A METHOD FOR OBTAINING THE CREDIBILITY OF A SIMULATION MODEL

Ke Fang[a], Ming Yang[b], Kaibin Zhao[c]

[a],[b],[c]Control & Simulation Center, School of Astronautics, Harbin Institute of Technology, Harbin, China

[a]hitsim@163.com, [b]myang@hit.edu.cn, [c]kaibin.zhao.HIT@hotmail.com

ABSTRACT
Model validation is one of the key problems of simulation systems V&V (Verification & Validation). How to obtain the credibility of a simulation model is still the core issue prior to its application. Based on the introduction of model credibility, a method of workflow for obtaining the credibility of a simulation model is brought forward, the essentials including factor space establishment, similarity analysis, result transformation and defect tracing are explained, then a case study which provides a walk through application of the method is presented. The method corrects and clarifies some misunderstanding in model validation, and provides a practical way to obtain the model credibility.

Keywords: model validation, V&V, credibility, factor space, similarity analysis, result transformation

1. INTRODUCTION

A simulation model uses mathematical modeling and coding techniques to build a run-able program to simulate the object in the real world. There is a quantitative similarity between the simulation model and its origin in application domain named "credibility", to measure at what level the model behaves as the origin. Research and practice have been being conducted to compute the credibility of a simulation model, which is the goal of model V&V (Verification & Validation) (Sargent 2015). However it is not an easy job.

Many recent studies focus on the problem of model validation, and propose constructive walkthroughs to conduct it. Roy addressed the uncertainty quantification problem in model validation, and proposed a comprehensive framework for model V&V (Roy 2011). You summarized several statistics similarity analysis methods and proposed a decision process in quantitative model validation (You 2012). Eek established a dependency graph of a simulator at Saab Aeronautics, and developed a quantified table for credibility evaluation (Eek 2015). Kutluay focused on vehicle dynamics simulation models, and gave a literature survey on validation approaches (Kutluay 2014). Song used Colored Petri Nets to rebuild the model of train-to-train distance measurement system, and established a multi-layer factor space to reveal the model validation aspects (Song 2017).

Since every simulation model has different modeling strategies, it is hard to propose a straight mathematical expression of credibility in general. Researchers and engineers often use decision making method to synthesize the final result, after they obtain the influence level of all attributes related to the credibility. See the expression below.

$$F = \{<N, V>; <L, A>\}$$

$$C = f(v_1, v_2, ..., v_k) \tag{1}$$

In Eq.(1), F is a factor space (Zhou, Sun, and Li 2016) created by the model decomposition, where $N = (n_1, n_2, ..., n_k\}$ is the node set, which contains all the credibility influencing factors to the model; $V = \{v_1, v_2, ..., v_k\}$ is the value set, which contains all partial credibility of each factor and maps to N; $L = \{l_1, l_2, ..., l_k\}$ is the link set, which contains all relationship between each two factors; and A is the attribute set, which contains all attributes on each link such as weight etc. and maps to L. C is the credibility, and it is a synthesized result of $v_1 \sim v_k$ handled by f.

Obviously, how to establish factor space F, how to get partial credibility v_k, and how to select a synthesis function f are the key jobs. Especially, f will influence the final result significantly. An weighted average algorithm and variations are frequently used as f. However, the relations $l_1 \sim l_k$ between factors are not bound to be linear, and the synthesis way of partial credibility to the final one should be determined by the computational process of the model but not a general expression.

There is a misunderstanding that factor space F is for providing a pathway to synthesize the final credibility, and it should be built through structural decomposition. This is not true. Model credibility is mainly affected by its computational process but not the structure, so the usage of the factor space in model V&V has to be corrected.

The main idea of this paper is to bring forward a practical method for obtaining the credibility of a simulation model, which may change the way of using factor space and similarity analysis etc. to get model credibility. A case study of a flying vehicle's mass center motion model is also presented to describe a walk through application of the method.

2. THE WORKFLOW OF THE METHOD

The fact is, the direct outputs of the simulation model have to be pointed out before acquiring the credibility, and if the credibility traceability is required, the intermediate outputs of the simulation model also have to be demonstrated. This actually forms the factor space of the model V&V.

By going through the factor space, each node of model output can be analyzed of the similarity between model and its origin by simulation and observed data, and then the result is transformed into the partial credibility.

If the credibility of direct outputs is acceptable, the work is done. If not, a defect tracing should be conducted to analyze lower level nodes of intermediate outputs. When all the analysis is done a final conclusion is drawn to show the grand credibility and nodes of defect.

The figure below shows the workflow of the method for obtaining the credibility of a simulation model, where sold line represents execution flow, and dot line represents data flow.

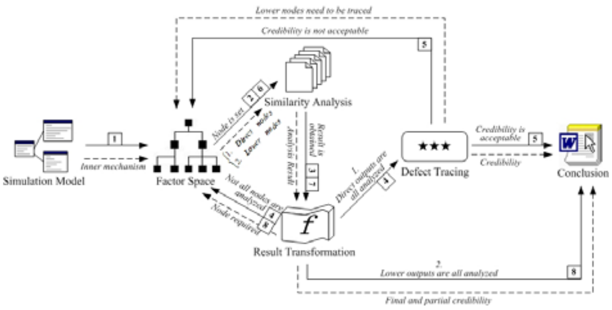


Figure 1: The Workflow for Obtaining Model Credibility

The workflow should be conducted as the following procedures:

1. Establish the factor space for evaluating the model credibility by reproducing the model computational process, to point out all factors which influence the credibility.
2. Use time domain analysis such as TIC (Theil Inequality Coefficient) (Andrei, Gary and Tryphon 2014), GRA (Gray Relational Analysis) (Wei and Li 1997) etc., frequency domain analysis such as Welch's periodogram analysis (Jiang and Mahadevan 2011), maximum entropy spectrum analysis (Mullins, Ling, Mahadevan and Sun 2015) etc., and statistics analysis such as parameter estimation, hypothesis testing etc. to perform the similarity analysis on higher nodes of direct outputs.
3. Use appropriate transformation formula to transform the similarity analysis result to credibility description.
4. If direct outputs are all analyzed, go to next step. If not, go back to Step 2 and continue with other direct outputs.
5. If the credibility is acceptable, draw the conclusion and the work is done. If not, perform defect tracing and go back to the factor space.
6. Similar to Step 2 but the nodes analyzed are lower nodes of intermediate outputs.
7. Similar to Step 3 but the nodes handled are lower nodes of intermediate outputs.
8. If traced lower nodes are all analyzed, gather partial and grand credibility to locate the defect nodes and draw the conclusion. If not, go back to Step 6 to continue with other nodes.

3. THE ESSENTIALS OF THE METHOD

3.1. Factor Space Establishment

The conventional way of building a factor space is the tree view, which is suitable to express the decomposition of a simulation model. However, a tree view is unable to reveal the computational process of the model, which is important to the grand credibility.

Actually, the tree view method only uses partial credibility on leaf nodes and applies weighted average algorithm etc. to synthesize the grand credibility. It does not need similarity analysis among branch nodes, which is problematic. The computing of credibility is not bound to be linear, and the direct outputs similarity should not be synthesized through intermediate outputs similarity. To resolve this, MADN (Multi-Attribute Decision Network) (Fang, Ma and Yang 2012) is used here.

See the example of a flying vehicle's motion model below. When the branches reach direct outputs, the factor space cannot be further built by structural decomposition but has to reproduce the model's computational process.

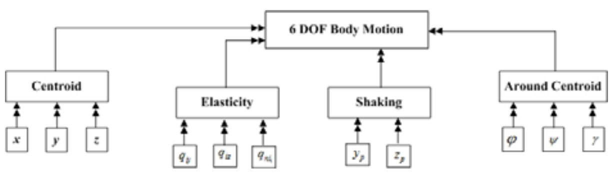


Figure 2: Part of A 6-DOF Body Motion Model's Factor Space

Take the vehicle's position x, y, z as an example. Further expand the factor space as below. The meaning of variables in Figure 3 is presented in Table 1 in the following Section 4.1. The process is conducted by revealing the computation process of the model but not the structural components. All the expanded nodes are credibility influencing factors n_k in Eq. (1), and the links between them are factors relationship l_k .

ISBN 978-88-97999-85-0; Affenzeller, Bruzzone, Jiménez, Longo and Piera Eds.

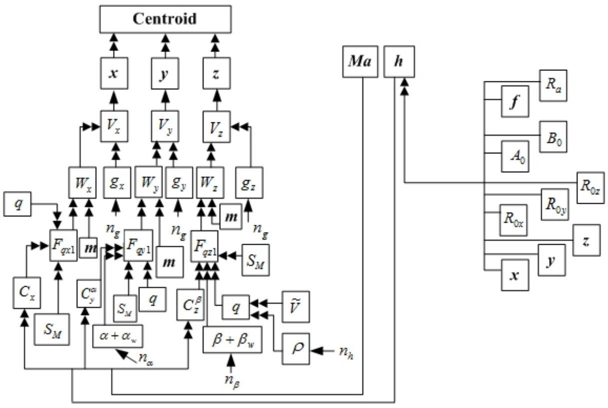


Figure 3: Expanded Factor Space of A 6-DOF Body Motion Model

A factor space establishment should be conducted as the following procedures:

1. Decompose the simulation model to the nodes which cannot be divided due to the model structure.
2. Continue to expand the factor space by computational process of the model. Position the outputs at upper level, and the inputs at lower level.
3. Terminate the expansion when the nodes which represent the model's initial inputs are revealed.
4. Use compose, derive, surpass, substitute, contradict, inherit etc. types of link (Fang, Yang and Zhang 2011) to connect the nodes in the factor space.
5. Determine the parameters for the factor space, such as weights for composition links, thresholds for surpassing links, conditions for uncertain links etc.

3.2. Similarity Analysis

A model's credibility has to be analyzed by comparing the simulation outputs with the real ones. In factor space, it has to be performed similarity analysis on the nodes, such as the vehicle's position x, y, z in Figure 2. According to the outputs feature, various analysis methods are correspondingly used.

If the output is time irrelevant, such as miss distance, hit probability etc., statistics analysis like *u* testing is often used. If the output is a series of sequential data which is gradually changing along with the elapsing time, time domain analysis like GRA is often used. If the output is changing frequently, frequency domain analysis like maximum entropy spectrum analysis is often used.

For example, the formula below shows the grey relational degree of a single output which comes from the simulation and the real world. (Wei and Li 1997)

$$\gamma = \frac{1}{n}\sum_{k=1}^{n}\frac{\min\limits_{k=1\sim n}|s_k - o_k| + \rho\max\limits_{k=1\sim n}|s_k - o_k|}{|s_k - o_k| + \rho\max\limits_{k=1\sim n}|s_k - o_k|} \quad (2)$$

where s_k is the simulation data series, and o_k is the observed data series. ρ is the resolution coefficient. $\min\limits_{k=1\sim n}|s_k - o_k|$ selects the minimum value in the series result set $|s_k - o_k|$, $k = 1 \sim n$, which is a single value, and vice versa for $\max\limits_{k=1\sim n}|s_k - o_k|$. The conclusion γ is named as grey relational degree, which indicates the similarity of the two series of data. The higher γ comes, the higher similarity is obtained.

Other methods have specific computational ways of acquiring the similarity between simulation and observed data respectively. Some have a Boolean result of "acceptable" or "unacceptable" like hypothesis testing and spectrum analysis. Some have a quantitative result like TIC. However, these similarity analysis results have to be transformed into credibility description.

3.3. Result Transformation

The result formation is various due to the similarity analysis method. It has different ranges, monotonies etc. In order to judge and synthesize the similarity analysis results, they have to be transformed into a unanimous credibility description, which has the range of [0, 1] and is monotonically increasing with the credibility.

The transformation should be realized by the nature of similarity analysis method itself. It has to adapt to the physical meaning of the method. For example, the T testing possesses a Boolean analysis result, which bases on an original and alternative hypothesis below:

$$H_0 : \mu_1 = \mu_2, H_2 : \mu_1 \neq \mu_2 \quad (3)$$

There is a β which represents the pseudo probability when the original hypothesis is accepted. Thus there comes a conclusion that when the alternative hypothesis is accepted, the transformed analysis result into credibility should be 0, and when the original hypothesis is accepted, the result should be 1-β.

If the significance level is α, according to the T distribution and the probability density function, 1-β can be calculated as below: (Zhang 2011)

$$1\text{-}\beta = 2 - T[T^{-1}(1-\frac{\alpha}{2}) - \frac{\delta}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}}] - T[T^{-1}(1-\frac{\alpha}{2}) + \frac{\delta}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}}]$$

$$S_w = \sqrt{\frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}} \quad (4)$$

where δ is the maximum tolerance of deviation, $\mu_1 = \mu_2 + \delta$; n and m are sample numbers of

Proceedings of the European Modeling and Simulation Symposium, 2017
ISBN 978-88-97999-85-0; Affenzeller, Bruzzone, Jiménez, Longo and Piera Eds.

simulation data and observed data; s_1 and s_2 are the variances of two samples.

Other methods have their own transformation formula, but meet the range of [0, 1] and the increased monotony. Some is simple, like 1-ρ for TIC. Some is relatively complex, like the transformation of F testing result for maximum entropy spectrum analysis (Zhang 2011).

3.4. Defect Tracing

If a node of higher level results in unacceptable credibility, we want to know what induces that "bad" credibility. In this case, a further and deeper validation is required, which should be conducted as the following procedures:

1. Perform similarity analysis on nodes in the factor space downwards, and locate those unacceptable nodes at upper level.
2. Perform further similarity analysis till the leaf nodes are reached.
3. On analysis route, those unacceptable nodes are defect points of the model, and the unacceptable leaf nodes are the origins of the defect.
4. If all sibling nodes of the upper unacceptable node are analyzed but no more defect points are found, that means more strict acceptability criteria (Oldrich and Andreas 2014) is required.

4. CASE STUDY

Use a flying vehicle's mass center motion model as an example, to present a walkthrough of the method. The model computes the position of the flying vehicle via aerodynamics coefficients and geophysics constants.

4.1. Factor Space Establishment

According to the computational process of the model, place final outputs at upper levels, intermediate outputs and inputs at lower levels, and build the factor space of model validation. Figure 3 shows the factor space. The table below shows the meaning of the variables:

Table 1: The Meaning of Variables

Variable	Meaning	Variable	Meaning
x,y,z	Position	ρ	Air density
V_x,V_y,V_z	Velocity	$\alpha+\alpha_w$	Velocity transition
W_x,W_y,W_z	Acceleration	$\beta+\beta_w$	Velocity transition
g_x,g_y,g_z	Acceleration of gravity	M_a	Mach
$F_{qx_1},F_{qy_1},F_{qz_1}$	Aerodynamic force	h	Height
m	Mass	f	Wind velocity
$C_x,C_y^{\alpha},C_z^{\beta}$	Aerodynamic coefficient	R_a	Equator radius
q	Dynamic pressure	A_0	Launch direction
$\widetilde{V}$	Relative flow velocity	B_0	Launch spot latitude
S_M	Sectional area	R_{0x},R_{0y} R_{0z}	Launch spot core radius

4.2. Similarity Analysis

First perform similarity analysis on higher nodes, that is, the flying vehicle's position x, y, z . The tables below show the simulated and observed data of the trajectory.

Table 2: Simulation & Observed Data of Position x

time(s)	simulation data x_sim(m)	observed data x_ref(m)
0.005	6022942.548	6022942.549
0.010	6023002.710	6023002.713
0.015	6023062.829	6023062.834
……	……	……
4.515	6064014.043	6069515.665
4.520	6064019.502	6069531.364
4.525	6064024.717	6069546.836
……	……	……
6.135	6055118.998	6058912.748
6.140	6055118.998	6058802.077
6.145	6055118.998	6058691.265

Table 3: Simulation & Observed Data of Position y

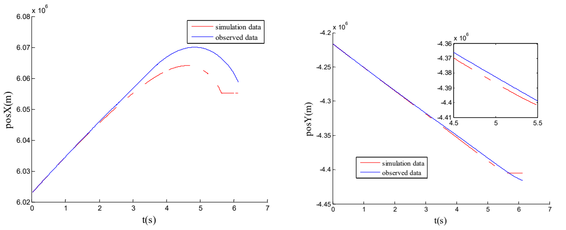
time(s)	simulation data y_sim(m)	observed data y_ref(m)
0.005	-4216050.798	-4216050.798
0.010	-4216224.172	-4216224.171
0.015	-4216397.545	-4216397.543
……	……	……
4.515	-4370194.423	-4366433.045
4.520	-4370366.670	-4366598.879
4.525	-4370538.900	-4366764.726
……	……	……
6.135	-4404860.781	-4415681.560
6.140	-4404860.781	-4415755.711
6.145	-4404860.781	-4415828.882

Table 4: Simulation & Observed Data of Position z

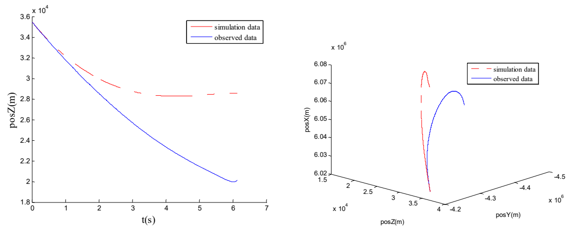
time(s)	simulation data z_sim(m)	observed data z_ref(m)
0.005	35471.429	35471.429
0.010	35452.188	35452.187
0.015	35432.944	35432.943
……	……	……
4.515	28272.912	22350.570
4.520	28273.351	22341.140
4.525	28273.801	22331.720

……	……	……
6.135	28528.336	20086.733
6.140	28528.336	20098.080
6.145	28528.336	20109.940

According to the simulation and observed data, draw the flying vehicle's trajectory as below:



(a) Position x (b) Position y



(c) Position z (d) 3D trajectory

Figure 4: Trajectory of the Flying Vehicle

Use GRA method of Eq.(2) to perform the similarity analysis of trajectory. Set $\rho = 0.5$, and we get $\gamma_x = 0.727,\ \gamma_y = 0.789,\ \gamma_z = 0.623$.

4.3. Result Transformation

When Eq. (2) meets $\gamma \geq \gamma_{th}$, it can be considered as the simulation and observed data have acceptable similarity. Set C_{th} as the acceptability criteria, the formula below can be used to transformed the GRA result into credibility:

$$C(\gamma) = \begin{cases} \frac{1-C_{th}}{1-\gamma_{th}}(\gamma - \gamma_{th}) + C_{th}, & \gamma \in [\gamma_{th}, 1] \\ \frac{C_{th}\gamma}{\gamma_{th}}, & \gamma \in [0, \gamma_{th}] \end{cases} \quad (5)$$

Select $\gamma_{th} = 0.8, C_{th} = 0.7$, the figure below shows the transformation:

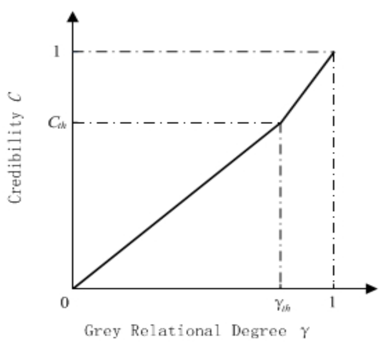


Figure 5: Transformation from γ to C

Use Eq.(5) to perform a transformation with the results we get in Chapter 4.2, the credibility of *x*, *y*, *z* can be derived as $C_x = 0.636,\ C_y = 0.691,\ C_z = 0.545$. It indicates the credibility is not acceptable, and the model validation needs a defect tracing.

4.4. Defect Tracing

Make further validation downwards along the factor space routes shown in Figure 3. Here to explain the process in general, we only analyze node *h* and its sub-networks due to the limited content of the paper.
Based on the principle of tracing downwards, the node of height *h* is analyzed first. Its simulation and observed data are shown as the following table.

Table 5: Simulation and Observed Data of Height *h*

time(s)	simulation data *h_sim(m)*	observed data *h_ref(m)*
0.005	10000.0778	10000.078
0.010	9999.489	9999.490
0.015	9998.888	9998.890
……	……	……
4.505	6340.700	8424.085
4.510	6324.899	8412.098
4.515	6309.027	8400.041
……	……	……
6.135	21.894	110.314
6.140	21.894	68.455
6.145	21.894	26.649

According to the simulation and observed data, draw the figure of height *h* as below:

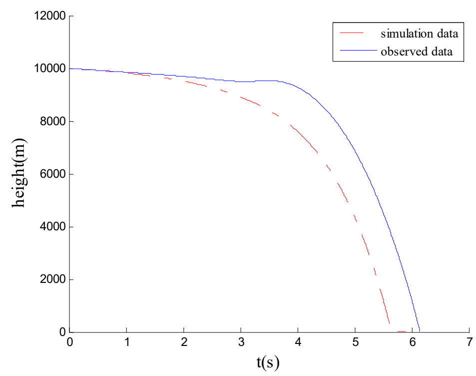


Figure 6: Height of the Flying Vehicle

Use GRA method of Eq.(2) to perform the similarity analysis of height *h*. Set $\rho = 0.5$, and we get $\gamma_h = 0.7161$. Use Eq.(5) to perform a result transformation, and we get $C_h = 0.6266$, which shows it is unacceptable, and needs further defect tracing.
The sibling nodes of height *h* include equator radius R_a, launch direction A_0, launch spot latitude B_0 and

launch spot core radius vector R_{0x}, R_{0y}, R_{0z}, which are equal to each other respectively between simulation and observed data. Meanwhile, the initial values of the iteration variable x, y, z are also equal to each other respectively between simulation and observed data. Then we focus on the last sibling node wind velocity *f*. The table below shows the simulation and observed data of *f*.

Table 6: Simulation and Observed Data of *f*

Time(s)	simulation data *f_sim*(m/s)	observed data *f_ref*(m/s)
0.005	69.230	-69.230
0.010	69.227	-69.227
0.015	69.223	-69.223
……	……	……
4.505	47.560	-59.899
4.510	47.467	-59.828
4.515	47.373	-59.757
……	……	……
5.515	16.608	-36.353
5.520	16.373	-36.186
5.525	16.137	-36.018

According to the simulation and observed data, draw the figure of wind velocity *f* as below:

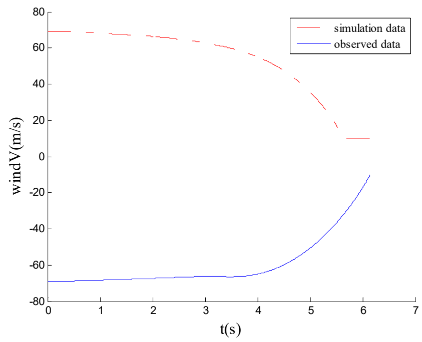


Figure 7: Wind Velocity of the Model

Use GRA method of Eq.(2) to perform the similarity analysis of wind velocity *f*. Set $\rho = 0.5$, and we get $\gamma_f = 0.5150$. Use Eq.(5) to perform a result transformation, and we get $C_f = 0.4506$, which shows the node of wind velocity *f* has unacceptable credibility.

4.5. Result Analysis

First, since x, y, z, the direct outputs of the model get unacceptable credibility $C_x = 0.6363 < 0.7$, $C_y = 0.6906 < 0.7$, and $C_z = 0.5447 < 0.7$, the model failed to pass the validation, which indicates the model is lack of credibility and has to be revised prior to its application.
Second, when analyze further along the factor space routes, the vehicle's height h gets unacceptable credibility $C_h = 0.6266 < 0.7$, and wind velocity f gets unacceptable credibility $C_f = 0.4506 < 0.7$ at leaf level. Other constants and variables are proved to be credible beneath the node of vehicle's height *h*. (The parts other than height *h* are omitted in the paper)
According to the procedures of defect tracing, it can be concluded that wind velocity f is the main reason which causes the direct outputs x, y, z has unacceptable credibility. In another word, as long as the sub-model of wind velocity is corrected, the model of mass center motion will have a good chance to be credible.

5. CONCLUSION

There is some misunderstanding in model validation nowadays, which hinders us to get an objective credibility of the model. Factor space has to be built by structural and computational decomposition combined, especially with the latter one on emphasis of obtaining objective results. Meanwhile, similarity analysis result has to be transformed into credibility description by the nature of similarity analysis method itself. Finally, the computational decomposition is mainly used to fulfill the potential defect tracing of the model, but not to make credibility synthesis.
Model validation is one of the key problems of simulation systems V&V. The method explained here provides a practical way to obtain the credibility of a simulation model, and corrects and clarifies some misunderstanding in model validation. Since the method is insensitive to the model which is being validated, it can be applied to all kinds of simulation models, no matter the model is built by continuous-time nature or discrete event. Further study should be conduced on how to determine an accurate acceptability criteria for the nodes in factor space according to the computational process of the model, and how to solve the problem of validate the model with iteration operation inside etc. Meanwhile, we should notice it is hard to get observed data on some intermediate outputs of the model, such as aerodynamic force etc. On these nodes subjective evaluation is often used to substitute the objective similarity analysis.

ACKNOWLEDGMENTS

The paper is supported by the National Natural Science Foundation of China (Item No. 61374164).

REFERENCES

Sargent R. G., 2015. Model verification and validation. London: Springer.
Roy C. J., Oberkampf W. L, 2011. A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. Computer Methods in Applied Mechanics & Engineering, 200(25–28):2131-2144.
You L., Mahadevan S., 2013. Quantitative model validation techniques: New insights. Reliability Engineering & System Safety, 111(2):217-231.
Eek M., Kharrazi S., Gavel H., et. al., 2015. Study of industrially applied methods for verification, validation and uncertainty quantification of

simulator models. International Journal of Modeling Simulation & Scientific Computing, 6(02): 1-29.
Kutluay E., Winner H., 2014. Validation of vehicle dynamics simulation models - A review. Vehicle System Dynamics, 52(2): 186-200.
Song H., Liu J., Schnieder E., 2017. Validation, verification and evaluation of a Train to Train Distance Measurement System by means of Colored Petri Nets. Reliability Engineering & System Safety, 164:10-23.
Zhou L, Sun K, Li H., 2016. Multifactorial decision making based on type-2 fuzzy sets and factor space approach. Journal of Intelligent & Fuzzy Systems, 30(4): 2257-2266.
Andrei D., Gary J. B., Tryphon T. G., 2014. Validating aircraft models in the gap metric. Journal of Aircraft, 51(6): 1665-1672
Wei H. L., Li Z., 1997. Grey relational analysis and its application to the validation of computer simulation models for missile system. System Engineering and Electronics, (2): 55-61
Jiang X. M., Mahadevan S., 2011. Wavelet spectrum analysis approach to model validation of dynamic systems. Mechanical Systems and Signal Processing, 25(2): 575-590
Mullins J., Ling Y., Mahadevan S., Sun L., 2015. A strachan separation of aleatory and epistemic uncertainty in probabilistic model validation. Reliability Engineering & System Safety, 147: 49-59
Fang K., Ma P., Yang M, 2012. The MAD network for virtual protocol systems credibility evaluation. CIMS, 18(5): 1054-1060
Fang K., Yang M., Zhang Z, 2011. The MAD network for credibility evaluation of computer simulation. Proceedings of The 13th IEEE Joint International Computer Science and Information Technology Conference, pp.636-642., Chongqing, China.
Zhang J. Y., 2011. Validation methods and assistant tools based on coherence of data. Thesis (Master). Harbin Institute of Technology.
Oldrich P., Andreas B., 2014. A new approach to define criteria for rail vehicle model validation. Vehicle System Dynamics, 52(1): 125-141

AUTHORS BIOGRAPHY

Ke Fang is an Associate Professor of the Control & Simulation Center, School of Astronautics at Harbin Institute of Technology, China. He holds a Ph.D. in control science and engineering from Harbin Institute of Technology, and was a visiting scholar at Arizona State University from year 2014 to 2015 in AZ, USA. His research interests include complex simulation systems, model validation and VV&A. His e-mail address is hitsim@163.com.

Ming Yang is a Professor of the Control & Simulation Center, School of Astronautics at Harbin Institute of Technology, China. He holds a Ph.D. in control science and engineering from Harbin Institute of Technology. He involved in major simulation conference and committees in China. His research interests include distributed simulation, hardware-in-loop simulation and VV&A etc. His e-mail address is myang@hit.edu.cn.

Kaibin Zhao received the M.S. degree in control science and engineering from Harbin Institute of Technology, China, in 2013. He is currently pursuing the Ph.D. degree in control science and engineering at HIT. His current research interests is the model validation for complex system. His e-mail address is kaibin.zhao.HIT@hotmail.com.