lambda_5^* = 0.412$.

Average discrepancies between these estimates for optimal values $\lambda_r^*$, $r = 1, \ldots, R$, are presented in the Table 3.

Comparing Tables 2 and 3 we conclude that goodness of fit of CVaR regression in cases when all weights are equal and when optimal values $\lambda_r^*$, $r = 1, \ldots, R$ are used is approximately the same.

## 4 Comparison of VaR, and CVaR Risk Measures in Screening Process

As described in [3], the alloy development process includes three rounds of laboratory melting, processing and testing of candidate steels. Round 1 involves screening of several exploratory compositions. Round 2 consists of additional laboratory testing of new heats for the best compositions from Round 1. The most promising steel from Round 2 (the prime candidate steel) is the subject of more extensive testing in Round 3, in which several heats are tested. The ranking and selection of best alloys are based on combinations of yield strength and CVN at $-84\,^{\circ}$C.

In [1] we have shown that the tensile yield strength can be accurately predicted with ordinary linear regression. At the same time, we encountered difficulty in attempting to build a linear regression model for CVN at $-84\,^{\circ}$C. Whereas each specimen is characterized by a single value of yield strength, three values are observed for CVN. This setting and the fact that there is a substantial variability in these values complicates the process of screening specimens with respect to CVN.

As specified in [3], an acceptable specimen must satisfy the following toughness requirement: the average of three values of CVN must be greater than specified threshold $c$, with no single value below $c$ by more than some specified value $d$.

It is clear that this ranking criterion is not perfect, since it is based on a very small number of random observations (only three values of CVN). In fact, screening

with this rule is random and unstable. To overcome this shortcoming we (see [1]) screened specimens with respect to CVN with quantile (percentile) regression. Quantile regression involves CVN values from testing of several specimens (not only one specimen). Therefore, ranking criterion based on quantile regression is more stable.

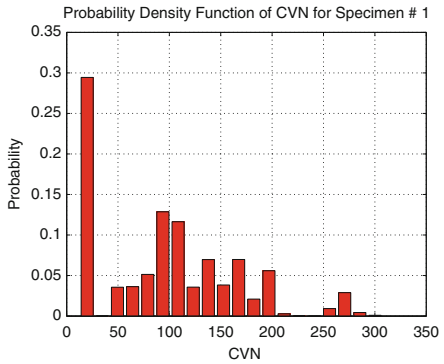
In order to determine which quantile corresponds to the smallest of the three values of CVN, we successively drew three random values from a standard normal distribution, each time selecting only the smallest value. This was repeated 10,000 times, producing an average value of 0.19. The interpretation in [1] was that the smallest of three such CVN values approximately corresponds to the 20th percentile of the CVN distribution. The ranking criterion in [1] was based solely on the 20th percentile of the CVN distribution.

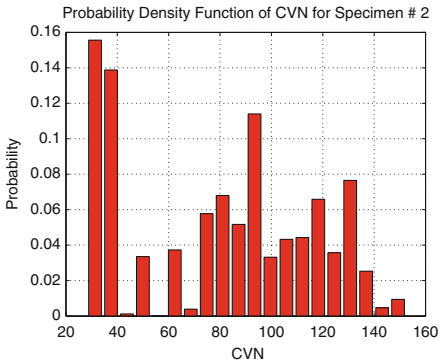
Nevertheless, quantile regression does not evaluate the tail of CVN distribution. In contrast to quantile, the CVaR is an average of outcomes exceeding this quantile. Conceptually, CVaR is preferable to quantile for screening specimens with respect to CVN.

We analyzed probabilistic models of CVN for different specimens using the following rules of specimens screening:

1. Rule suggested in [3] which is based on the average and the 20th VaR of the CVN distribution;
2. Rule suggested in [1] which is based on the 20th VaR of the CVN distribution;
3. Rule which is based on the 20th CVaR of the CVN distribution.

Figure 1 shows that specimen 2 has larger value of 20 % percentile (20 %–VaR) of the CVN distribution than the specimen 1, but smaller values of average and 20 % CVaR. Therefore, the screening rule 1 could not select the better specimen from these two specimens; rule 2 should classify the specimen 2 as better one than

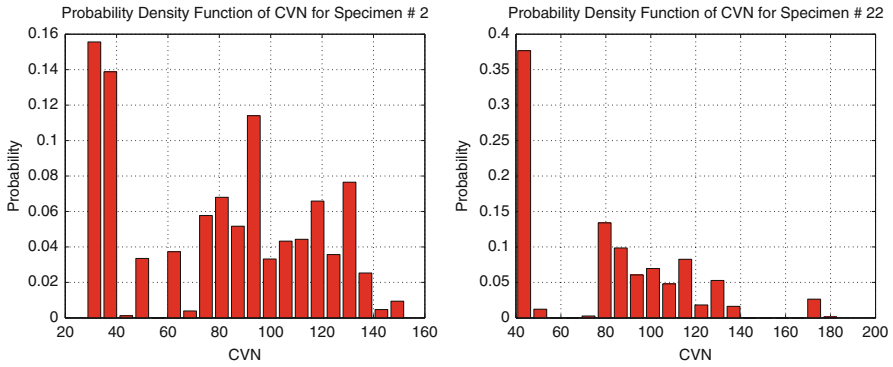



| specimen # | 20%-VaR | 20%-CVaR | Mean |
|---|---|---|---|
| 1 | 19.21 | 117.55 | 97.1 |

| specimen # | 20%-VaR | 20%-CVaR | Mean |
|---|---|---|---|
| 2 | 35.37 | 91.75 | 79.63 |

**Fig. 1** Probabilistic models for specimen 1 and 2, and their characteristics




| specimen # | 20%-VaR | 20%-CVaR | Mean | specimen # | 20%-VaR | 20%-CVaR | Mean |
|---|---|---|---|---|---|---|---|
| 2 | 35.37 | 91.75 | 79.63 | 22 | 45.00 | 89.42 | 80.13 |

**Fig. 2** Probabilistic models for specimen 2, and 22, and their characteristics

specimen 1; rule 3 should classify the specimen 1 as better one than specimen 2. Since CVaR is preferable to quantile for screening specimens with respect to CVN, the correct classification provides only rule 3.

Figure 2 shows that the specimen 22 has larger value of 20 % percentile (20 %–VaR) and the mean of the CVN distribution than the specimen 2, but smaller value of 20 % CVaR. Therefore, the screening rules 1 and 2 should classify the specimen 22 as a better one than specimen 2, but rule 3 should classify the specimen 2 as a better one than specimen 22. Since CVaR is preferable to quantile for screening specimens with respect to CVN, the correct classification provides only rule 3.

These examples demonstrate that results of screening with rules 1, 2, and 3 contradict each other. Rule 3 which is based on the CVaR of the CVN distribution is more appropriate and it is recommended for screening process. We also emphasize that in regression and in probabilistic modeling, all available information should be used to increase accuracy of screening process. For instance, screening of specimens in Round 1 can use CVN distributions derived from data generated during this Round. According to observations from Round 1, these CVN distributions are considered as "correct". Round 2 generates additional information, which should be combined with information from Round 1 and utilized for derivation of new CVN distributions. These new CVN distributions are used for screening of specimens in Round 2. Finally, screening of specimens in Round 3 should use CVN distributions derived from data generated during three rounds.

We compared performance of the two approaches to estimation of CVaR in the process of screening with respect to CVN by using 33 pairs of specimens.

The first approach, which is based on a *Mixed Quantile Quadrangle*, utilizes original experimental data for direct estimating of 20 %-CVaR for $\ln(\text{CVN})$. Therefore, in screening two specimens with numbers $j_1$, and $j_2$ the first approach should compare two values $\text{CVaR}_{0.2}(\ln(\text{CVN}_{j_1}))$, and $\text{CVaR}_{0.2}(\ln(\text{CVN}_{j_2}))$, $j_1$, $j_2 = 1,\ldots,J$.

The second approach estimates 20 %-CVaR in two steps. At the first step, the original experimental data are used for building distributions of CVN for each specimen. At the second step we directly calculated 20 %-CVaR for each specimen using these CVN distributions. In screening, two specimens with numbers $j_1$, and $j_2$ the second approach should compare two values $\mathrm{CVaR}_{0.2}(\mathrm{CVN}_{j_1})$, and $\mathrm{CVaR}_{0.2}(\mathrm{CVN}_{j_2})$, $j_1, j_2 = 1, \ldots, J$.

Since logarithm is an increasing function, the inequalities $\mathrm{CVaR}_{0.2}(\mathrm{CVN}_{j_1}) > \mathrm{CVaR}_{0.2}(\mathrm{CVN}_{j_2})$, and $\mathrm{CVaR}_{0.2}(\ln(\mathrm{CVN}_{j_1})) > \mathrm{CVaR}_{0.2}(\ln(\mathrm{CVN}_{j_2}))$, are equivalent. Both approaches provide the same result only in screening of 20 pairs. In the 13 pairs results of screening were different. Despite the fact that the average discrepancy between estimates of $\mathrm{CVaR}_{0.2}(\ln(\mathrm{CVN}))$ obtained by using these two approaches do not exceed 4.55 % (see results in the Table 2), discrepancy in results of screening exceeds 39 %. Since the second approach uses 10,000 CVN values for each specimen, while the first approach uses only 3 values, we consider the second approach is more accurate.

## 5 Precision of Estimates of CVaR Derived From Small Samples

Section 3 describes procedure for building a CVN distribution for each specimen and determining characteristics of these distributions. The procedure is based on generating large samples of CVN values for each specimen by using quantile regressions. If size of the sample is large, it can be considered as a population corresponding to a specimen, and characteristics of CVN distribution such as average, 20 %-quantile, and 20 %-CVaR can be considered as "true" characteristics corresponding to the specimen. Therefore, these characteristics can be used in the screening process. We considered that samples containing 10,000 values can be classified as large, although for high precision, sample size 100,000 or 1,000,000 may be needed. For such large sample sizes, procedure of building CVN distributions is time consuming.

However, usually, samples of small and moderate size are used. In this case estimates of average, VaR, and CVaR derived from small samples (with size 100, 200, 500) may have large uncertainty, which should be taken into consideration in the screening process. This uncertainty can be quantitatively characterized by 90 %-th or 95 %-th confidence intervals. The "confidence interval probability" is the probability that the calculated confidence interval encompasses the true value of the population parameter (average, VaR, or CVaR). The size of the confidence interval, and the confidence levels provide information on the accuracy of the estimation.

In order to build confidence interval, we randomly sample from a population of size $N$ to produce $k$ new samples of fixed size $n$. Each sample consists of random realizations of random variables $\xi_1, \ldots, \xi_n$. Functions of these random variables, average, 20 % VaR, and 20 % CVaR, called statistics, are also random.

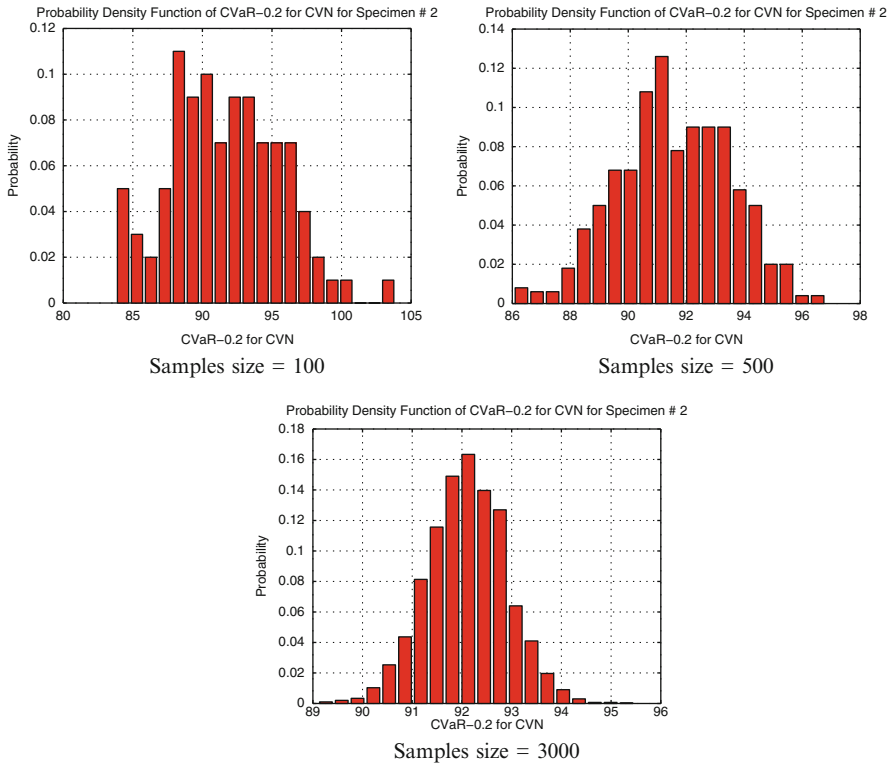



**Fig. 3** Empirical sampling distributions of 20 % CVaR for CVN for Specimen 2 (number of generated samples = 10,000, sample sizes =100; 500; 3,000)

From each generated sample we calculated random realizations of these statistics. With a large number of new samples, $k = 10{,}000$, we generated empirical sampling distributions separately for average, 20 % VaR, and 20 % CVaR. For each of these empirical sampling distributions we determined average, 5 %- and 95 %-quantiles, which specify point estimate as well as lower and upper confidence limits for corresponding statistics.

In our calculations we used the following sampling parameters:

1. population size $N = 10{,}000$;
2. number of generated small samples $k = 10{,}000$;
3. sizes of small samples $n = 100, 200, 300, 500, 1{,}000, 3{,}000$;

Figures 3 and 4 show empirical sampling distributions of the population parameters (20 % CVaR, and 20 % VaR for CVN) in case when random samples of fixed sizes ($n = 100$, 500, and 3,000) were drawn from the population. Figure 3 demonstrates that sample 20 % CVaR distribution gradually converges to the normal distribution when sample size increases from 100 to 3,000. But this is not the case for 20 % VaR distributions (see Fig. 4).

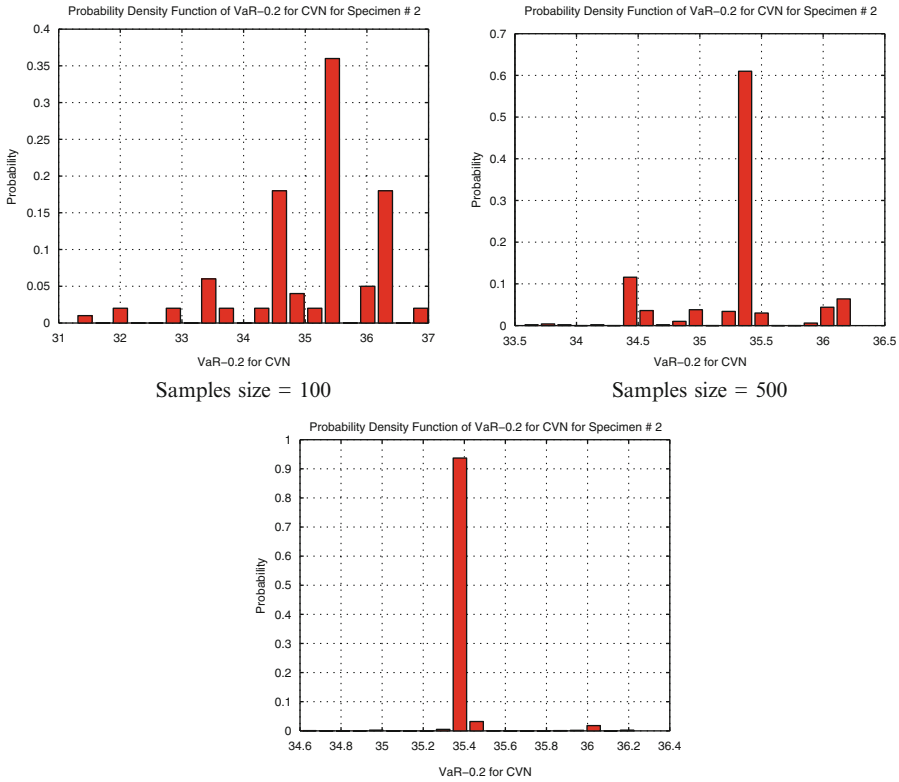



**Fig. 4** Empirical sampling distributions of 20 % VaR for CVN for Specimen 2 (number of generated samples = 10,000, sample sizes =100; 500; 3,000)

Empirical sampling distributions of the population parameters were used for determining their point estimates, 90 % confidence intervals, and the accuracy of estimation calculated as the ratio of length of the confidence interval to the corresponding point estimate. The point estimate for a population parameter was determined as the mean of its empirical sampling distribution. Estimates of 20 % VaR for CVN distribution for Specimen 2 obtained for samples of size 100, 200, 300, 500, 1,000, and 3,000 are in Table 4.

The Table 4 shows the rapid growth of accuracy of the 20 % VaR for CVN (Specimen 2) estimation while a sample size increases from 100 to 3,000. The best accuracy, 0.20 %, is achieved for sample size = 3,000, it is significantly better than for sample size = 1,000.

Similar results of estimation for 20 % CVaR for CVN (Specimen 2) are presented in Table 5.

**Table 4** Estimates of 20 % VaR for CVN (Specimen 2)

| Sample size | Point estimate | 90 % Confidence interval | | Accuracy of estimation (%) |
|---|---|---|---|---|
| | | 5 %-Quantiles | 95 %-Quantiles | |
| 100 | 35.30 | 33.43 | 36.65 | 9.11 |
| 200 | 35.28 | 33.81 | 36.23 | 6.86 |
| 300 | 35.31 | 34.44 | 36.23 | 5.07 |
| 500 | 35.34 | 34.44 | 36.23 | 5.06 |
| 1,000 | 35.36 | 34.61 | 36.05 | 4.10 |
| 3,000 | 35.39 | 35.37 | 35.44 | 0.20 |

**Table 5** Estimates of 20 % CVaR for CVN (Specimen 2)

| Sample size | Point estimate | 90 % Confidence interval | | Accuracy of estimation (%) |
|---|---|---|---|---|
| | | 5 %-Quantiles | 95 %-Quantiles | |
| 100 | 91.82 | 84.51 | 99.02 | 15.80 |
| 200 | 91.79 | 86.63 | 96.96 | 11.25 |
| 300 | 91.74 | 87.51 | 95.95 | 9.20 |
| 500 | 91.77 | 88.57 | 94.99 | 7.00 |
| 1,000 | 91.75 | 89.43 | 94.06 | 5.05 |
| 3,000 | 92.09 | 90.76 | 93.40 | 2.86 |

The Table 5 shows the moderate growth of accuracy of the 20 % CVaR for CVN (Specimen 2) estimation while a sample size increases from 100 to 3,000. However, the comparison of Tables 4 and 5 shows that estimates of 20 % VaR are more accurate than estimates of 20 % CVaR for all sample sizes. For instance, for samples size = 3,000 accuracy of estimation of 20 % CVaR is 10 times worse than accuracy of estimation of 20 % VaR. Therefore, in the screening procedure based on estimation of CVaR derived from samples of moderate sizes we have to use confidence intervals.

All values in the 90 % confidence interval for CVaR are plausible values for the CVaR with probability 0.9, whereas values outside the interval are rejected as plausible values for CVaR. Therefore, in screening process based on estimates of CVaR, which were derived from samples of moderate sizes, we should compare confidence intervals of two specimens instead of their point estimates as described in Sect. 4. If these confidence intervals do not overlap, these specimens are necessarily significantly different. In this case, the specimen, corresponding to the confidence interval with lower values, should be disqualified. However, if these specimens have overlapping confidence intervals, the screening procedure cannot select the best specimen. In this case, we should increase sample size and repeat the procedure of estimation of confidence interval.

For instance, consider confidence intervals for 20 %-CVaR, corresponding to specimen 2 and specimen 10, shown in Fig. 5.

Figure 5 shows that, if sample size equals 100, then these specimens have overlapping confidence intervals. In this case, the screening procedure cannot select

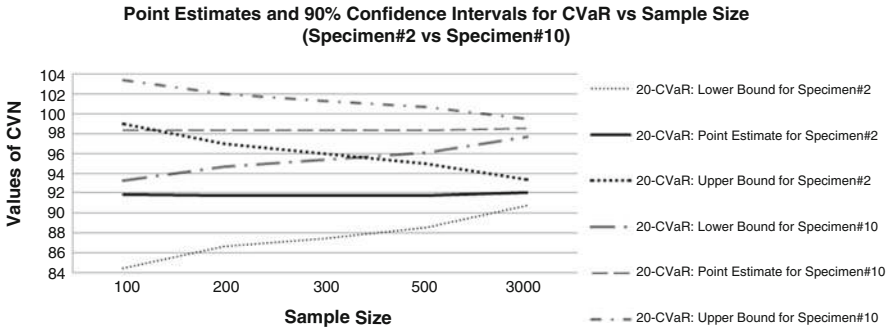



**Fig. 5** Confidence intervals for 20 %-CVaR, corresponding to specimens 2 and 10, which were derived from 10,000 samples of fixed size 100, 200, 300, 500, and 1,000

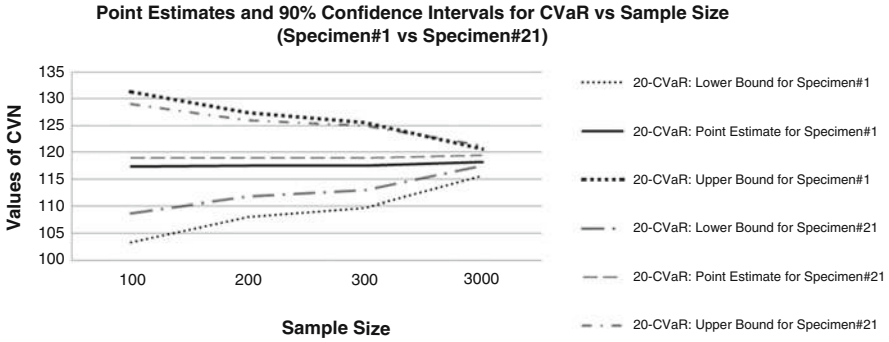



**Fig. 6** Confidence intervals for 20 %-CVaR, corresponding to specimens 1 and 21, which were derived from 10,000 samples of fixed size 100, 200, 300, and 3,000

the best specimen. Then, we produced 10,000 samples of fixed size 200 and derived new confidence intervals for these specimens. Figure 5 shows that for this sample size, specimens 2 and 10 have also overlapping confidence intervals. Successively building confidence intervals for samples of fixed size 300 and 500, we found that in the latter case these confidence intervals do not overlap. Therefore, for sample size 500 screening rule should classify the specimen 10 as a better one than specimen 2.

However, not always increasing of sample size results in separation of confidence intervals for two specimens.

Figure 6 shows that specimens 1 and 21 have overlapping confidence intervals for sample sizes 100, 200, 300, and 3,000. In this case, the screening procedure cannot select the best specimen. Therefore, both these specimens should be classified as promising steel samples.

This section demonstrated that the approach to CVaR estimation based on distribution built with percentile regression allows calculating confidence interval in addition to point estimate for CVaR. Utilization of confidence intervals in the screening procedure enables to reduce screening errors related with uncertainty in

point estimates of CVaR. For this purpose, we should compare confidence intervals of two specimens instead of their point estimates, as described in Sect. 4. If these confidence intervals do not overlap, these specimens are significantly different. In this case, the specimen, corresponding to the confidence interval with lower values, should be disqualified. However, if these specimens have overlapping confidence intervals, the screening procedure cannot select the best specimen. In this case, we should increase the sample size and repeat the procedure of estimation of confidence interval.

## 6 Summary and Conclusions

We presented new statistical approaches for modeling mechanical properties and screening specimens in development of new materials with highly variable properties. Paper [1] suggested to use quantile regression for screening specimens. However, quantile regression does not take into account tails of distributions. In contrast to quantile, CVaR is an average of observations exceeding quantile. Therefore, CVaR which takes into account tails may be preferable to quantile for screening specimens.

We investigated two approaches for CVaR estimation. The first approach, based on a Mixed-Quantile Quadrangle, uses experimental data for direct estimating 20 %-CVaR. In particular, we considered a Mixed-Quantile regression for $\ln(CVN)$ with the Rockafellar error function. The second approach estimates 20 %-CVaR in two steps. The first step uses quantile regression for generating a large number of samples of CVN values and building distributions of CVN for each specimen. The second step calculates 20 %-CVaR for each specimen from these distributions.

Accuracy of generated CVN distribution was evaluated with the coefficient of multiple determination, $R^1(\alpha)$, which measures the proportion of the 'variability' in the response that is accounted for the fitted $\alpha$-th quantile surface. We found that for a wide range of probability levels ($0.05 \leq \alpha \leq 0.832$) the corresponding quantile regression functions capture more than 25 % of the variability in the response (that is accounted for by the fitted $\alpha$-th quantile surface). The most accurate portion of the generated CVN distribution corresponds to the range of probability levels $0.351 \leq \alpha \leq 0.533$. The quantile regression functions corresponding to this range of $\alpha$ account for 44–48.71 % of the variability in the response.

Discrepancy between estimates of CVaR based on Rockafellar error and on distribution from the percentile regression was measured by the MAE. We found that two approaches to CVaR estimation provide similar estimates of $\mathrm{CVaR}_{0.2}(\ln(\mathrm{CVN}))$ for all specimens. The numerical experiments demonstrated that the average discrepancy between these estimates does not exceed 4.55 %.

We compared performance of these two approaches for CVaR estimation for screening specimens with respect to CVN. The dataset included 33 pairs of specimens. Both approaches resulted in the same ranking for 20 pairs. However,

results of screening were different for 13 pairs. Although the average discrepancy between estimates of $\mathrm{CVaR}_{0.2}(\ln(\mathrm{CVN}))$ with these two approaches does not exceed 4.55 % (see results in the Table 2), discrepancy in screening exceeds 39 %. Since in the second approach we considered 10,000 sampled CVN values for each specimen, while in the first approach we used only 3 values, the second approach is more accurate.

We also analyzed probabilistic models of CVN for different specimens using the following rules of specimens screening:

1. Rule suggested in [3], based on the average and the 20th VaR of the CVN distribution;
2. Rule suggested in [1], based only on the 20th VaR of the CVN distribution;
3. Rule based only on the 20th CVaR of the CVN distribution.

Results of screening by using Rules 1, 2, and 3 sometimes contradicted each other. We think that the Rule 3 which is based on the CVaR of the CVN distribution is the more appropriate, and it should be used for screening process.

CVaR estimation based on distribution (built with percentile regression) allows for calculating confidence interval, in addition to point estimate of CVaR. The confidence intervals in the screening procedure reduce screening errors coming from the uncertainty in estimates of CVaR. We compared confidence intervals of two specimens, instead of their point estimates as described in Sect. 4. If confidence intervals do not overlap, the specimens are significantly different. However, for specimens with overlapping confidence intervals the screening procedure cannot select the best specimen. In this case, we can increase the sample size and repeat the procedure of estimation of confidence interval.

Approaches proposed in this paper could identify promising compositions of materials and processing parameters for developing better steels. This may reduce cost of experimental programs, shifting resources to cost-effective computer modeling techniques.

## References

1. Golodnikov, A., Macheret, Y., Trindade, A., Uryasev, S., Zrazhevsky, G.: Statistical Modeling of Composition and Processing Parameters for Alloy Development. Model. Simul. Mater. Sci. Eng. **13**, 633–644 (2005)
2. McClintock, F.A., Argon A.S.: Mechanical Behavior of Materials. Addison-Wesley, Reading, MA (1966)
3. Goldren, A.P., Cox, T.B.: AMAX Report, CPR-2, AMAX Materials Research Center, Ann Arbor, MI (1986)
4. Corwin, W.R., Houghland, A.M.: Effect of Specimen Size and Material Condition on the Charpy Impact Properties of 9Cr-1Mo-V-Nb Steel. In: Corwin, W.R., Lucas, G.E. (eds.) The Use of Small-Scale Specimens for Testing Irradiated Material. ASTM STP 888, Philadelphia, PA, pp. 325–338 (1986)

5. Lucon, E., et al.: Characterizing Material Properties by the Use of Full-size and Sub-size Charpy Tests. In: Siewert, T.A., Manahan, M.P. (Sr. eds.) Pendulum Impact Testing: A Century of Progress. ASTM STP 1380, pp. 146–163. American Society for Testing and Materials, West Conshohocken, PA (1999)
6. Todinov, M.T.: Uncertainty and Risk Associated with the Charpy Impact Energy of Multi-run Welds. Nucl. Eng. Des. **231**, 27–38 (2004)
7. Koenker, R., Bassett, G.: Regression Quantiles. Econometrica **46**, 33–50 (1978)
8. Rockafellar, R.T., Uryasev, S.: The Fundamental Risk Quadrangle in Risk Management, Optimization, and Statistical Estimation. Surv. Oper. Res. Manag. Sci. **18**, 33–53 (2013)
9. Rockafellar, R.T., Uryasev, S.: Optimization of Conditional Value-at-risk. J. Risk **2**, 21–42 (2000)
10. Rockafellar, R.T., Uryasev, S.: Conditional Value-at-risk for General Loss Distributions. J. Bank. Financ. **26**, 1443–1471 (2002)
11. Koenker, R., Machado, J.A.F.: Goodness of Fit and Related Inference Processes for Quantile Regression. J. Am. Stat. Assoc. **94**, 1296–1310 (1999)
12. Taylor, J.W., Bunn, D.W.: Combining Forecast Quantiles Using Quantile Regression: Investigating the Derived Weights, Estimator Bias and Imposing Constraints. J. Appl. Stat. **25**, 193–206 (1998)
13. Bassett, G., Koenker, R.: An Empirical Quantile Function for Linear Models with iid Errors. J. Am. Stat. Assoc. **77**, 407–415 (1982)
14. Koenker, R., d'Orey, V.: Computing Regression Quantiles. Appl. Stat. **43**, 410–414 (1994)