

PD ISO/TR 16250:2013



BSI Standards Publication

Road vehicles — Objective rating metrics for dynamic systems

bsi.

...making excellence a habit.
AVA Documents™
Document Sharing Hub

National foreword

This Published Document is the UK implementation of ISO/TR 16250:2013.

The UK participation in its preparation was entrusted to Technical Committee AUE/15, Safety related to vehicles.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2013. Published by BSI Standards Limited 2013

ISBN 978 0 580 73294 2

ICS 43.020; 43.180

Compliance with a British Standard cannot confer immunity from legal obligations.

This Published Document was published under the authority of the Standards Policy and Strategy Committee on 31 July 2013.

Amendments issued since publication

Date	Text affected
------	---------------

Road vehicles — Objective rating metrics for dynamic systems

*Véhicules routiers — Mesures pour l'évaluation objective des
systèmes dynamiques*



Reference number
ISO/TR 16250:2013(E)

AVA **Documents**
Document Sharing Hub
© ISO 2013



COPYRIGHT PROTECTED DOCUMENT

© ISO 2013

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword	iv
Introduction	v
1 Scope	1
2 Terms and definitions	1
3 Symbols and abbreviated terms	1
3.1 General abbreviated terms	1
3.2 General symbols and subscripts	2
3.3 CORA	2
3.4 EARTH and EEARTH	3
3.5 Model reliability metric	4
3.6 Bayesian confidence metric	5
3.7 Overall ISO rating	5
4 General requirements to the data	5
5 CORA metric	6
5.1 Corridor rating	6
5.2 Cross-correlation rating	8
5.3 Step-by-step procedure	10
6 EARTH metric	11
6.1 EARTH phase score	12
6.2 EARTH magnitude score	13
6.3 EARTH slope score	14
6.4 Overall EARTH score	15
6.5 Step-by-step procedure	15
7 Model reliability metric	16
8 Bayesian confidence metric	16
9 ISO metric	18
9.1 CORA corridor method	18
9.2 EEARTH method	18
9.3 Calculation of the overall ISO rating	23
9.4 Meaning of the objective rating score	24
10 Pre-processing of the data	24
10.1 Sampling rate	25
10.2 Filtering	25
10.3 Interval of evaluation	25
11 Limitations	26
11.1 Type of signals	26
11.2 Metrics validation	26
11.3 Meaning of the results	26
11.4 Multiple responses	27
Annex A (informative) Child restraint example	28
Annex B (informative) Sled test example	46
Annex C (informative) Case studies	51
Bibliography	65

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2. www.iso.org/directives

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received. www.iso.org/patents

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

The committee responsible for this document is ISO/TC 22, *Road vehicles*, Subcommittee SC 10, *Impact test procedures*, and SC 12, *Passive safety crash protection systems*.

Introduction

Computer-Aided Engineering (CAE) has become a vital tool for product development in the automobile industry. Various computer programs and models are developed to simulate dynamic systems. To maximize the use of these models, their validity and predictive capabilities need to be assessed quantitatively. Model validation is the process of comparing CAE model outputs with test measurements in order to assess the validity or predictive capabilities of the CAE model for its intended usage. The fundamental concepts and terminology of model validation have been established mainly by standard committees including the United States Department of Energy (DOE),^[6] the American Institute of Aeronautics and Astronautics (AIAA),^[1] the Defense Modeling and Simulation Office (DMSO) of the US Department of Defense (DOD),^[5] the American Society of Mechanical Engineers Standards Committee (ASME) on verification and validation of Computational Solid Mechanics,^[2] Computational Fluid Dynamics and Heat Transfer,^[3] and various other professional societies.^{[4][22][23]}

One of the critical tasks to achieve quantitative assessment of models is to develop a validation metric that has the desirable metric properties to quantify the discrepancy between functional or time history responses from both physical test and simulation result of a dynamic system.^{[7][19][20]} Developing quantitative model validation methods has attracted considerable researchers' interest in recent years.^{[12][13][14][18][20][21][26][28][29][32]} However, the primary consideration in the selection of an effective metric should be based on the application requirements. In general, the validation metric is a quantitative measurement of the degree of agreement between the physical test and simulation result.

In this Technical Report, four state-of-the-art objective rating metrics are investigated and they are: CORrelation and Analysis (CORA) metric,^{[10][30][31]} Error Assessment of Response Time Histories (EARTH) metric,^{[28][34]} model reliability metric,^{[18][27][35]} and Bayesian confidence metric.^{[14][16][36]} Multiple dynamic system examples for both tests and CAE models are used to show their advantages and limitations. Further enhancements of the CORA corridor rating and the development of an Enhanced Error Assessment of Response Time Histories (EEARTH) metric are proposed to improve the robustness of these metrics. A new combined objective rating metric is developed to standardize the calculation of the correlation between two time history signals of dynamic systems. Multiple vehicle safety case studies are used to demonstrate the effectiveness and usefulness of the proposed metric for an ISO Technical Report.

Road vehicles — Objective rating metrics for dynamic systems

1 Scope

This Technical Report specifies a method to calculate the level of correlation between two non-ambiguous signals. The focus of the methods described in this Technical Report is on the comparison of time-history signals or functional responses obtained in all kinds of tests of the passive safety of vehicles and the corresponding numerical simulations. It is validated with signals of various kinds of physical loads such as forces, moments, accelerations, velocities, and displacements. However, other applications might be possible too, but are not in the scope of this Technical Report.

2 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

2.1

filtering

smoothing of signals by using standardized algorithms

2.2

goodness or level of correlation

similarity of two signals

2.3

interval of evaluation

time domain that is used to calculate the correlation between two signals

2.4

rating

rating score

calculated value that represents a certain level of correlation (objective rating)

2.5

sampling rate

recording frequency of a signal

2.6

time sample

pair values (e.g. time and amplitude) of a recorded signal

2.7

time-history signal

physical value recorded in a time domain; those signals are non-ambiguous

3 Symbols and abbreviated terms

3.1 General abbreviated terms

CAE Computer-Aided Engineering

CORA CORrelation and Analysis

DTW Dynamic Time Warping

EARTH	Error Assessment of Response Time Histories
EEARTH	Enhanced Error Assessment of Response Time Histories
SME	Subject Matter Expert

3.2 General symbols and subscripts

$C, C(t)$	analysed signal (CAE signal)
$T, T(t)$	reference signal (test signal)
t	time signal (axis of abscissa)
Δt	interval between two time samples
t_0	time zero of an event (e.g. test, crash, impact, etc.)
t_{start}	starting time of the interval of evaluation
t_{end}	ending time of the interval of evaluation
N	total number of sample points (e.g. time steps) between the starting time, t_{start} , and ending time, t_{end}
$N_{>0}$	all natural numbers without zero

3.3 CORA

Z_{CORA}	CORA rating
Z_1	corridor rating
$Z_1(t)$	corridor rating at time t (curve)
Z_2	cross-correlation rating
Z_{2a}	phase-shift rating
Z_{2b}	size rating
Z_{2c}	shape (progression) rating
w_{Z1}	weighting factor of the corridor rating, Z_1
w_{Z2}	weighting factor of the cross-correlation rating, Z_2
w_{Z2a}	weighting factor of the phase-shift rating, Z_{2a}
w_{Z2b}	weighting factor of the size rating, Z_{2b}
w_{Z2c}	weighting factor of the shape rating, Z_{2c}
k_{Z1}	exponent factor for calculating the corridor rating between the inner and outer corridors
k_{Z2a}	exponent factor for calculating phase-shift rating, Z_{2a}
k_{Z2b}	exponent factor for calculating size rating, Z_{2b}
k_{Z2c}	exponent factor for calculating shape rating, Z_{2c}
T_{norm}	absolute maximum amplitude of the reference signal, T
a_0	relative half width of the inner corridor

b_0	relative half width of the outer corridor
δ_i	half width of the inner corridor
δ_o	half width of the outer corridor
$\delta_i(t)$	lower/upper inner corridor at time t (curve)
$\delta_o(t)$	lower/upper outer corridor at time t (curve)
$D_{\min}$	coefficient of the allowable lower limit of the phase shift
$D_{\max}$	coefficient of the allowable upper limit of the phase shift
F_C	sum of the square of the area for the time-shifted evaluated curve, C
F_T	sum of the square of the area for the reference curve, T
$INT_{\min}$	percentage of the minimum remaining overlapping time of the reference and evaluation curves after time shift
m	shift of a signal along the axis of abscissa
$m_{\min}$	minimum m shift of a signal
$m_{\max}$	maximum m shift of a signal
n	number of samples
$n_{\min}$	time step shifted to get the maximum cross correlation
ρ	cross correlation
$\rho(m)$	cross correlation at shift, m
δ	phase-shift time at the maximum cross correlation, ρ
$\delta_{\min}$	lower limit of CORA phase shift
$\delta_{\max}$	upper limit of CORA phase shift

3.4 EARTH and EEARTH

E_E	overall EARTH score
E_M	EARTH magnitude score
E_P	EARTH phase score
E_S	EARTH slope (topology) score
w_M	weighting factor of the magnitude score, E_M
w_P	weighting factor of the phase score, E_P
w_S	weighting factor of the slope score, E_S
k_M	exponent factor for calculating the magnitude score, E_M
k_P	exponent factor for calculating the phase score, E_P
k_S	exponent factor for calculating the slope score, E_S
ε_{mag}	EARTH magnitude error

ε_{slope}	EARTH slope error
ε_M^*	maximum allowable magnitude error
ε_P^*	maximum allowable percentage of time shift
ε_S^*	maximum allowable slope error
$\bar{C}(t)$	mean value of CAE curve
$C^{ts}, C^{ts}(i)$	truncated and shifted CAE curve
C^{ts+d}	derivative CAE curve, C^{ts}
C^{ts+w}	warped CAE curve, C^{ts}
C^{ts+d+w}	derivative warped CAE curve, C^{ts}
$\bar{T}(t)$	mean value of test curve
$T^{ts}, T^{ts}(j)$	truncated and shifted test curve
T^{ts+d}	derivative test curve, T^{ts}
T^{ts+w}	warped test curve, T^{ts}
T^{ts+d+w}	derivative warped test curve, T^{ts}
ρ_E	maximum cross correlation of all $\rho_L(m)$ and $\rho_R(m)$
$\rho_L(m)$	cross correlation — signal is moved to the left
$\rho_R(m)$	cross correlation — signal is moved to the right
$d(i, j)$	local cost function to perform the dynamic time warping
m	time steps moved to evaluate the EARTH phase error
n_ε	number of time shifts to get, ρ_E

3.5 Model reliability metric

T_{norm}	absolute maximum amplitude of the reference signal, T
a	reliability target
b	threshold factor of the reliability assessment
ε_L	lower bound of the threshold interval
ε_U	upper bound of the threshold interval
ε_Φ^L	lower bound of the Bayesian interval hypothesis in probabilistic principal component analysis space (PPCA)
ε_Φ^U	lower bound of the Bayesian interval hypothesis in PPCA space
r	model reliability
P	cumulative probability

$\Phi, \Phi(t)$ $p \times n$ reduced data matrix

3.6 Bayesian confidence metric

A constant vector

B_{iM} Bayes factor for multivariate case

δ likelihood function

ε_Φ predefined threshold vector

H_0 null hypothesis

H_1 alternative hypothesis

K confidence of accepting the model

κ measure of confidence within an interval

Λ variance of error variable, ε_i^*

μ_Φ p mean values obtained from Φ^*

N normal distribution

$N(\rho, \Lambda)$ normal distribution of δ with mean vector, p , and variance matrix, Λ

π_0 prior probability of hypothesis

ρ prior mean, δ

Σ_Φ variance matrix of Φ^*

$\Phi, \Phi(t)$ difference curve between test curve, T , and CAE curve, C

$f(\delta)$ prior density function of δ

3.7 Overall ISO rating

R combined rating of EEARTH and the CORA corridor method

E EEARTH rating score

Z CORA corridor rating ($Z = Z_1$)

w_E weighting factor of the EEARTH rating, E

w_Z weighting factor of the CORA corridor rating, Z

r rank of the sliding scale of the ISO metric

$SC_{lower}(r)$ lower threshold of rank, r

$SC_{upper}(r)$ upper threshold of rank, r

4 General requirements to the data

The metrics described in this Technical Report require non-ambiguous curves (e.g. time-history curves). Furthermore, it is required that the reference curve, $T(t)$, and the evaluated curve, $C(t)$, are both defined between starting time, t_{start} , and ending time, t_{end} . Both curves shall have the same number of sample points, N , with a constant time interval, Δt , within the evaluation interval.

5 CORA metric

The objective evaluation metric called CORA — correlation and analysis^{[10][30][31]} — uses two independent sub-ratings, a corridor rating, and a cross-correlation rating to assess the correlation of two signals. The rating structure of CORA is shown in [Figure 1](#).

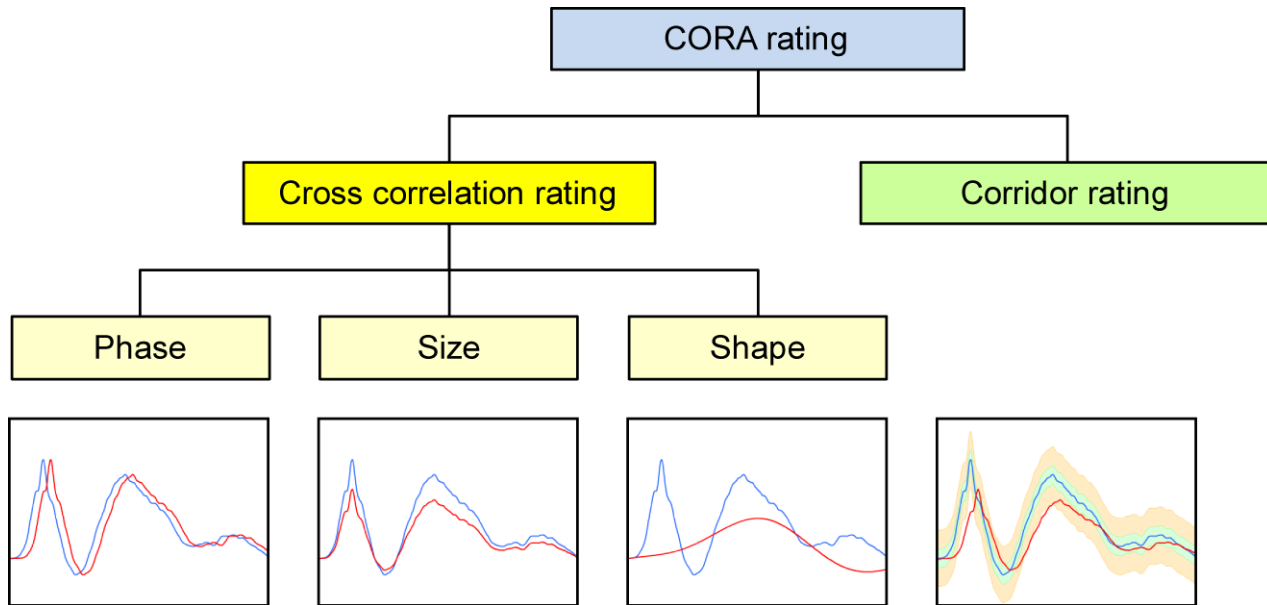


Figure 1 — CORA rating structure

The corridor and cross-correlation ratings are used to compensate each other's disadvantages, and the CORA rating tool is trying to separate an engineer's knowledge from the objective rating metric by using external parameters. However, it is possible to fine-tune the evaluation to the specific needs of the applications by adjusting those metric parameters to reflect the SME's knowledge of the applications.

The corridor rating, Z_1 , calculates the deviation between both curves with the help of user-defined or automatically generated corridors. The cross-correlation rating, Z_2 , analyses specific curve characteristics, such as phase shift, Z_{2a} , size, Z_{2b} , and shape of the signals, Z_{2c} . The rating results range from "0" (no correlation) to "1" (perfect match). The influence of the sub-ratings on the global rating is adjusted by user-defined weighting factors. Formulae (1) and (3) show how to calculate the CORA rating by using weighting factors [see Formulae (2) and (4)]. Details of each sub-rating are introduced in the following subsections.

$$Z_{CORA} = w_{Z1} \cdot Z_1 + w_{Z2} \cdot Z_2 \quad (1)$$

$$w_{Z1} + w_{Z2} = 1 \quad (2)$$

$$Z_2 = w_{Z2a} \cdot Z_{2a} + w_{Z2b} \cdot Z_{2b} + w_{Z2c} \cdot Z_{2c} \quad (3)$$

$$w_{Z2a} + w_{Z2b} + w_{Z2c} = 1 \quad (4)$$

5.1 Corridor rating

The corridor rating calculates the deviation between two signals by means of corridor fitting. The two sets of corridors, the inner and the outer corridors, are defined along the mean curve. If the evaluated curve (e.g. CAE curve) is within the inner corridor bounds, a score of "1" is given, and if it is outside the outer corridors, the rating is set to "0". The assessment declines from "1" to "0" between the bounds of

inner and outer corridors resulting in three different rating zones as shown in Figure 2. This transition is user-defined. The compliance with the corridors is calculated at each specific time, t , and the final corridor rating, Z_1 , of a signal is the average of all ratings, $Z_1(t)$, at specific times, t .

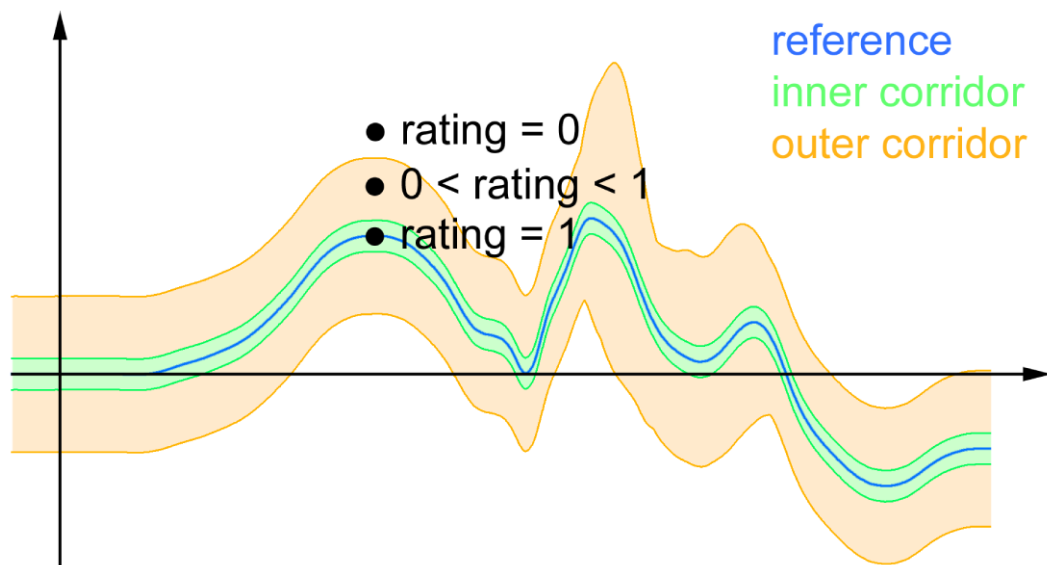


Figure 2 — Rating zones of the corridor method (corridors of constant width)^[10]

The philosophy is to use a narrow inner corridor and a wide outer corridor.^[17] It limits the number of “1” ratings to only good correlations and gives the opportunity to distinguish between poor and fair correlations. If the outer corridor is too narrow, too many curves of a fair or moderate correlation would get the same poor rating of “0”, like signals of almost no correlation with the reference. The width of the corridors can be adjusted in order to reflect the specific signal characteristic, and it can be constant for the whole duration of the dynamic responses or vary at the different time steps.

This Technical Report applies the most common approach of using the constant corridor widths for the whole duration of the dynamic response. The parameters a_0 and b_0 define the relative half width of the inner and the outer corridors. Both shall be between “0” and “1”, and a_0 must be less than b_0 . The absolute half widths of both corridors are defined as the product of relative half width and the absolute maximum amplitude, T_{norm} , of the reference signal, T . Formula (5) shows the calculation of T_{norm} .

$$T_{norm} = \max\{|\min(T)|, |\max(T)|\} \quad (5)$$

The absolute half width of the inner corridors (absolute distance from reference signal to outer bounds of the inner corridors) is defined by Formula (6). The calculation of the absolute half width of the outer corridors [see Formula (7)] is similar to that of the inner corridors.

$$\delta_i = a_0 \cdot T_{norm} \quad 0 \leq a_0 \leq 1 \quad (6)$$

$$\delta_o = b_0 \cdot T_{norm} \quad 0 \leq b_0 \leq 1 \quad \text{and} \quad a_0 < b_0 \quad (7)$$

Based on these definitions, the upper and lower bounds of the inner corridors are defined by Formula (8) and the upper and lower bounds of the outer corridors are defined by Formula (9).

$$\delta_i(t) = T(t) \pm \delta_i \quad (8)$$

$$\delta_o(t) = T(t) \pm \delta_o \quad (9)$$

Formula (10) shows the calculation of the corridor rating for the correlation between the reference signal, T , and the analysed signal, C , at each evaluation time, t . If the absolute difference between the signals T and C is less than the half width of the inner corridors, δ_i , then the rating is set to "1". The rating is calculated by Formula (10) when the absolute difference between both signals is in between $\delta_i \leq |T(t) - C(t)| \leq \delta_o$. If the absolute difference between both signals is greater than the half width of the outer corridors, δ_o , then the rating is set to "0". The parameter k_{Z1} assesses the location of the analysed signal within the outer corridor and it applies the appropriate penalty on the rating score. A linear ($k_{Z1} = 1$), quadratic ($k_{Z1} = 2$), cubical ($k_{Z1} = 3$), or any other regression relationship can be defined accordingly.

$$Z_1(t) = \begin{cases} 1 & \text{if } |T(t) - C(t)| < \delta_i \\ \left(\frac{\delta_o - |T(t) - C(t)|}{\delta_o - \delta_i} \right)^{k_{Z1}} & \text{if } \delta_i \leq |T(t) - C(t)| \leq \delta_o \\ 0 & \text{if } |T(t) - C(t)| > \delta_o \end{cases} \quad k_{Z1} \in N_{>0} \quad (10)$$

The final corridor rating, Z_1 , is calculated by averaging all single time step ratings, $Z_1(t)$, as shown in Formula (11). The parameter N represents the total number of sample points (e.g. time steps) between the starting and ending times of the interval of evaluation.

$$Z_1 = \frac{\sum_{t=t_{start}}^{t_{end}} Z_1(t)}{N} \quad (11)$$

One of the advantages of the corridor rating is the simplicity and the clearness of the algorithm. It reflects criteria which are used intuitively in engineering judgment. Sometimes, this simplicity may be the disadvantage of the method. For example, a small distortion of the phase can lead to a undesirable rating.

5.2 Cross-correlation rating

The cross-correlation rating may compensate for the disadvantages resulting from the corridor rating by analysing the characteristics of signals. Three sub-ratings (phase, size, and shape) with individual weighting factors are implemented.

The calculated maximum cross correlation is the base of the analysis of phase shift, size, and shape of the signals.

5.2.1 Maximum cross correlation

In general, the cross-correlation metric moves the test signal, T , by multiples of Δt in relation to the CAE signal, C . The cross-correlation value, $\rho(m)$, is calculated at each shifted state. The time shift with the maximum ρ is the base of the calculation of the cross-correlation rating.

Formula (12) shows the calculation of the cross correlation at each time shift, m . Curve C is moved by multiples $m \in (m_{min}, m_{max})$ of Δt between the minimum and the maximum shift [see Formulae (13) and (14)]. The range of the time shift is limited by the parameter INT_{min} . Therefore, the signals T and C are

at least overlapping within the interval $INT_{\min} \cdot (t_{\text{end}} - t_{\text{start}})$. The parameter n in Formula (12) is not constant but is reduced to $n_{\min}$ [see Formula (15)].

$$\rho(m) = \frac{\sum_{i=0}^{n-1} C(t_{\text{start}} + (m+i) \cdot \Delta t) \cdot T(t_{\text{start}} + i \cdot \Delta t)}{\sqrt{\sum_{i=0}^{n-1} C^2(t_{\text{start}} + (m+i) \cdot \Delta t) \cdot \sum_{i=0}^{n-1} T^2(t_{\text{start}} + i \cdot \Delta t)}} \quad -1 \leq \rho \leq 1 \quad (12)$$

$$m_{\min} = \frac{(INT_{\min} - 1) \cdot (t_{\text{end}} - t_{\text{start}})}{\Delta t} \quad 0 < INT_{\min} < 1 \quad (13)$$

$$m_{\max} = \frac{(1 - INT_{\min}) \cdot (t_{\text{end}} - t_{\text{start}})}{\Delta t} \quad 0 < INT_{\min} < 1 \quad (14)$$

$$n_{\min} = INT_{\min} \cdot n \quad (15)$$

As described above, the cross correlation ρ is the maximum of all $\rho(m)$ [see Formula (16)].

$$\rho = \max\{\rho(m)\} \quad (16)$$

5.2.2 Phase-shift rating

The phase-shift rating requires the parameters $D_{\min}$ and $D_{\max}$ to limit the phase shift. Both are defined within (0, 1). The thresholds of the phase shift are calculated by Formulae (17) and (18).

$$\delta_{\min} = D_{\min} \cdot (t_{\text{end}} - t_{\text{start}}) \quad 0 < D_{\min} \leq 1 \quad (17)$$

$$\delta_{\max} = D_{\max} \cdot (t_{\text{end}} - t_{\text{start}}) \quad 0 < D_{\max} \leq 1 \quad (18)$$

The phase-shift rating is calculated by Formula (19) at the maximum cross correlation, ρ , and the corresponding time shift, δ . The parameter k_{Z2a} describes the decline of the rating between “1” and “0”. A linear ($k_{Z2a} = 1$), quadratic ($k_{Z2a} = 2$), cubical ($k_{Z2a} = 3$), or any other regression relationship can be defined accordingly.

$$Z_{2b} = \begin{cases} 1 & \text{if } |\delta| < \delta_{\min} \\ \left(\frac{|\delta_{\max} - |\delta||}{\delta_{\max} - \delta_{\min}} \right)^{k_{Z2a}} & \text{if } \delta_{\min} \leq |\delta| \leq \delta_{\max} \\ 0 & \text{if } |\delta| > \delta_{\max} \end{cases} \quad k_{Z2a} \in N_{>0} \quad (19)$$

5.2.3 Size rating

The size of the signals is analysed by comparing the area below the two curves after the phase shift. It is a necessary evaluation but it may not be sufficient to evaluate the overall level of correlation. For instance, the area below a signal with high and narrow peak could be identical to the area of a curve

with low but wide peak. The size method would evaluate this example with “1” although the shape of the signals is completely different.

$$\frac{F_C}{F_T} = \frac{\sum_{i=1}^n C^2(t_{start} + \delta + i \cdot \Delta t)}{\sum_{i=1}^n T^2(t_{start} + i \cdot \Delta t)} \quad (20)$$

$$Z_{2b} = \begin{cases} \left(\frac{F_C}{F_T} \right)^{k_{Z2b}} & \text{if } F_T > F_C \\ \left(\frac{F_T}{F_C} \right)^{k_{Z2b}} & \text{if } F_T \leq F_C \end{cases} \quad k_{Z2b} \in N_{>0} \quad (21)$$

The analysis of size is done by comparing the square of the areas between the curves and the time axis. Due to equidistant supporting points of both signals, the area can be represented by Formula (20). Finally, the size rating is calculated by Formula (21). The meaning of k_{Z2b} is similar to k_{Z2a} (see 5.2.2).

5.2.4 Shape rating

As shown in Formula (22), the shape rating, Z_{2c} , is derived from the maximum cross correlation, ρ , described in 5.2.1. The meaning of k_{Z2c} is similar to k_{Z2a} (see 5.2.2).

$$Z_{2c} = \frac{1}{2}(\rho + 1)^{k_{Z2c}} \quad k_{Z2c} \in N_{>0} \quad (22)$$

5.3 Step-by-step procedure

The signals shall be pre-processed as described in [Clause 10](#). After preparing the signals for the analysis and defining the interval of evaluation, the maximum absolute amplitude, T_{norm} , of the reference signal, T , shall be determined within this interval. It is used to calculate the inner and outer corridors. The actual corridor and cross-correlation assessment can be executed within the defined interval. The overall rating ranges between “0” and “1”. A score of “1” does not mean that both signals are identical. Solely, their correlation is mathematically perfect within the defined tolerances.

To summarize, the following step-by-step procedures shall be followed to calculate CORA rating:

- 1) Pre-process both signals according to [Clause 10](#).
- 2) Calculate T_{norm} within this interval by using the reference signal.
- 3) Calculate the inner and the outer corridors.
- 4) Calculate the corridor rating, $Z_1(t)$, at every specific time, t , within the interval of evaluation.
- 5) Calculate the corridor rating, Z_1 , based on $Z_1(t)$ and the number, N , of time samples.
- 6) Calculate the maximum cross correlation, ρ , between T and C .
- 7) Calculate phase rating, Z_{2a} .
- 8) Calculate size rating, Z_{2b} .
- 9) Calculate shape rating, Z_{2c} .
- 10) Calculate the cross-correlation rating, Z_2 .

11) Calculate the overall CORA rating, Z_{CORA} .

6 EARTH metric

Another objective rating metric called Error Assessment of Response Time Histories (EARTH) was developed for dynamic system applications.[28][34] Figure 3 shows the structure of the EARTH error measures.

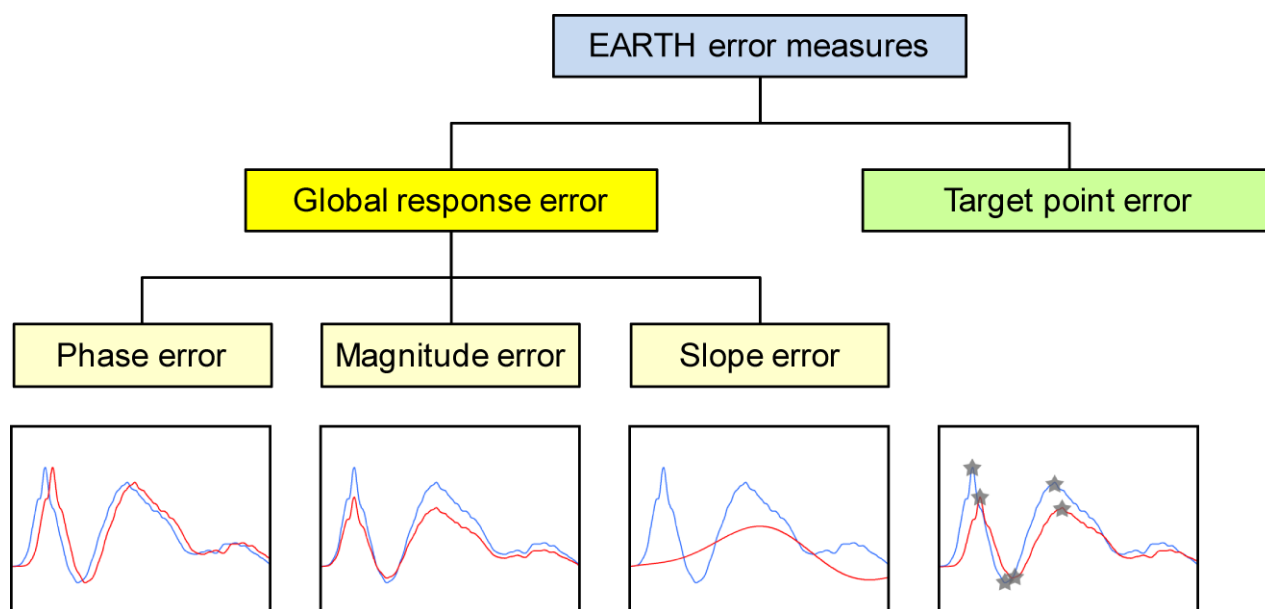


Figure 3 — EARTH error measures structure[28]

The EARTH metric is divided into two categories: global response error and target point response error. The global response error is defined as the error associated with the complete time history with equal weight on each point. The three main components of the global response error are phase error, magnitude error, and topology (or so-called slope) error. The target point error is defined as the error associated with a certain localized phenomenon of interest, such as peak error and time-to-peak error. The target point error represents the characteristic of a part of the time history, but does not indicate an overall performance of the entire time history. In addition, the target point error is generally application-dependent; hence, it is not the focus of this Technical Report.

Quantifying the errors associated with these features of phase, magnitude, and topology (slope) separately is challenging because there are strong interactions among them. For example, to quantify the error associated with magnitude, the presence of a phase difference between the time histories may result in a misleading measurement. A unique feature dynamic time warping (DTW)[25] of the EARTH metric is used to separate the interaction of phase, magnitude, and topology (slope) errors. DTW is an algorithm for measuring discrepancy between time histories. It aligns peaks and valleys as much as possible by expanding and compressing the time axis according to a given cost (distance) function.[9]

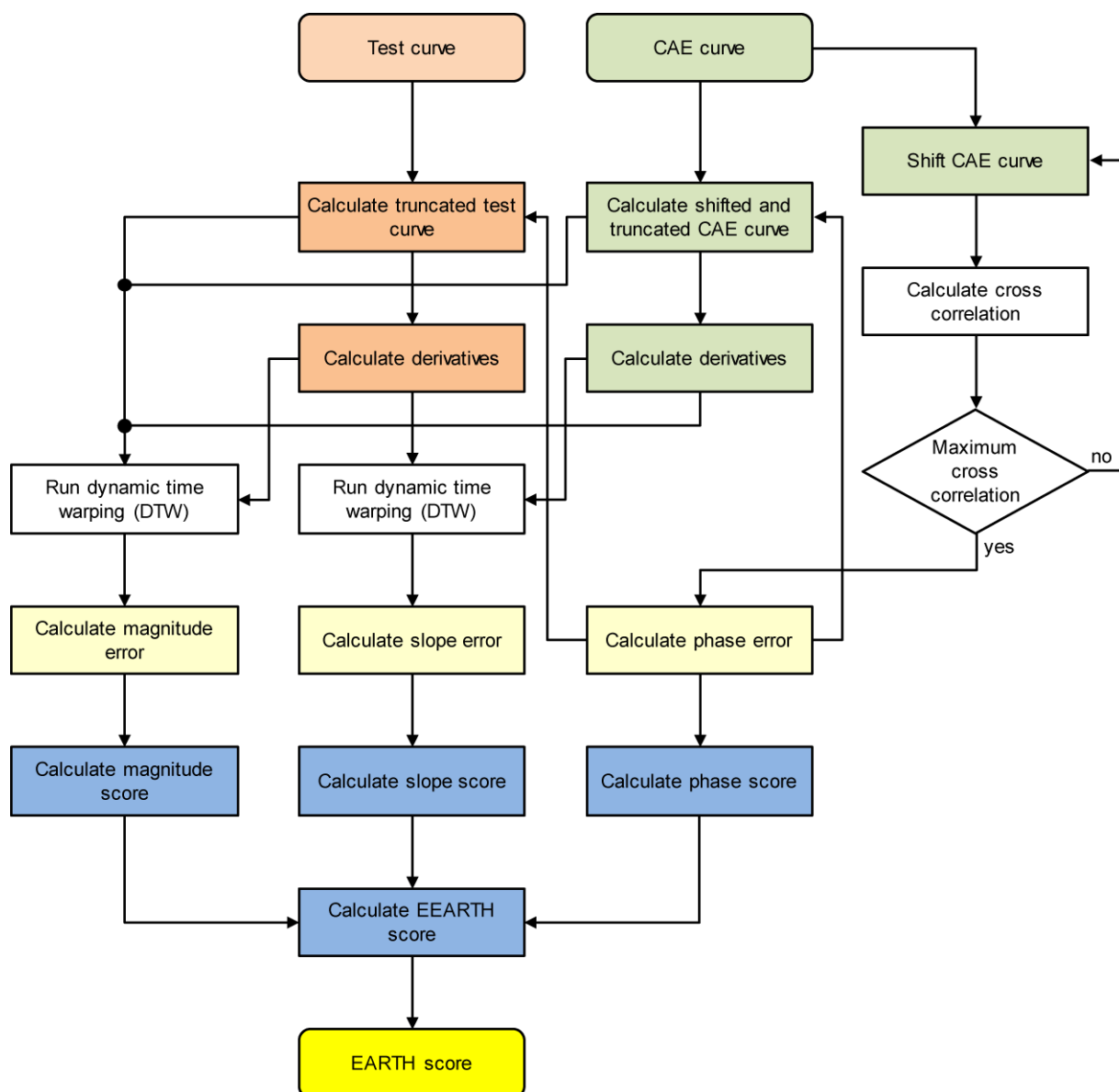


Figure 4 — Workflow of the revised EARTH metric

Since the ranges of the three errors are quite different and there is no single rating that can provide a quantitative assessment alone, the initial EARTH metric study employs a linear regression method to combine the three errors into one score. Numerical optimization method is employed to identify the linear coefficients so that the resulting EARTH rating can match the SME's ratings closely for a specific application.[8][24] However, the resulting linear combination of the EARTH metric is mainly numerical-based and application-dependent; hence, it may not be scalable to other applications. In order to provide more intuitive ratings while maintaining the advantages of the original EARTH metric, a revised EARTH metric is proposed.[28] Figure 4 shows the flowchart of the revised EARTH metric, and the details of the algorithms are described in the following subclauses.

6.1 EARTH phase score

The EARTH phase score, E_P , is used to measure the phase lag between the two time histories. The maximum allowable percentage of time shift is ε_P^* , and it is predefined. In this step, the initial curve, C , is shifted left then right one step at a time to the original test data, and the cross correlation between the

truncated test curve and shifted and truncated C are calculated until reaching the maximum allowable time-shift limits $\varepsilon_P^* \cdot (t_{end} - t_{start})$.

When the initial curve, C , is moved to the left by m time steps, the number of overlap points of the two time histories after time shift $m \cdot \Delta t$ is reduced to $N - m$ and the corresponding cross-correlation value, $\rho_L(m)$, is calculated by Formula (23).

$$\rho_L(m) = \frac{\sum_{i=0}^{n-1} [(C(t_{start} + (m+i) \cdot \Delta t) - \bar{C}(t)) \cdot (T(t_{start} + i \cdot \Delta t) - \bar{T}(t)))]}{\sqrt{\sum_{i=0}^{n-1} [C(t_{start} + (m+i) \cdot \Delta t) - \bar{C}(t)]^2} \cdot \sqrt{\sum_{i=0}^{n-1} [T(t_{start} + i \cdot \Delta t) - \bar{T}(t)]^2}} \quad (23)$$

When the initial curve, C , is moved to the right by m time steps, the number of overlap points after time shift $m \cdot \Delta t$ is reduced to $N - m$ and the corresponding cross-correlation value, $\rho_R(m)$, is calculated by Formula (24).

$$\rho_R(m) = \frac{\sum_{i=0}^{n-1} [(C(t_{start} + i \cdot \Delta t) - \bar{C}(t)) \cdot (T(t_{start} + (m+i) \cdot \Delta t) - \bar{T}(t)))]}{\sqrt{\sum_{i=0}^{n-1} [C(t_{start} + i \cdot \Delta t) - \bar{C}(t)]^2} \cdot \sqrt{\sum_{i=0}^{n-1} [T(t_{start} + (m+i) \cdot \Delta t) - \bar{T}(t)]^2}} \quad (24)$$

The maximum cross correlation, ρ_E , is the maximum of all $\rho_L(m)$ and $\rho_R(m)$. The number of the time-shifting steps that yields the maximum cross correlation, ρ_E , is defined as the EARTH phase error, n_ε . The corresponding shifted and truncated CAE curve C is recorded as C^{ts} and the corresponding truncated test curve is recorded as T^{ts} .

The EARTH phase score, E_P , is calculated by Formula (25). The best EARTH phase score is “1”, which means there is no need to shift CAE curve to reach the maximum cross correlation between the initial test and CAE curves. If the time shift, n_ε , is equal to or greater than the maximum allowable time-shift threshold $\varepsilon_P^* \cdot N$, then the EARTH phase score is “0”. In between, the EARTH phase score is calculated by a regression method; it is either linear ($k_P = 1$), quadratic ($k_P = 2$), or cubical ($k_P = 3$).

$$E_P = \begin{cases} 1 & \text{if } n_\varepsilon = 0 \\ \left(\frac{\varepsilon_P^* \cdot N - n_\varepsilon}{\varepsilon_P^* \cdot N} \right)^{k_P} & \text{if } 0 < n_\varepsilon < \varepsilon_P^* \cdot N \\ 0 & \text{if } n_\varepsilon \geq \varepsilon_P^* \cdot N \end{cases} \quad k_P \in \{1, 2, 3\} \quad (25)$$

6.2 EARTH magnitude score

The magnitude error is a measure of discrepancy in the amplitude of the two time histories. The magnitude error is defined as the difference in amplitude of the two time histories when there is no time lag between them. Before calculating the magnitude error, the difference between the time histories caused by error in phase and topology (slope) are minimized by using dynamic time warping. The local cost function, $d(i, j)$, for DTW used considers both the distance and the slope as shown in Formula (26). Once the local cost matrix is built, the algorithm finds the alignment path which runs through the low-

cost areas on the cost matrix. This alignment path defines the corresponding elements of both $C^{ts}(i)$ and $T^{ts}(j)$ that will lead to the minimum accumulated cost function.

$$d(i, j) = (C^{ts}(i) - T^{ts}(j))^2 + (t_i - t_j)^2 \cdot \left| \left(\frac{dC^{ts}(i)}{dt} \right)_{t=t_i} - \left(\frac{dT^{ts}(j)}{dt} \right)_{t=t_j} \right| \quad (26)$$

The derivatives at each data points of the truncated test curve, T^{ts} , and the shifted and truncated CAE curve, C^{ts} , are first calculated, which yields the truncated test slope curve, T^{ts+d} , and the shifted and truncated CAE slope curve, C^{ts+d} . Next, DTW is performed to the truncated test curve, T^{ts} , and the shifted and truncated CAE curve, C^{ts} , which results in the truncated and warped test slope curve, T^{ts+w} , and the shifted, truncated, and warped CAE slope curve, C^{ts+w} . The EARTH magnitude error, ε_{mag} , is calculated by Formula (27).

$$\varepsilon_{mag} = \frac{\|C^{ts+w} - T^{ts+w}\|_1}{\|T^{ts+w}\|_1} \quad (27)$$

Formula (28) is used to calculate the EARTH magnitude score, E_M , where ε_M^* is the maximum allowable magnitude error, and k_M defines the order of the regression. The best EARTH magnitude score is “1”, which means there is no difference in the amplitudes after phase shift and dynamic time warping. If the EARTH magnitude error, ε_{mag} , is equal to or greater than the maximum allowable magnitude error threshold, ε_M^* , then the EARTH magnitude score is “0”. In between, the EARTH magnitude score is calculated by regression method.

$$E_M = \begin{cases} 1 & \text{if } \varepsilon_{mag} = 0 \\ \left(\frac{\varepsilon_M^* - \varepsilon_{mag}}{\varepsilon_M^*} \right)^{k_M} & \text{if } 0 < \varepsilon_{mag} < \varepsilon_M^* \\ 0 & \text{if } \varepsilon_{mag} \geq \varepsilon_M^* \end{cases} \quad k_M \in \{1, 2, 3\} \quad (28)$$

6.3 EARTH slope score

The topological error is a measure of discrepancy in topology (slope) of the two time histories. The topology of a time history is defined by the slope at each point. In order to ensure that the effect of global time shift is minimized, the slope is calculated from the time-shifted histories T^{ts} and C^{ts} . Thus, by taking the derivative at each point, the derivative time-shifted histories, represented by T^{ts+d} and C^{ts+d} , are obtained. Dynamic time warping is performed on T^{ts+d} and C^{ts+d} , which results in the truncated and warped test slope curve, T^{ts+d+w} , and the shifted, truncated, and warped CAE slope curve, C^{ts+d+w} . These two curves are then used to calculate the EARTH slope error, ε_{slope} , by Formula (29).

$$\varepsilon_{slope} = \frac{\|C^{ts+d+w} - T^{ts+d+w}\|_1}{\|T^{ts+d+w}\|_1} \quad (29)$$

Formula (30) is used to calculate the EARTH slope score, E_S , where ε_S^* is the maximum allowable slope error, and k_S defines the order of the regression. The best EARTH slope score is “1”, which means there is no difference between the two slope curves. If the slope error, ε_{slope} , is equal to or greater than the

maximum allowable slope error, ε_S^* , then the EARTH slope score is “0”. In between, the EARTH slope score is calculated by regression method.

$$E_S = \begin{cases} 1 & \text{if } \varepsilon_{slope} = 0 \\ \left(\frac{\varepsilon_S^* - \varepsilon_{slope}}{\varepsilon_S^*} \right)^{k_S} & \\ 0 & \text{if } \varepsilon_{slope} \geq \varepsilon_S^* \end{cases} \quad k_S \in \{1, 2, 3\} \quad (30)$$

6.4 Overall EARTH score

The above three EARTH scores are combined into one EARTH score by using weighting factors shown in Formulae (31) and (32).

$$E_E = w_P \cdot E_P + w_M \cdot E_M + w_S \cdot E_S \quad (31)$$

$$w_P + w_M + w_S = 1 \quad (32)$$

The goal is to translate the initial three EARTH errors into a standardized score between “0” and “1”. Due to the similarity with the cross-correlation rating in the CORA metric, the consistent weighting factors with the CORA cross-correlation rating are used. In addition, the parameters of the EARTH metric, such as thresholds for phase, magnitude, and topological errors, are defined by matching the EARTH rating with the SME’s knowledge using a validation database and optimization technique.^[28]

6.5 Step-by-step procedure

The following step-by-step process shall be followed to calculate the EARTH rating:

- 1) Pre-process both signals according to [Clause 10](#) (T and C).
- 2) Calculate the phase error in terms of time steps, n_e , by maximizing cross correlation.
- 3) Calculate the phase score, E_P .
- 4) Calculate the shifted and truncated time history curves, T^{ts} and C^{ts} .
- 5) Calculate the slope curves of the shifted and truncated time history curves, T^{ts+d} and C^{ts+d} .
- 6) Perform dynamic time warping to the shifted and truncated time history curves to generate the shifted, truncated, warped time history curves, T^{ts+w} and C^{ts+w} .
- 7) Calculate the magnitude error, ε_{mag} , between T^{ts+w} and C^{ts+w} .
- 8) Calculate the magnitude score, E_M .
- 9) Perform dynamic time warping to the slope curves to generate the shifted, truncated, warped slope time history curves, T^{ts+d+w} and C^{ts+d+w} .
- 10) Calculate the slope error, ε_{slope} , between T^{ts+d+w} and C^{ts+d+w} .
- 11) Calculate the slope score, E_S .
- 12) Calculate the overall EARTH rating, E_E .

7 Model reliability metric

A model reliability-based validation metric was developed for dynamic system applications.^[14] Figure 5 shows the illustration of this metric. The difference between CAE and test curves is taken as the validation feature [see Formula (33)]. The threshold factor, b , is defined by the SME's experience; the lower and upper bounds of the threshold interval, $[\varepsilon_L, \varepsilon_U]$, are defined as the product of the threshold factor, b , and the absolute maximum amplitude, T_{norm} , of the reference signal, T [see Formulae (34), (35), and (36)]. A simple validation metric $r = P(\varepsilon_\Phi^L < \Phi < \varepsilon_\Phi^U)$ is defined to indicate the model reliability, which represents the probability that the observed difference, $\Phi(t)$, is within a small interval of $[\varepsilon_\Phi^L, \varepsilon_\Phi^U]$. If the adequacy or confidence requirement, $r > a$, is met, where a is a predefined reliability target, the model prediction is accepted. Since this difference time-history curve has much better normality than those of the test and CAE curves, a normal distribution of Φ can then be assumed and the model reliability metric can be simply calculated by Formula (37).

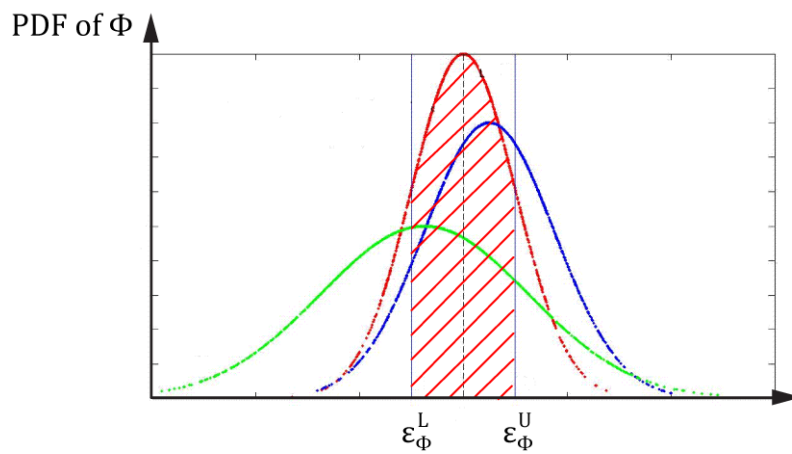


Figure 5 — Illustration of model reliability metric

$$\Phi(t) = C(t) - T(t) \quad (33)$$

$$T_{norm} = \max\{|\min(T)|, |\max(T)|\} \quad (34)$$

$$\varepsilon_\Phi^L = -b \cdot T_{norm} \quad (35)$$

$$\varepsilon_\Phi^U = b \cdot T_{norm} \quad (36)$$

$$r = P(\Phi < \varepsilon_\Phi^U) - P(\Phi < \varepsilon_\Phi^L) \quad (37)$$

There are only two adjusting parameters: threshold factor, b , and reliability target, a , and both have clear physical meanings. The model reliability metric is one of the simplest metrics for dynamic system applications, and it is easy to understand and interpret.

8 Bayesian confidence metric

A Bayesian confidence-based validation metric was developed for dynamic system applications.^[13] Similar to the model reliability metric, the difference curve, $\Phi(t)$, between the test and CAE curves are