Journal of Loss Prevention in the Process Industries 23 (2010) 300–307

Contents lists available at ScienceDirect

Journal of Loss Prevention in the Process Industries

journal homepage: www.elsevier.com/locate/jlp

An uncertainty and sensitivity analysis of dynamic operational risk assessment model: A case study

Xiaole Yang, M. Sam Mannan*

Mary Kay O'Connor Process Safety Center, Artie McFerrin Department of Chemical Engineering, Texas A&M University System, College Station, TX 77843-3122, USA

ARTICLE INFO

Article history:
Received 25 August 2009
Received in revised form
9 November 2009
Accepted 9 November 2009

Keywords:
Uncertainty characterization
Sensitivity analysis
Dynamic operational risk assessment
Goodness-of-fit

ABSTRACT

In Dynamic Operational Risk Assessment (DORA) models, component repair time is an important parameter to characterize component state and the subsequent system-state trajectory. Specific distributions are fit to the industrial component repair time to be used as the input of Monte Carlo simulation of system-state trajectory. The objective of this study is to propose and apply statistical techniques to characterize the uncertainty and sensitivity on the distribution model selection and the associated parameters determination, in order to study how the DORA output that is the probability of operation out-of-control, can be apportioned by the distribution model selection. In this study, eight distribution fittings for each component are performed. Chi-square test, Kolmogorov–Smirnov test, and Anderson-Darling test are proposed to measure the goodness-of-fit to rank the distribution models for characterizing the component repair time distribution. Sensitivity analysis results show that the selection of distribution model among exponential distribution, gamma distribution, lognormal distribution and Weibull distribution to fit the industrial data has no significant impact on DORA results in the case study.

1. Introduction

For Dynamic Operational Risk Assessment (DORA) models (Fig. 1), component state sojourn time distributions are the inputs for system-state trajectory simulation. Among the three component states (Fig. 2), the third state (*abnormal detected and under repair*) sojourn time distribution is obtained by fitting distribution to industry component repair time data. In the case that the collected data is sufficient enough, distribution fitting is statistically satisfied with accepted uncertainty. However, it is not always possible to find enough data. In a highly reliable system, a single failure may occur at a frequency in order of 10^{-6} or $10^{-7}$and repair happens at a corresponding low frequency so that the repair data is usually not enough for a good distribution fitting. Given the limited repair time data points, the major concern on uncertainty includes: 1) what distribution type should be selected; 2) whether the distribution type is a sensitive factor for DORA results.

Therefore the problem first becomes to select a distribution model and the associated distribution parameters for characterizing component state sojourn time that have the best representation among a class of distributions. There are several techniques to examine how well a sample of data agrees with a given distribution as its population. In those goodness-of-fit techniques, hypothesis test is based on measuring the discrepancy or consistency of the sample data to the hypothesized distribution. Chi-square test is used to measure how well the fit matches the data if the data are represented by discrete points with Gaussian uncertainties (Bock & Krischer, 1998). However, the value of the chi-square test statistic depends on how the data is binned. Another disadvantage of the chi-square is that it requires an adequate sample size for the approximations to be valid. Pearson's Chi-square test is distinguished from the case with Gaussian errors, and is applied if the data are represented by integer numbers of events in discrete bins, following Poisson statistics rule (Cowan, 1998). Kolmogorov–Smirnov (K–S) test is a goodness-of-fit measurement technique for one-dimensional data samples. It is used to test whether the data sample comes from a population with a specific distribution (Chakravarti, Laha, & Roy, 1967). Anderson-Darling test (Stephens, 1974) is a modification of K–S test and gives more weight to the tails than does the K–S test. There are several others, such as the Shapiro–Wilk test (Shapiro & Wilk, 1965) and the probability plot (Chambers, Cleveland, Kleiner, & Tukey, 1983) for goodness-of-fit measurement.

When a failure occurs to a component, the component must be repaired and it is then unavailable for processing during a certain amount of time called the repair time (Dallery & Gershwin, 1992). In

* Corresponding author. Tel.: +1 979 862 3985.
E-mail address: mannan@tamu.edu (M. Sam Mannan).

0950-4230/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jlp.2009.11.001

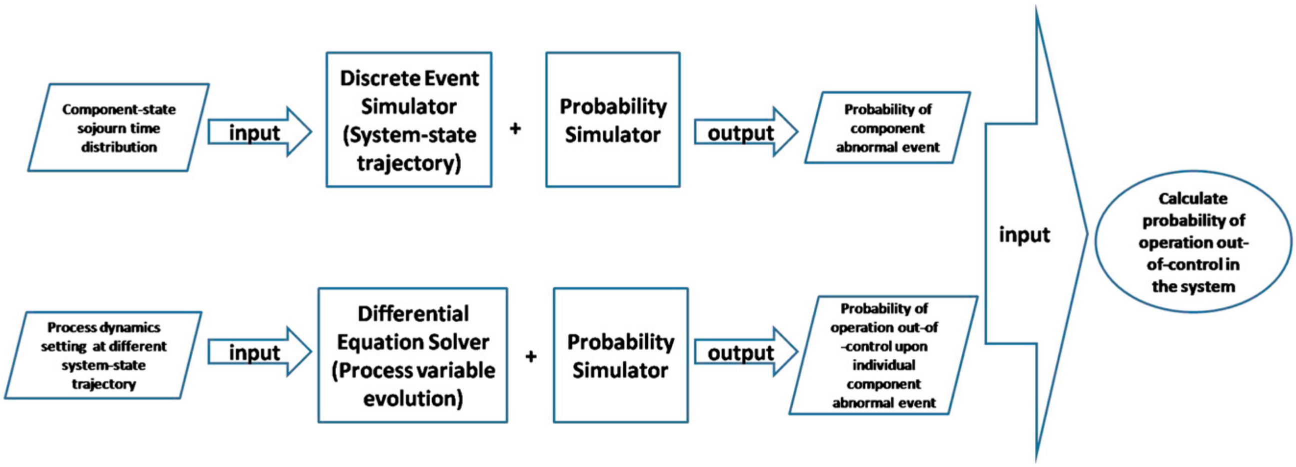

Fig. 1. Scheme of DORA model.

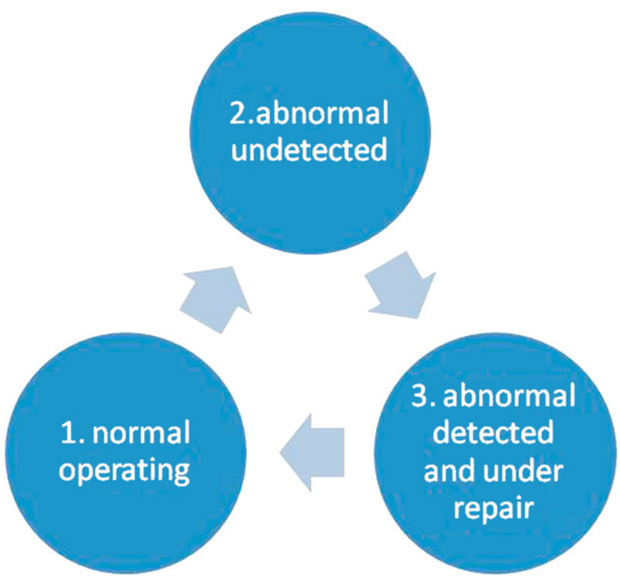

Fig. 2. Component state flow diagram.

reliability engineering, random variables from exponential distribution, gamma distribution, lognormal distribution or Weibull distribution are usually assumed to characterize the time-to-repair distribution in most of the models. By selecting candidates from those distribution families, epistemic uncertainty is reduced using engineer expert judgment. The uncertainty is further reduced by selecting the distribution model according to the rank of goodness-of-fit. In Section 2 of this paper, eight distributions are fitted into the component repair time data. In Section 3 the goodness-of-fit is measured using Chi-square test, K–S test and A-D test. A sensitivity analysis is performed in Section 4 to study the impact of distribution type selection to the dynamic operational risk assessment results. A summary of this research study is presented in Section 5.

2. Distribution fittings for DORA input

In a DORA study on the level control in an oil/gas separator, three components are in the study scope, pump, control valve (CV) and level transmitter (LT). Each of the components can be specified at any time by defining its performance behavior state as time being: *normal operating*, *abnormal undetected*, and *abnormal detected and under repair* (Fig. 2). A random variable is assigned to represent the sojourn time a component spends in any of the states. X is a random variable characterizing the sojourn time of component in State 1. Therefore the exponential distribution density function of the first state sojourn time is given by:

$$f(x;\lambda) = \lambda e^{-\lambda x} \quad (1)$$

If the component inspection interval is T, and Y is random variable to characterize the sojourn time of component State 2, the cumulative distribution function of State 2 sojourn time is given by:

$$P(Y<y) = \frac{e^{-\lambda(T-y)} - e^{-\lambda T}}{1-e^{-\lambda T}} \quad (2)$$

For the third component state, we obtain prior knowledge about the sojourn time by collecting component repair time data. Four widely used distribution types in reliability engineering modeling, exponential, gamma, lognormal, and weibull are used to fit the collected repair time data. Each type has two distributions with different number of parameters. They are exponential with single parameter (exponential), exponential with two parameters

Table 1
Distribution parameters of the eight distributions fitted to the pump, CV and LT repair time data.

Distribution	Pump – Parameters	CV – Parameters	LT – Parameters
Exponential	$\lambda = 0.11744$	$\lambda = 0.16901$	$\lambda = 0.32258$
Exponential (2P)	$\lambda = 0.13307$ $\gamma = 1.0$	$\lambda = 0.29268$ $\gamma = 2.5$	$\lambda = 0.90909$ $\gamma = 2.0$
Gamma	$\alpha = 0.23273$ $\beta = 36.586$	$\alpha = 0.98773$ $\beta = 5.9901$	$\alpha = 7.3923$ $\beta = 0.41935$
Gamma (3P)	$\alpha = 0.33912$ $\beta = 24.11$ $\gamma = 1.0$	$\alpha = 0.51275$ $\beta = 6.7362$ $\gamma = 2.5$	$\alpha = 0.46567$ $\beta = 1.754$ $\gamma = 2.0$
Lognormal	$\sigma = 0.99585$ $\mu = 1.4947$	$\sigma = 0.64201$ $\mu = 1.5107$	$\sigma = 0.30246$ $\mu = 1.0832$
Lognormal (3P)	$\sigma = 1.3804$ $\mu = 1.0937$ $\gamma = 0.82505$	$\sigma = 6.7709$ $\mu = -0.46155$ $\gamma = 2.5$	$\sigma = 0.78934$ $\mu = 0.04695$ $\gamma = 1.7088$
Weibull	$\alpha = 1.1698$ $\beta = 7.0744$	$\alpha = 3.7703$ $\beta = 3.8824$	$\alpha = 4.2115$ $\beta = 2.8792$
Weibull (3P)	$\alpha = 0.54255$ $\beta = 5.7422$ $\gamma = 1.0$	$\alpha = 0.61366$ $\beta = 2.5638$ $\gamma = 2.5$	$\alpha = 0.88087$ $\beta = 1.1366$ $\gamma = 2.0$