ELSEVIER

Reliability Engineering and System Safety 73 (2001) 167–176

RELIABILITY ENGINEERING & SYSTEM SAFETY

www.elsevier.com/locate/ress

CDF sensitivity analysis technique for ranking influential parameters in the performance assessment of the proposed high-level waste repository at Yucca Mountain, Nevada, USA

Sitakanta Mohanty[a,*], Y-T. (Justin) Wu[b,1]

[a]Center for Nuclear Waste Regulatory Analyses, Southwest Research Institute, 6220 Culebra Road, San Antonio, TX 78238, USA
[b]Structural Integrity and Reliability Section, Southwest Research Institute, 6220 Culebra Road, San Antonio, TX 78238, USA

Received 25 July 2000; accepted 14 May 2001

Abstract

A cumulative distribution function (CDF)-based method has been used to perform sensitivity analysis on a computer model that conducts total system performance assessment of the proposed high-level nuclear waste repository at Yucca Mountain, and to identify the most influential input parameters affecting the output of the model. The performance assessment computer model referred to as the TPA code, was recently developed by the US nuclear regulatory commission (NRC) and the center for nuclear waste regulatory analyses (CNWRA), to evaluate the performance assessments conducted by the US department of energy (DOE) in support of their license application. The model uses a probabilistic framework implemented through Monte Carlo or Latin hypercube sampling (LHS) to permit the propagation of uncertainties associated with model parameters, conceptual models, and future system states. The problem involves more than 246 uncertain parameters (also referred to as random variables) of which the ones that have significant influence on the response or the uncertainty of the response must be identified and ranked. The CDF-based approach identifies and ranks important parameters based on the sensitivity of the response CDF to the input parameter distributions. Based on a reliability sensitivity concept [AIAA Journal 32 (1994) 1717], the response CDF is defined as the integral of the joint probability-density-function of the input parameters, with a domain of integration that is defined by a subset of the samples. The sensitivity analysis does not require explicit knowledge of any specific relationship between the response and the input parameters, and the sensitivity is dependent upon the magnitude of the response. The method allows for calculating sensitivity over a wide range of the response and is not limited to the mean value.

Keywords: Sensitivity analysis; Risk assessment; Uncertainty analysis; Waste management; Influential parameters

1. Introduction

The US nuclear regulatory commission (NRC) and the center for nuclear waste regulatory analyses (CNWRA) recently developed a general computer model [2,3] for conducting total-system performance assessments of the proposed high-level waste repository at Yucca Mountain (YM) in Nevada, USA. This computer model is a tool for evaluating the performance assessments conducted by the US department of energy (DOE) in support of their license application. The relationship between the output and the input parameters in such a model is highly nonlinear and involves strong interactions among the input parameters. Since it is often difficult to describe precisely the governing physical processes, coupling among these processes, parameters defining the physical system, and the evolution of the repository system over the long time period of interest (TPI), significant uncertainties exist in the computer model. The performance assessment model is then designed in such a way that these uncertainties can be analyzed by using alternative conceptual models and assigning uncertainties to model parameters. The performance assessment model also has a large number of input parameters that are described by ranges of variance and probability distribution functions representing uncertainties and variability. Sensitivity analysis of the performance assessment model is conducted to determine the uncertainty in the output due to uncertainties in the model (not considered in the paper) and input parameters and to determine the most influential input parameters that control the behavior of the output. Knowledge of the most influential input parameters is

* Corresponding author. Tel.: +1-210-522-5160; fax: +1-210-522-5155.
E-mail address: smohanty@swri.org (S. Mohanty).
[1] Current address: Applied Research Associates, 811 Forest Road, Raleigh, NC 27609, USA.

0951-8320/01/$ - see front matter
PII: S0951-8320(01)00054-0

important because it can be used for (i) providing an insight on where more efforts should be devoted to reduce the uncertainties of the output to significantly improve the understanding of the system, (ii) comparing results from calculations conducted by different research groups involved in solving the same problem, and (iii) aiding in design improvements by interactively improving the design criteria or requirements to reduce vulnerability to a particular design parameter.

Methods that are available to identify influential parameters such as linear and step-wise regression [4], non-parametric approaches [5,6], one-at-a-time approaches [7,8], response surface analysis, and Fourier amplitude sensitivity test (FAST) have their advantages and disadvantages. For example, while the linear regression-based sensitivity analysis methods provide ranking of the parameters regardless of the values of the response (an advantage), these methods are most suitable when the response is approximately a linear function of the input parameters (a disadvantage). Similarly, while a one-at-a-time approach has an unambiguous way to relate output parameter sensitivity to an input parameter, the method may not be able to study the interaction effects well. The limitation of other methods are documented in Mohanty et al. [9] and Lu and Mohanty [10].

This paper presents an alternative approach to identify and rank influential parameters based on the sensitivity of the cumulative distribution function (CDF) of model response to the parameter distribution. The approach does not assume a linear or other explicit functional relationship between the response and the input parameters, and the sensitivity is dependent upon the magnitude of the response. The feature of this method is that it allows calculating sensitivity over a wide range of the response and is not limited to the mean value. This approach is more general and provides more information than the regression-based methods. In the alternative approach, the responses and the corresponding random samples of the parameters are ordered so that for each selected response percentile, a corresponding subset of the ordered samples can be identified. Based on a reliability sensitivity concept [1], the response CDF is defined as the integral of the joint probability-density-function of the parameters, with a domain of integration defined by a subset of the samples. The response CDF sensitivities are then calculated from the derivatives of the probability integral. The derivatives are statistically estimated from the samples and used to identify and rank the influential random variables.

Section 2 presents the major functions and components of the performance assessment model for a geologic high-level nuclear waste repository, to give the reader a background of the computer model to be analyzed. Section 3 presents the details of the CDF-based method that is used to conduct sensitivity analysis on the computer model. The application of the CDF-based method to the performance assessment model is presented in Section 4 along with verification of the results, and the conclusion is presented in Section 5.

2. The performance assessment computer model

The performance assessment computer model used in this study was developed by the NRC and the CNWRA to study the uncertainty in estimating the post-closure performance of the proposed high-level waste (HLW) repository at Yucca Mountain (YM) over long time periods (e.g. 10,000 years). The computer code, referred to as the total-system performance assessment (TPA) code [2,3], considers uncertainties and spatial variability of system attributes, model parameters, and future system states (i.e. scenario classes). To capture spatio-temporal and temporal variability and future system state, the TPA code operates in a probabilistic manner (i.e. input parameter sampled from assigned probability distributions). The TPA code estimates dose from released radionuclides during specified time periods (e.g. regulatory compliance TPI) at designated receptor locations. This paper presents only a brief discussion of the processes modeled in the code. For a complete description of the methods and assumptions, the reader is referred to Mohanty and McCartin [2,3].

The TPA code simulates physical and chemical processes of the repository system. The calculations of the most likely scenario involve the degradation of a waste package (WP) in which HLW (e.g. spent nuclear fuel) is disposed approximately 300 m below the surface of YM in the engineered barrier system (EBS), the release of radionuclides when the water infiltrating the ground surface contacts exposed SNF and transports the radionuclides through the partially water-saturated geologic medium beneath the repository and subsequently in the saturated zone to a designated receptor group 20 km down-gradient of the repository. The system-level computation involves estimation of (i) time varying precipitation resulting from climate changes, water percolation from the land surface to the subsurface and subsequently into the emplacement drifts and onto WPs; (ii) the evolution of the chemical and physical processes (e.g. temperature, humidity, pH, chloride ion concentration, and carbonate ion concentration) affecting engineered barrier degradation including the WPs and radionuclide releases; (iii) the time of failure and the number of WPs failed based on corrosion phenomena and rockfalls in the unbackfilled repository; (iv) rockfalls induced by seismic events; (v) the rate of radionuclide release from the EBS to the groundwater pathway as a function of time; (vi) the rate of flow and transport of radionuclides through the unsaturated zone and the saturated zone (water tables); and (vii) radioactive dose to a designated receptor group from groundwater based on the group's lifestyle characteristics.

In addition to calculating the physical processes associated with the most likely scenario, the TPA code also computes the processes that involve high-consequence-low-probability events that include the dose consequences of (i) the WP failure from the displacement of yet unknown faults intersecting the repository; and (ii) intrusive igneous activities (WPs failed in place after coming in contact with