A Practitioner's View of Verification and Validation

Jayne T. Lyons, The MITRE Corporation
Craig T. Doescher, The MITRE Corporation
Rosemary T. Seykowski, Potomac Institute for Policy Studies
Keith P. Curtis, The MITRE Corporation

The MITRE Corporation
1820 Dolley Madison Boulevard
McLean, VA 22102-3481
(703) 883-6000 (Phone); (703) 883-1279 (Fax)

Potomac Institute for Policy Studies
1600 Wilson Boulevard, Suite 1200
Arlington, VA 22209
(703) 525-0770 (Phone); (703) 525-0229 (Fax)

jtlyons@mitre.org, cdoesche@mitre.org, rseykowski@potomacinstitute.org, kcurtis@mitre.org

67th MORS Symposium
Working/Composite Group Number 25 - Test and Evaluation
19 November 1999

ABSTRACT: *The Navy's Program Executive Office for Theater Surface Combatants (PEO TSC) has embarked on a program to use advances in simulation technology to significantly improve system acquisition through the use of validated federated simulations. As a result PEO(TSC) has developed an engineering-level High Level Architecture federation focused on Integrated Ship Defense (ISD). The ISD federation development approach uses a model-test-model paradigm to evaluate the federation's potential to support Test and Evaluation (T&E). While this project is not complete, this paper discusses preliminary lessons learned from the federation verification and validation (V&V) effort, and comments on the complex, evolving inter-relationship between simulation, V&V, and T&E.*

The ISD federation is intended to closely represent the actual anti-air warfare combat system installed on Navy ships and includes both tactical code-in-the-loop and detailed physics-based simulations. The ISD federation simulates system-level effectiveness, specifically the measures of effectiveness and measures of performance associated with the development and operational tests for the Ship Self Defense System Mark I. Using the data collected from these tests and comparing results to simulated test scenarios, we employ a model-test-model approach to conduct V&V on the federation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
25-11-99

2. REPORT TYPE
67th MORS Symposium Paper

3. DATES COVERED *(From - To)*
22-24 June 1999

4. TITLE AND SUBTITLE
"A Practitioner's View of Verification and Validation"

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER
0799V23A

5e. TASK NUMBER
AA

5f. WORK UNIT NUMBER

6. AUTHOR(S)

Jayne T. Lyons, The MITRE Corporation
Craig T. Doescher, The MITRE Corporation
Rosemary T. Seykowski, Potomac Institute for Policy Studies
Keith P. Curtis, The MITRE Corporation

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The MITRE Corporation
1820 Dolley Madison Blvd
McLean, VA 22102-3481

Potomac Institute for Policy Studies
1600 Wilson Blvd, Suite 1200
Arlington, VA 22209

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Navy's Program Executive Office for Theater Surface Combatants (PEO TSC) has embarked on a program to use advances in simulation technology to significantly improve system acquisition through the use of validated federated simulations. As a result PEO(TSC) has developed an engineering-level High Level Architecture federation focused on Integrated Ship Defense (ISD). The ISD federation development approach uses a model-test-model paradigm to evaluate the federation's potential to support Test and Evaluation (T&E). While this project is not complete, this paper discusses preliminary lessons learned from the federation verification and validation (V&V) effort, and comments on the complex, evolving inter-relationship between simulation, V&V, and T&E.
The ISD federation closely represents the actual anti-air warfare combat system installed on Navy ships and includes both tactical code-in-the-loop and detailed physics-based simulations. The ISD federation simulates system-level effectiveness, specifically the measures of effectiveness and measures of performance associated with the development and operational tests for the Ship Self Defense System Mark I. Using the data collected from these tests and comparing results to simulated test scenarios, we employ a model-test-model approach to conduct V&V on the federation. Thus, the tools used to analyze live test data will also be used to conduct V&V of the federation.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Craig Doescher
U	U	U	Unlimited	9	19b. TELEPHONE NUMBER *(include area code)* (703)883-6742

1. Introduction

The use of simulation to support system acquisition is not new to the Program Executive Office for Theater Surface Combtants (PEO TSC). An established "sneaker net" of models and simulations has been used for some time to estimate the performance of the Integrated Ship Defense (ISD) system. What is new is establishing a closer relationship between the operational view, the system under development and test, and the simulation representing the system. This relationship is made possible through the use of the High Level Architecture (HLA) and model-test-model paradigm. An overview of this new and developing relationship is shown below in Figure 1.

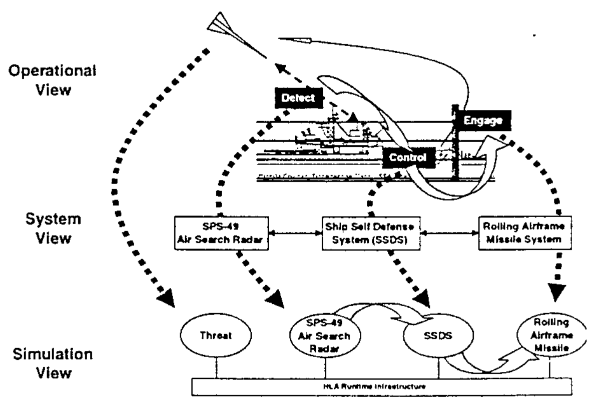

Figure 1. Ship Defense Views

The desire to take greater advantage of simulation for system acquisition, the "newness" of HLA, and the progressive demand for explicit simulation Verification and Validation (V&V) throughout the Department of Defense (DoD) community provided the mandate for the V&V effort in the ISD federation development. Although V&V has been part of the federation development process from the start, the importance of V&V has risen as the Navy looks toward supporting the Test and Evaluation (T&E) community with valid simulations of the system under test. As the simulations representing the system components become more robust, they will become increasingly valuable to the system development and testing process. Through reconfigurable and extensible simulation, system capability can be evaluated in operational contexts not possible during live testing.

Since 1997, PEO (TSC) has led an effort to explore the use of distributed simulation as a tool to improve their acquisition process, specifically in the area of T&E. As part of this exploration, the V&V process must be sufficiently understood so as to properly support the T&E community with appropriate simulation tools. The ISD project makes use of HLA as the underlying architecture for connecting the simulations, known as federates, together into a federation. This paper presents our planning process, accomplishments, and the lessons learned thus far in our efforts to develop a practical V&V approach.

2. The Problem

Before a simulation federation becomes an accredited resource for the program manager, it must be verified and validated. The focus of our effort is to determine an executable V&V process that is cost-effective, addressing limited schedule and constrained resources, things of great importance to the program manager. This process must be able to generate, collect, organize, analyze, and report to the stakeholders those data and findings pertinent to establishing an appropriate degree of confidence in the results of the ISD federation studies.

The simulation V&V strategies identified, the rationale for the V&V exercises described, and the set of V&V activities specified will contribute to an audit trail by which the representativeness of the federation and the credibility of its predictions may be demonstrated. Most importantly, we want to provide feedback to the simulation community on which other federations may base their V&V programs.

2.1 Stakeholders

There are many agencies within the Navy's Integrated Ship Defense community acting as stakeholders. The primary stakeholder is PEO (TSC) in that the success of this project from a functional standpoint can lead to great advances in how systems are tested and subsequently acquired, leading to reduced costs. In the future, we look to the T&E community supporting PEO (TSC) as the potential end user of a tool such as the ISD federation.

Other stakeholders include those developing systems in support of the ISD mission such as Naval Surface Weapons Center, Dahlgren Division (NSWC-DD), Naval Research Laboratory (NRL), Johns Hopkins University/ Applied Physics Laboratory (JHU/APL), Naval Air Warfare Center Weapons Division, China Lake and Point Mugu, CA (NAWC-WPNS), and Naval Warfare Assessment Station, Corona, CA (NWAS). As system developers, these organizations would be greatly impacted by a federation such as the ISD federation for it potentially provides a new method for evaluating system performance. This is only possible if the federation results can be validated. Organizations, such as Litton/PRC, Trident Systems, and The MITRE Corporation, have participated in either the development of federate software or integration into a federation and have a stake in the federation's success and therefore the V&V process.

The final stakeholder is the distributed simulation community at large. It is our expectation that what we learn can help others who are attempting to build federations to support the acquisition of military systems. If distributed simulation can prove to be successful in acquiring cheaper and better systems, it will be because the resulting simulation is valid and credible. It will be because we have succeeded in verifying and validating the resulting federation so that it is a credible representation of the ISD system.

2.2 Scope

The ISD mission is to provide a coordinated self-defense system for ships that traditionally have not had organic defenses to protect against low-flying cruise missiles in open ocean and littoral environments. The Navy community refers to the self-defense timeline as a "detect-control-engage" sequence. The typical ISD system centers around a combat system that "controls" a suite of sensors to "detect" incoming threats and weapons to "engage" threats. Because the more stressing operational requirements ISD system must satisfy are not usually tested during live testing, PEO(TSC) intends to use simulation to explore areas not previously possible. The development of the capability is planned to be executed in three phases with functionality delivered over time. This paper focuses on Phase I of the ISD federation which includes the critical components of the ISD suite of systems completing the cycle of detect, control, and engage in a test environment. HLA is used to facilitate interoperability between the following component simulations: Ship Self Defense System (SSDS) Mk I, SPS-49 air search radar, Rolling Airframe Missile (RAM) Blk 0, Close In Weapon Support (CIWS), and SLQ-32 Electronic Support (ES) System. The federation also includes both aerial test targets as well as reactive real world threats to explore scenarios not possible in live tests. The federation is designed to calculate the measures of effectiveness (MOEs) and measures of performance (MOPs) used by the test community during acceptance testing. This type of analysis will allow us to compare simulation results to test results which is the basis for our V&V process.

Like most programs, this V&V effort was bounded by funding shortfalls and integration issues, therefore a subset of the Phase I federation was used as a V&V testbed to explore the V&V issues. We concentrated on the detection portion of the ISD functional timeline by including SSDS, SPS-49 search radar, and aerial test target federates in the federation. Although not optimal, we believe the federation will still provide us with lessons learned in formulating a practical V&V process. Issues concerning the comparison of test data to simulated data when all simulation components are not present will arise and we will examine their impact as part of this effort.

3. Goals of V&V Effort

This V&V effort had two goals. The first was to demonstrate a feasible and practical V&V process on a federation of simulations using a subset of the ISD federation. Most of the current V&V efforts throughout DoD focused on the V&V process required for a simulation that stands alone. Technical and programmatic issues in bringing together multiple simulations to meet a single objective. Our effort seeks to understand what practical requirements are needed to V&V a federation of simulations rather than a stand-alone simulation.

The second goal backs the stakeholder, PEO (TSC) in this case, whose objective is to make better acquisition decisions and support the T&E community responsible for acceptance testing for new combat systems. The long-term vision is to provide PEO (TSC) decision-makers with a level of detail on combat system effectiveness that exceeds their current capability and to provide a tool to identify potential risk areas prior to testing the system at sea. The purpose of this effort then is to understand what level of validity a federation must have to support the T&E community.

4. Plan Development

Our first step was to understand the types of data we had available to us to conduct verification and validation. Considerable data had been collected during the SSDS Mark I Operational Tests (OT) conducted in June 1997. We structured the federation to include key test components, emulate the test scenarios, and calculate the critical MOEs and MOPs to the greatest extent possible. We refer to this approach as model-test-model paradigm to V&V. Essentially, we intend to use the test scenarios to aid in our federation construction, compare simulation results with test data, and then refine the federation as a result of comparisons.

Our next step in the V&V process development was to survey the literature of the simulation community for documented V&V processes and techniques to gain insight into the successes and failures of others. From this survey of V&V plans and efforts, it was determined that many plans to conduct V&V have been written, but few plans have been thoroughly executed. Most plans appeared incomplete and unexecutable due to lack of M&S requirements and acceptance thresholds, and lack of specific techniques to validate the analysis process.

When developing our process we started with the fundamentals and took the process outlined by Law and Kelton (1991) and mapped it to the Secretary of the Navy directive, SECNAV Instruction 5200 (1999). This mapping is shown in Figure 2. Law and Kelton's process is a non-specific overarching approach that is widely accepted. The SECNAV Instruction adds more detail, not, however, to the point of a "how to" guide. Our effort focuses on the "Federation" and "Correct Results Available" blocks of Law and Kelton and the last two blocks of the SECNAV Instruction, "System Verification" and "Results Validation". It is in these areas that we are trying to add insights and a "how to" methodology for a federation.

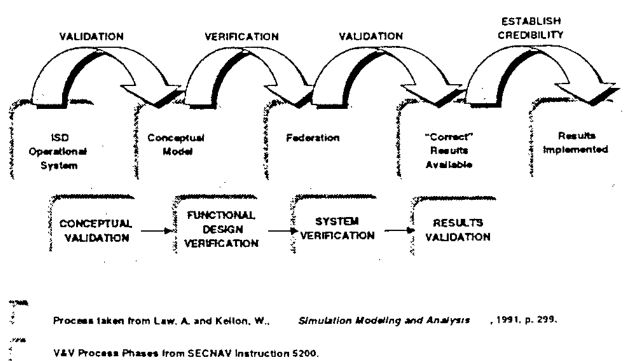

Figure 2. V&V Process

Our process initially supported Principle 6 of the DoD VV&A Recommended Practices Guide (1996) which states "V&V of each submodel or federate does not imply overall simulation or federation credibility and vice versa." However, throughout the evolution of our effort, we have determined that while having knowledge of the federate V&V status is useful in the federate selection process, it is not necessary that federates be verified and validated prior to integration into the federation. We understand that federates need to be verified and validated in the context of the federation. Our V&V process supports this hypothesis.

5. The Process

"Verification" and "validation" is traditionally defined as follows: "verification" is assessing whether we have built the model correctly; "validation" is assessing whether we have built the correct model (Law & Kelton, 1991). Our process separates the verification steps from the validation steps and treats them as separate processes. This separation compliments the use of different scenarios to conduct the testing. For the verification testing, very simple scenarios were constructed that are not necessarily based on real test data. They include parameters such as simple aerial target trajectories and stationary ship platforms. For the validation steps, the test scenarios emulate the operational tests as closely as possible to facilitate comparison between the simulation results and the actual test results. The overall process is diagrammed in Figure 3.

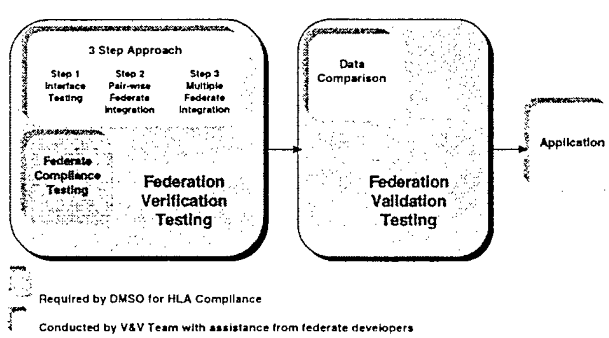

Figure 3. ISD V&V

The verification testing was further divided into three steps that mirror what we have done in successful federation integration testing. In the first step, single federate operation is verified by stimulating the federation with "stub" or "sister" federates. The goal is to ensure that message data in the form of parameters of interactions and attribute updates contain reasonable values. We assess the "reasonableness" through Subject Matter Experts (SMEs). SMEs also check that the timeline of events is correct. Lastly, using pre-computed values, we verify that each federate is making coordinate transformations correctly.

Step 2 of federation verification evaluates pair-wise interactions between federates using data collection tools. During this step we check that the parameters of interactions between federates comply with what is stipulated in the Federation Object Model (FOM) which is based on system interface documentation. We verify that messages "make sense" using SMEs and performance data for individual systems. We also verify coordinate system translations and time management parameters where necessary.

The third step exercises all federates repeating the tests in step 2.

It is noted in the diagram that federate compliance testing, required by the Defense Modeling and Simulation Office (DMSO) for HLA compliance, is done in this phase. It can be conducted either individually by federate developers before federation integration or it can be conducted in the context of the resulting federation. Our plan is to conduct compliance testing on individual federates after they have been integrated into the federation because the tight coupling of the federate messages does not allow federates to be run separately.

Validation is the determination that the resulting federation accurately represents the real system. It will answer the questions, "have we constructed the correct federation to support Navy T&E and can we emulate a test?" If there is an existing system, call it the base system, an ideal way to validate the model is to compare its output to that of the base system (Banks, 1998). In our case, the SSDS MK I was thoroughly tested and data were collected and reported by NWAS.

Using this test data, we selected eight scenarios that vary in complexity and components. We will run the simulation multiple times collecting combat system message data and federate communication data for each scenario. This simulation data will then be compared to test data captured in a series of reports produced by the Navy. Specifically, we will examine the message traffic between the SSDS and SPS-49 radar (refer to Figure 1). We will verify message content, ordering, and time stamps. Dependence on SMEs to interpret data is necessary.

As mentioned earlier, one of the challenges will be to determine what effect the differences in the base system and the federation have and to validly draw conclusions from each systems data.

6. Tools

The tools used for our V&V process are based on the tools currently used by the T&E analysis facility in the at-sea OT tests of the combat system. This is a unique feature of our federation and one that allows us to readily compare simulated data to real test data.

Several tools will be used in our process. A configurable data extraction tool, called DX, in the combat system software was retained as a capability in the SSDS federate and will be used to collect SSDS message traffic while the simulation is running. A test data analysis system, called "the Tool", enables the analyst to evaluate combat system weapons message traffic and manipulate data into the required measures of effectiveness and performance. A passive data logger, developed specifically for this project, will also be used to collect any data exchanged between federates during run time. This tool can also be used to verify aerial target flight paths and event timelines. Tools such as MATLAB® will allow us to plot and do statistical analysis on resulting data.

7. Progress to Date

Currently, we have completed portions of the verification testing. Of the three steps that we describe, we completed the first two verification steps, individual federate interfaces and pair-wise federate testing. For both steps, we relied on our data logger to verify that each federate conforms to the FOM. In the pair-wise testing, we verified the message types as well as the ordering of messages exchanged to insure the time evolution in the simulations were correct. We also verified that the coordinate system translation algorithms used by each federate was done consistently and correctly. Lastly, we performed compliance testing as we integrated federates into the federation.

8. Lessons Learned to Date

Because the federation is incomplete at present and because we have yet to apply the process fully, the lessons that we have learned are most important to the simulation community as a whole. Firstly, the V&V community would be concerned about the validity of this paper if the first two lessons were not mentioned.

We have re-discovered that V&V cultures (terminology, methods, etc.) differ across the simulation community and there is a need to account for differences with education and documentation at the onset of a considerable V&V effort. We did not do this initially until we as a team were at a standstill and could not go forward. We recommend this be done at the beginning of the federation development process.

One cannot start conducting V&V early enough in the federation development process. V&V should be of primary interest when making federate selections, however, the V&V status of federates is not however a pre-requisite for inclusion in a federation.

Operational or tactical software, used as federates, introduces an interesting V&V debate on the need and level of V&V required. Some view the need for V&V to be negligible because the simulation is the operational software. Based on our experience, significant V&V issues arise when the operational software is changed to enable integration into the federation. These changes require the federate to be validated.

Legacy simulations have V&V histories that are often based on their accepted usage over time rather than rigorous methods and are rarely documented sufficiently. One could spend inordinate amounts of time piecing together documentation and determining their status. We recommend that one not take time sorting it out, rather the time would be better spent understanding the capabilities and limitations of the federate.

V&V of individual federates is not sufficient; the federation needs to be verified and validated as a single system. A federate may have a V&V history but unless it was used in a federation for the same purpose, it needs to be re-verified and validated in the context of the federation. We spent too much effort requiring federate developers to V&V individual federates and provide supporting documentation. We realize that this was wasteful and time would have been better spent conducting the V&V on individual federates while running in the federation where inputs, environments and assumptions are different from running standalone.

We neglected early in the federation development process to document our conceptual model sufficiently (refer to Figure 2). At the time, we had not realized how important this would be to the V&V process. The conceptual model dictates what is being built and for what reason. For this reason, the document supports the V&V process. We recommend that federation developers carefully construct their conceptual model with the V&V process in mind. We also recommend that the conceptual model itself be validated.

When designing a V&V process, it is imperative to understand what types of data are needed to conduct V&V and to influence the data collection process as much as is possible. In our case we had considerable amounts and types of data coming from the SSDS MK I operational testing at our disposal for V&V. We then designed our V&V approach around exploiting that data. Problems arise, however, when data update rates are not sufficient or important data from a simulation perspective is not part of the test collection. As a result, we depended on system experts rather than quantitative data to validate the simulation's performance.

Similarly, thought should be given to data formats, coordinate system translations, and how "ground truth" is defined. There is considerable data manipulation that is done in test data reduction. Understanding these processes is imperative when using test data for validation purposes.

9. Conclusions

We have employed a model-test-model approach to our V&V process for the ISD federation. Using data collected during system operational testing and following guidance from literature we formulated a plan that allows us to compare simulation results to live test data. We realize that our V&V process is predicated on the existence of test data and additional efforts are required to understand how to use simulation to support acquisition decisions prior to the testing phases.

While we are learning about the V&V requirements for Navy test and evaluation of systems, we are basing our processes on the fact that we have data to exploit. Issues arise when data does not yet exist. This will need to be addressed by the Program Executive Office for Theater Surface Combatants and Expeditionary Warfare as they face complex combat system development for new surface ships. They are planning to construct a federation to support the acquisition of the new SSDS Mk. 2. Their focus will be on estimating the performance of the SSDS system, and in particular, the platform-level operational requirement for ship self defense, Probability of Raid Annihilation (PRA) which is exceedingly difficult and costly to analyze based on at-sea and land-based testing only. Understanding the V&V requirements for federations built around new systems is imperative to building a useful tool for the decision-maker.

10. References

[1] Averill Law, W. David Kelton, 1991 *Simulation Modeling and Analysis*, McGraw-Hill, New York.

[2] Secretary of the Navy, 1999, *SECNAV Instruction 5200*, Washington, DC.

[3] Jerry Banks, 1998, *Handbook of Simulation*, John Wiley & Sons, Inc., New York.

[4] DoD Instruction 5000.61, 1996, *DoD M&S VV&A*, Washington, DC.

[5] Defense Modeling and Simulation Office, 1996, *DoD VV&A Recommended Practices Guide*, Alexandria, VA.

Acronyms

CIWS	Close In Weapons System
DMSO	Defense Modeling and Simulation Office
DoD	Department of Defense
ES	Electronic Support
FOM	Federation Object Model
HLA	High Level Architecture
ISD	Integrated Ship Defense
JHU/APL	Johns Hopkins University Applied Physic Laboratory
MOE	Measure of Effectiveness
MOP	Measure of Performance
NAWC	Naval Air Warfare Center
NRL	Naval Research Laboratory
NSWC-DD	Naval Surface Weapons Center - Dahlgren Division
NWAS	Naval Warfare Assessment Station
OT	Operational Test
PEO(TSC)	Program Executive Office for Theater Surface Combatants
PRA	Probability of Raid Annihilation
RAM	Rolling Airframe Missile
SME	Subject Matter Expert
SSDS	Ship Self Defense System
T&E	Test and Evaluation
V&V	Verification and Validation

Author Biographies

JAYNE LYONS, an electrical engineer at the MITRE Corporation, was tasked with the technical leadership of the integration team for the ISD federation. She was a member of the ISD Verification and Validation Team.

CRAIG DOESCHER, an operations research analyst at the MITRE Corporation, was a member of the ISD Verification and Validation Team.

ROSEMARY SEYKOWSKI, formerly an operations research analyst at The MITRE Corporation and currently a Research Associate at the Potomac Institute for Policy Studies, was the lead member of the ISD Verification and Validation Team.

KEITH CURTIS, an operations research analyst for The MITRE Corporation, was the ISD program manager.

List of Descriptors

Verification and Validation
Engineering-Level Simulation
Federation Development Process
Navy Test and Evaluation
Model-Test-Model
System Acquisition