Appendix 2

Best Practices

in the Development of Simulation Scenarios

for Validation Activities in Fast and Real-Time Simulation

Prepared By:

Acknowledgments

The persons listed below attended the March 2003 workshop in Rome, Italy. This paper is a collection of their ideas, experience, and expertise. The authors would like to thank the attendees for their dedicated involvement and contributions to this effort.

AENA, Spain:, José Miguel de Pablo, Manuel Dorado
BAE Systems, Washington: Donald Eddy
CENA, France: Caroline Chabrol, Yann LeFablec, Vincent Kapp, Didier Pavet
CNA, Washington, D.C: Joseph Post
CSSI, Washington, D.C.: Stephane Mondoloni
Deep Blue, Italy: Marinella Leone, Patrizia Marti, Alberto Pasquini
DFS, Germany: Andreas Tautz
DGAC, France: Bernard Bonnafoux
DLR, Germany: Franz Knabe
Eurocontrol Experimental Centre, France: Peter Crick, Andrew Harvey, Yann Kermarquer, Nigel Makins, Conor Mullan
Eurocontrol HQ, Belgium: Ulrich Borkenhagen, Keith Vickery, Hans Wagemans
ENAV, Italy: Giancarlo Ferrara. Christiana Abbate, Giorgio Matrella, Francesco Podiari
ERAU,Florida: Ian Wilson
FAA, Washington: Dino Piccione, Diana Liang, Sharon Kurywchak
FAA Tech Center: Karen Buondonno, Parimal Kopardekar, Albert Schwartz
Fraport, Germany: Oliver Dellman
ISA Software, France: Ian Crook
LMI, Virginia: Shahab Hasan
MITRE, Virginia: Fred Wieland, Urmila Hiremath
NASA: Sandy Lozito, Karlin Roth, Gary Lohr
NATS UK: Joanne Thornton
NLR, Netherlands: Juergen Teutsch
OECD- HRP, Norway: Gyrd Skraaning
SICTA, Italy: Roberto Caramanica, Giovanni Esposito, Rossana Loffredo
SJSU, California: Kevin Corker
TransSolutions, Texas: Belinda Hargrove
Univ. of Sienna, Italy: Antonio Rizzo
Volpe, Massachusetts: Kim Cardosi

Table of Contents

Page

List of Tables

List of Figures

Executive Summary

The second Federal Aviation Administration (FAA)/EUROCONTROL Action Plan 5 Workshop, held in conjunction with Action Plan 9, was conducted on March 11-13, 2003, to develop and standardize 'best practices' related to the development of simulation scenarios for validation exercises. The workshop was well attended by 32 European and 19 United States (US) practitioners, all of whom were experienced in fast-time and real-time simulation for concept validation. The workshop focused on the development of simulation scenarios for both fast-time and real-time simulation and covered the following topic areas:

1. Coordination of Fast & Real-time Simulation Activities
2. Scenario Uses
3. Scenario Considerations
4. Building of Traffic (or Traffic Building)
5. Scripting of Specific Events (or Specific Event Scripting)
6. Analysis Considerations

Practitioners introduced each of the topic areas with a short briefing, which was then followed by a discussion session facilitated by a moderator. After the initial discussions on each topic, practitioners were divided into working groups to further discuss and identify elements of best practices for the topics to which they were assigned. Their consolidated recommendations were presented to all workshop participants for final discussion and consensus.

The resultant best practices for developing and validating simulation scenarios are presented in this appendix. Their purpose is to serve as supplemental guidelines for experienced practitioners who perform concept validation activities. Additional material concerning scenario development and coordination of fast and real-time activities outside the scope of the topic areas listed above is also included in this appendix, based on discussions held during the workshop.

Overall, the participating practitioners felt that the workshop was very valuable and useful for their research endeavors. They suggested that further workshops be arranged to discuss additional topics such as Metrics and Measures, more specific topics on Scenario Development, and Reporting Results.

1 Background

The Federal Aviation Administration (FAA)/EUROCONTROL R&D Committee was established in December 1995 during the second FAA/EUROCONTROL R&D Symposium, held in Denver, Colorado. The focus of the FAA/EUROCONTROL R&D Committee was to define priorities in terms of common actions and agendas of both organizations. The

Committee identified areas of mutual interest where the FAA and EUROCONTROL could work together in R&D and defined several R&D Cooperative Tasks, which are referred to as 'Action Plans'.

The goal of Action Plan 5 (AP5) is to determine a unified strategy for validating and verifying the performance, reliability, and safety of Air Traffic Management (ATM) systems. One objective of Action Plan 9 (AP9) is to promote a mutual understanding between US and Europe on the use and development of fast-time simulation models for modelling of air traffic operational concepts. An objective shared by AP5 and AP9 is to develop detailed best practices for performing tasks associated with the verification and validation of ATM systems. The development and subsequent use of these best practices will allow for sharing of information and comparison of results among all researchers. Best practices covering the use of metrics, data collection, data analysis, and reporting are useful for each type of validation exercise, as they take into account several factors including resources, cost, and most importantly, prior lessons learned.

In order to capture 'lessons learned' from a range of research organizations, several workshops were conducted and more will follow. The attendees of these workshops include experienced practitioners in their respective fields and in the topics presented at the workshops. This appendix describes the best practices resulting from the Second US/Europe Practitioners' Workshop, where the topic for discussion was the development of fast- and real-time simulation scenarios for validation activities.

2 Second US/Europe Practitioners' Workshop on best practices in the development of fast-time and real-time simulation scenarios for validation activities.

The FAA and EUROCONTROL organized the second AP5 workshop in conjunction with AP9. The workshop took place in March 2003 in Rome, Italy. The objective of this second workshop was to discuss and identify best practices from practitioner experience in the development of simulation scenarios for validation activities.

During ATM concept validation, both fast-time and real-time simulation techniques are used. These techniques require the development of scenarios, but at present, no standardized guidelines exist to steer the researcher through the development of these scenarios. During this workshop, researchers identified best practices based on actual experiences in developing scenarios that are either unique to fast-and real-time techniques or common among them. This appendix documents those best practices. In addition, relevant best practices from the Operational Concept Validation Strategy Document (OCVSD) Appendix 1 ("Best Practices for Human-in-the-Loop Validation Exercises", here after referred as Appendix 1, are also included where appropriate.

The Second Workshop participants included 19 practitioners from the United States (US) and 32 practitioners from Europe representing both fast-time and real-time simulation fields. The US participants included researchers from the FAA William J. Hughes Technical Center, National Aeronautics and Space Administration (NASA) Ames Research Center, NASA Langley Research Center, Volpe Transportation Research Center, MITRE, BAE systems, Embry-Riddle Aeronautical University (ERAU), Logistics Management Institute (LMI), CNA Corporation, CSSI, San Jose State University, and TransSolutions.

The Europe participants included researchers from Aeropuertos Españoles y Navegación Aérea (AENA), Deutsche Flugsicherung GmbH (DFS), Direction Générale de l'Aviation Civile (D.G.A.C.), Sistemi Innovativi per il Controllo del Traffico Aereo (SICTA), Deutschen Zentrum für Luft- und Raumfahrt (DLR), EUROCONTROL, EUROCONTROL Experimental Centre, Frankfurt Airport Services Worldwide (Fraport), Deep Blue, Ente Nazionale di Assistenza al Volo (ENAV), Centre d'Etudes de la Navigation Aérienne (CENA), ISA Software, National Air Traffic Services (NATS), OECD Halden Reactor Project, and Nationaal Lucht en Ruimtevaartlaboratorium (NLR).

3 Relevant Terms & Definitions

3.1 Simulation Definitions

Because fast-time and real-time simulations are distinct approaches, each technique has its own terminology, data, and validation strategies. To clarify some of the differences, this section contains a definition of terms.

3.1.1 Real-Time Simulation

A real-time simulation is one in which a human operator (e.g., air traffic controller) interacts with and reacts to, simulated conditions in near real-time.

3.1.2 Fast-Time Simulation

A fast-time simulation is one in which there is no interactive human involvement in simulated conditions. Instead, scenarios unfold using rule-based decisions that control the interactions between simulated actors.

<u>Note</u>: There are hybrid simulation techniques, which allow human interaction, but do not have to run in real-time. For purposes of this appendix, these techniques are classified under fast-time simulation.

3.2 Scenario Definitions

Scenarios are well recognized as an important tool for assessing the impact of proposed changes on the ATM system, and in the case of human-in-the-loop studies, on human performance. There are different views on the definition of scenario, depending on the context in which the term is being used. The following paragraphs define the term scenario within various contexts. In ATM, the two most common definitions refer to Operational Concept Scenarios and Validation Scenarios.

3.2.1 Operational Concept Scenario

An operational concept scenario is a documented description of a sequence of events involving one or more ‘actors’ and focused on some specific ATM function or procedure. This type of scenario is generally implemented during initial concept design and development phases. It allows for the identification and refinement of issues for further testing and development.

3.2.2 Validation Scenario

A validation scenario is an extension of the operational scenario. It is applied in a simulation environment where the objective is to excite the performance and interactions described in the operational scenarios. The simulation environment refers to various configurations of airspace, traffic sample, weather, failure modes, and any other controllable variables that might affect the performance of the ATM system. In this way, a validation scenario will test the assumptions in the concept scenarios and thus the concept design. The validation scenario provides robust feedback in terms of whether or not a proposed concept can be implemented in the operational environment.

3.3 Scenarios in Simulations

There are several types of scenarios; each takes on a different meaning. The following definitions distinguish the various scenario terms.

3.3.1 Real-Time Simulation Scenario

A scenario for a real-time simulation is generally characterized by a specific or generic geographical environment, traffic sample, and special scripted events, such as weather. A variety of scenarios may be developed for a real-time activity in order to gain generality in the results. For example, traffic samples may vary by volume to test the limits of the human operators with new concepts, procedures, and/or equipment. In general, scenarios in real-time simulation are meant to challenge the human operator in some or all aspects of the system. They are used to allow for measurement of workload, for example, of a new concept on an air traffic controller.

3.3.2 Fast-Time Simulation Scenario

A scenario for a fast-time simulation is characterized by a given set of conditions that include a specific geographic environment (e.g., airspace, airport), traffic sample, platform configuration, event or set of events, procedures, and any other controllable variables (e.g., weather, failure modes) that are of interest to the outcome of the simulation. Scenarios in fast-time simulation typically represent normal operating conditions of a new concept for analysis. In other words, if a new procedure is being evaluated, the fast-time model runs assuming the procedure performs without fail throughout the scenario. This is an effective approach when predicting potential capacity gains, for example, of a new concept.

3.3.3 Baseline Scenario

A baseline scenario is a reference scenario that provides a benchmark against which the effects of the experimental conditions (i.e., proposed operational concepts) can be evaluated. The baseline scenario establishes a point of comparison. It implements the scope and control conditions for testing of the hypothesis. It does not implement the experimental conditions.

4 Scenario Uses

Scenarios are a core component of many types of evaluations and simulations. Their application is not limited to any one type or level of testing. Rather, they can provide the basis for exploratory studies, demonstrations, usability testing, and operational test and evaluation activities. The objectives of scenario use and the actual way in which they should be used can change with increasing maturity in the concept development. This section begins by discussing the use of scenarios in the typical Concept Validation stages (V1 -V3) of a concept's design and is based on the levels of maturity outlined in the OCVSD.

4.1 Scenarios and Levels of Maturity

4.1.1 V1 Establishment of Concept Principles

In the earliest stages of concept development, the concept designers will not have a well-formed idea of the concept principles or operational and technical requirements. The scenarios at this stage will be focused on understanding the potential of the concept. The purpose at this stage is to help form the principles, eliminate poor design choices and focus on further design. Scenarios will be focused on establishing the scope of the concept.

4.1.2 V2 Proof of Concept

In the next stage of development, basic assumptions should already have been tested and basic design already formulated. At this level, scenarios should focus on more advanced aspects of the concept design, proof of use, non-nominal cases or marginal capabilities. At this stage, there may still need to be a repetition of the type of analysis performed after V1, but mainly the scenarios here will not be so singularly focused. The outcome of this stage in the validation process should be a mature, stable concept design with an initial proof of concept. Scenarios will be focused on setting the limits of the concept, establishing procedures and phraseology and determining clear requirements to assist in producing a stable environment for the final pre-implementation phase.

4.1.3 V3 Concept Integration and Pre-Ops Trials

In the final stage of the Concept development, the scenarios will focus on repetition, replication and performance measurement. The experiments here will be assessing a mature concept in a rigorous, high fidelity experiments (large fast-Time or real-time). As such, the scenario will need to be designed with choice of indicators and metrics in mind. Large detailed scenarios will be built to model the whole system under assessment. The scenarios will need to be run repeatedly in order to generate enough data for performance measurement, with baseline scenarios taken into consideration. The outcome of this stage in the validation process will be detailed information about the expected performance of the system under a variety of normal conditions and the ability of the system to cope with various non-normal conditions.

4.2 Scenario Objectives

4.2.1 Exploration

Scenarios can be used to develop and refine a concept and explore what-if questions. Methods suitable to meeting these objectives include cognitive walkthroughs, part task simulations, and fast-time simulations. Exploratory scenarios are generally a low-cost means of identifying issues and refining the experimental approach for future studies on a proposed operational concept.

4.2.2 Demonstration

Scenarios can be used to show or exhibit a potential capability to a targeted audience. They can be incorporated into demonstrations, which although do not prove a concept, can give interested parties a view of how a concept might appear in the operational environment. With demonstrations, researchers may have to forego strict experimental design – e.g., procedures may evolve as demonstration progresses, but this does not preclude the value of the information gained from the demonstration and the scenarios that comprise it.

4.2.3 Usability Testing

Scenarios created for usability testing can help to determine the appropriateness of a proposed tool, procedure, and/or equipment. Usability testing typically examines user preference and performance for a specific use, and thus scenarios are generated correspondingly to address issues associated with the proposed concept.

4.2.4 Operational Test & Evaluation

Operational test and evaluation (OT&E) scenarios are used to confirm that the 'system' under test performs to predetermined standards and/or criteria when stressed. The OT&E scenario is designed to elicit performance to the level that it demonstrates the system works as specified under all foreseen exceptions and failure modes. It is also designed to answer questions related to the expected benefits and expected costs of implementing a new concept.

5 Coordination of Fast & Real-Time Simulation Activities

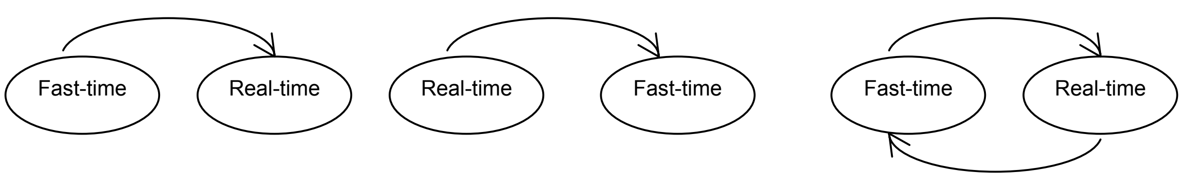
Table 1 presents a list of recommended best practices to be considered when co-ordinating scenarios between fast-time and real-time simulation activities. Each best practice is described in more detail following the table.

Table 1 - Best Practices for Coordination of Fast & Real-time Scenario Development

5	Coordination of Fast & Real-time Simulation Activities
bp 2-01	Communicate with fellow fast-time and real-time researchers
bp 2-02	Clearly state and document your design assumptions
bp 2-03	Clearly state the likely implications of your assumptions
bp 2-04	Standardize scenarios for use in both fast- and real-time simulations when possible

Traditionally validation has taken the approach suggested in the Validation Route Map, that first you do fast-time, followed by a real-time simulation. Although not as customary, one approach is to perform real-time experiments, followed by fast-time. In reality, there is no set pattern on which validation technique should follow which, providing the selected one is appropriate. A thorough analysis actually involves multiple iterations using different types of evaluation tools. The "Banana Model" provides a good basic illustration of this process, but it presents the illusion that it is a one-way flow. In fact, there are two-way flows to every node in the model. With each iteration, the researcher uses information gained from the previous analysis, revisits assumptions and inputs, refines approaches, and selects the appropriate tool for future analysis.

When considering the coordination of fast-time and real-time validation techniques, the link between them should reflect the iterative nature of the relationship, as demonstrated in Figure 1. This iterative approach suggests the need for greater coordination in the design, development and use of scenarios, and in particular adopting a common structure, data sharing and re-using information where possible. It is the aim of the best practices presented in this document to promote this coordination.



Traditional Approach Alternative Approach Reality

Figure 1 - Iterative nature of fast- and real-time analyses.

Certain software packages are already under development with the aim of combining the characteristics of the fast-time and real-time simulation techniques. This development highlights the need for greater coordination in scenario development and the need for a common structure and leads to the first Best Practice recommendation on this document.

bp 2-01. Communicate with fellow fast-time and real-time researchers

In the interest of better coordination, those working on real-time simulations and those working on fast-time simulations should communicate and interact with each other regarding their work. Often the two parties are working independently on the same or similar issues. However, they should communicate about their efforts, and share information on such issues as assumptions, results, and lessons learned.

5.1 Data Sharing and Re-Use between Fast-Time & Real-Time

Data sharing between fast- and real-time simulations can occur in two forms: one focused on the data input, and one on data output.

The first type is data originating from a single source is then used as input into the scenario design for both types of simulations. The benefits of using common data sources are faster scenario development times, common baselines, and the ability to compare results.

The second form of data sharing concerns the re-use of data used by or produced by another/previous simulation. Such data could include the following:

- System performance, response times (e.g., Controller-Pilot Data Link Communications (CPDLC)), and variability data
- Human error models
- Trajectory prediction inaccuracy, and aircraft position inaccuracy
- Battery of flow-upsetting parameters/events
- Traffic forecast for target years
- Baseline (Decision Support Tools (DST's), traffic, airspace, procedures, roles and responsibilities, operators, etc.))

The researcher should always use caution when considering sharing data in either form. He/she must ask questions: Are the chosen inputs suitable in both types of simulation? Is the specific type of data produced in the real-time simulation (e.g., controller response times) realistic enough to put into the fast-time model? What impact would a learning effect demonstrated in a real-time simulation data have on the fast-time model? Could the fast-time model erroneously magnify or overstate any of the impacts demonstrated in the real-time?

To combat these risks this document proposes the following best practises.

bp 2-02. Clearly state and document your design assumptions

One of the difficulties in coordinating fast and real-time simulation activities and sharing data is the different types of assumptions that go into each, which can cause problems when trying to share or re-use data. As such, it is very important that both fast and real-time researchers clearly state their assumptions before performing their analyses. Perhaps a common template could be developed. That way, even from the outside, an expert could look at the template, identify, and understand the assumptions. This is vital for interpretation.

bp 2-03. Clearly state the likely implications of your assumptions

The researcher should state not only assumptions, but also the implications of the assumptions. This helps in the correct interpretation of a model as often a model can be wrongly interpreted if the implications of its assumptions are not taken into account. This is particularly important if the assumptions themselves are implicit. Documentation would help eliminate these risks and would provide traceability by showing which outputs came from which sources, what assumptions affected what outputs, etc.

bp 2-04. Standardize scenarios for use in both fast- and real-time simulations when possible.

To help facilitate coordination between scenario development in fast-time and real-time simulations, a standardized scenario development process and characterization scheme could be developed. These could be designed around each technique with common processes and areas noted for coordination purposes. This would be useful when designing scenarios, building scenarios, and possibly even comparing results after the simulations.

5.2 Technology Advancement

Due to the different objectives of fast-time and real-time simulations and the differing techniques and expertise involved for each, coordination historically between fast-time and real-time has not been easy. For example, it is often very difficult to model certain concepts in the fast-time environment. A view has been taken that some concepts just are not suitable for evaluation in the fast-time environment. Part of the problem has been that the fast-time simulation technology has not evolved to a level suitable to simulate the concepts. Emerging technologies, however, may solve this problem. Fast-time simulation software is continually under development and may in the future be used increasingly for what were traditionally subjects for real-time assessments only. Researchers may see more hybrid fast-time/real-time models in the years to come.

6 Scenario Development Considerations

Table 2 presents a list of recommended best practices to be considered when developing and designing scenarios. Each best practice is described in more detail following the table.

Table 2 - Best Practices for Scenario Development and Design

6	Scenario Development Considerations
bp 2-05	Allow the research questions to drive scenario development
bp 2-06	Map out and prioritize all potential scenario characteristics
bp 2-07	Identify and maintain common scenario characteristics for comparison
bp 2-08	Check that the test environment can support your scenarios and be prepared to make compromises
bp 2-09	Start with baseline derived from current information (including traffic)

bp 2-10	Consider options such as "generic airspace"
bp 2-11	Seek subject matter expertise during the development of your scenarios
bp 2-12	Obtain involvement from all stakeholders during scenario design
bp 2-13	Define and maintain necessary levels of realism
bp 2-14	Strike a balance in the order of the scenario presentation
bp 2-15	Scenarios should carefully isolate key technologies that require large capital investments in order to be clear on potential marginal gains
bp 2-16	A generic list of 'minimum' scenarios should be formed

Researchers have many factors to consider when designing scenarios. They know that carefully designed scenarios produce the most useful test results and that poorly crafted scenarios can preclude the achievement of test objectives. As such, the following section provides guidance on points to consider in the scenario development process.

***bp 2-05.* Allow the research questions to drive scenario development**

While the researcher should strive to develop scenarios that model the operational environment, it is crucial to build scenarios to address the research questions and objectives. All perspectives (researcher, operational, and management) should be taken into consideration (Best Practices – Appendix 1, bp 1-16).

***bp 2-06*. Map out and prioritize all potential scenario characteristics**

Scenarios can be characterized on several different levels; by concept aspects (roles, procedures, and sequence of tasks), environment aspects (airspace, route structure, traffic volume), and event aspects (normal events, non-normal events). The researcher should develop a scheme to map out and prioritize potential scenario characteristics to aid in the development process.

***bp 2-07.* Identify and maintain common scenario characteristics for comparison**

One of the biggest challenges in scenario development is to create scenarios that are experimentally comparable, but that are different enough to present a "new" problem for the participants. The experimenters often have to design scenarios that are similar (in complexity) but not exactly same. The similar scenarios are essential part of the good experimental design principles as they help statistical comparisons. However, scenarios that are similar could produce (often undesirable) learning effects. The learning effects could negate the effects of experimental conditions. Simply changing the aircraft call signs are not always adequate since controller memory recognizes air traffic patterns as well. Learning effects may be controlled somewhat be randomizing the order of experimental conditions.

This Best Practice is taken from Appendix 1 (bp 1-17). For more information, please refer to the text in that appendix.

***bp 2-08*. Check that the test environment can support your scenarios and be prepared to make compromises**

In keeping with the FAA / EUROCONTROL principles of concept levels of maturity (V1 to V5) the scenarios should become more mature and rigorous as the concept becomes more mature. If the scenario changes, ensure that the simulation environment is capable of supporting the scenario requirements.

***bp 2-09*. Start with baseline derived from current information (including traffic)**

Baseline traffic samples should be derived from current traffic. The use of baselines developed by past information should be avoided, as it can be hard to get all the necessary data. Procedures, tools, and traffic schedules continuously change; therefore obtaining current information helps establish a more realistic baseline. Although in the fast-time environment, one has the ability to create a future baseline based on known future improvements. For example, if a runway is to be in place at an airport a year from now, modelling the current airport will have no value in comparing any new improvements.

***bp 2-10*. Consider options such as "generic airspace".**

In many real-time, HITL studies, actual airspace is replicated and real traffic scenarios are modelled to produce a familiar and realistic environment for data collection. For many exercises, using site-specific/existing airspace and traffic may be an essential requirement, such as when assessing the impact of a proposed new runway on operations at a specific airport. However, using site-specific airspace may induce constraints on the sample of subjects who participate in the study. Using generic airspace may be an option for studies that utilize participants from several different facilities or for general studies that do not need to be applied to a specific airspace.

This Best Practice is taken from Appendix 1 (bp 1-15). For more information, please refer to the text in that appendix.

***bp 2-11*. Seek subject matter expertise during the development of your scenarios**

The researcher should acquire subject matter expertise (SME) from air traffic control specialists, operational personnel, and/or individuals familiar with the domain of interest and corresponding traffic characteristics to assist in the development of scenarios. SMEs can help build and validate the scenarios. They can assist the researcher in determining the number of scenarios required, for example, to achieve the desired simulation results. They can also point out weaknesses in the scenarios that may lead to incomplete data or missed opportunities. Having an SME participate in the scenario development process will add credibility and fidelity to any simulation effort.

***bp 2-12*. Obtain involvement from all stakeholders during scenario design**

The researcher should try to obtain input from the various stakeholders involved with a simulation. Including sponsors and potential users of the operational concept under test, for example airport and airline representatives, and when possible, Airline Operation Center representatives, is key to gaining buy-in of simulation results. This best practice does not only apply to scenario development, but also to the overall simulation effort.

***bp 2-13*. Define and maintain necessary levels of realism**

While developing scenarios, researchers are usually interested in maintaining high realism related to aircraft mix, traffic density, sector geometry, routes, and procedures. Many times, the scenario development process starts from collecting actual operational flight plan data, which provides the realism. The operational data with realistic traffic density and aircraft types works well for near-term initiatives (e.g., Reduced Vertical Separation Minima). However, for future concepts, such as free manoeuvring, operational data may need considerable modifications in order to satisfy the experimental objectives. In such situations, a number of variables, such as future traffic load, future aircraft mix, and technologies, may need to be modified. Such modifications to better model the future environment may not be realistic in the present day operations, however will represent the future environment.

This Best Practice is taken from Appendix 1 (bp 1-19). For more information, please refer to the text in that appendix.

***bp 2-14*. Strike a balance in the order of the scenario presentation**

While sound experimental design in real-time simulation calls for randomization of scenario presentation, over-randomization can lead to an unfair variance in the results. Researchers should keep in mind that controllers are accustomed to a certain ‘order’ in their work. Over-randomizing the scenarios is unnatural. On the other hand, presenting scenarios in too similar of an order can lead to over-familiarization, or a learning effect. Balance is required. This is a problem in real-time, but not in fast-time, as a computer will not ‘learn’ the sample.

***bp 2-15*. Scenarios should carefully isolate key technologies that require large capital investments in order to be clear on potential marginal gains**

When developing scenarios that require an assessment of multiple concepts, it is essential that you isolate the effects and performance changes of each concept, thereby making clear the benefit of each. There are certain concepts that require a large investment in time and resources, therefore the cost/benefit analysis must show the benefit of these particular concepts without the additional or cumulative affects of any other concept in your scenario.

***bp 2-16*. A generic list of ‘minimum’ scenarios should be formed**

Most concept validation simulations are performed based on a single set of scenario conditions and does not take into account many aspects of the ATC environments. A solution to this problem would be a common set of scenarios based on the type of simulation performed.

For example, at a minimum you should be required to perform one bad weather experiment, one runway change, one go around, and one level bust. Granted, this solution does not always fit; time, resource, and cost play an important part in how many experiments can be run.

The use of this list will enable some cross-referencing among different concepts or simulations. The use of such a list is also particularly useful when assessing concepts of higher maturity by way of providing a set of 'bench-mark' tests.

7 Building of Traffic

Table 3 presents a list of recommended best practices to be considered when building traffic for scenarios. Each best practice is described in more detail following the table.

Table 3 - Best Practices for Building Traffic

7	**Building Traffic**
bp 2-17	Starting and/or ending scenarios slowly is not usually efficient
bp 2-18	Traffic peaks and troughs may be relevant to research
bp 2-19	Validate aircraft performance
bp 2-20	Don't exceed the airport projections when projecting enroute or terminal growth
bp 2-21	Consider city pairs, when possible
bp 2-22	Translate yearly forecasts (peak) into daily and hourly operations
bp 2-23	Try to identify aircraft types and mix

Developing scenarios usually involves a large investment in time and cost. A large percent of developing scenarios is in building the traffic sample. A 'basic' traffic sample is not so hard but a large percentage of time is spent on getting realism in the sample. For fast-time (e.g. RAMS) an estimate of total effort shows around 20-30% of total effort spent on traffic preparation.

***bp 2-17*. Starting and/or ending scenarios slowly is not usually efficient**

A common way of building a scenario is to initiate traffic gradually into the problem, allowing the controllers to ease themselves into the scenario. When the main part of the problem is over, traffic is usually tapered off. For an hour-long scenario, as much as 15 minutes on both ends of the scenario might be dedicated to “ramp up/down” time. Given the time and cost restraints to run a simulation, this is not the most effective or efficient use of scenario time. A more efficient means of building a scenario is to begin with a normal traffic load. Depending on the study, practitioners may also end the scenario in the middle of a conflict or other problem. (Best Practices – Appendix 1, bp 1-18)

***bp 2-18*. Traffic peaks and troughs may be relevant to research**

HITL practitioners often try to impose high workload levels on simulation participants to stretch the limits of the participant’s abilities, and to test the limits of new concepts and procedures on the air traffic system. To do this, they typically design traffic scenarios to represent peaks or high levels of traffic activity. Slower-manifesting problems are generally avoided since the practitioner wants to make the best use of valuable time. Prior simulation research has shown, however, that operational errors often occur in the beginning of troughs, or lower levels of traffic, immediately following very high levels of traffic activity (this results because of a temporarily perceived reduction of complexity and thus lower vigilance).

In order to better emulate the operational environment and to capture all possible conditions for human error, HITL practitioners should script a range of traffic activity into their scenarios; those which include both peak levels of traffic and troughs. (Best Practices – Appendix 1, bp 1-20)

***bp 2-19*. Validate aircraft performance**

The validation of aircraft performance is very important for realism, especially considering that the purpose of the simulation may be to introduce a new aircraft type. Controllers can detect if an aircraft performance is suspect by seeing the aircraft movements. Steps have been taken to standardize aircraft performance. However, even given this, realism of aircraft performance continues to be an issue.

***bp 2-20*. Don't exceed the airport projections when projecting enroute or terminal growth**

When developing traffic samples, don't exceed the level of traffic that is projected for the origin or destination airports. Exceeding the level of traffic at the origin or destination airport will cause unrealistic delays and throughput in the airspace system. For example, don't have 80 a/c per hour flying between airports with a 35 a/c per hour capacity.

***bp 2-21*. Consider city pairs, when possible**

A common practice in today's environment is to simulate each airport or section of airspace as a stand-alone system, without taking into account the affects on certain city pairs. When generating traffic samples, the origin and destination airports must be able to handle the level of traffic that is projected for each. New metrics designed around city-pair performance are becoming a requirement. Traffic should be adjusted based on terminal area forecasts and/or other business models, to provide a realistic future traffic sample. Certain models or simulators require city-pairs to execute properly, so having city-pair information will aid in re-use and/or real-time/fast-time coordination. Finally, traffic generated for a certain city could be used in metro areas.

***bp 2-22*. Translate yearly forecasts (peak) into daily and hourly operations**

The Terminal Area Forecasts (TAF) provided by the FAA does not provide daily or hourly operations. For most simulations, these forecasts must be translated to daily and hourly operations. To translate into daily and hourly operations a baseline of current traffic should be developed through data gathering and adjusted based on traffic forecast. These techniques need examination, since they have not been perfected or standardized, especially when it comes to smoothing peak periods of traffic.

A point of caution here is to take abnormal peaks or troughs into account. For various reasons (e.g. a once-yearly major event in the area) traffic levels may experience an abnormal growth at certain times. Such occasions, although not normal, must be considered.

***bp 2-23*. Try to identify aircraft types and mix**

Environmental studies require the analysis of specific aircraft types especially during noise studies, so individual aircraft types and fleet mix is essential. Some simulation models only require aircraft groups or classes, but when the traffic sample is used in other simulation models, it becomes very difficult to re-engineer. Identifying individual aircraft types also helps in coordination between fast and real-time simulations, where in real-time simulation the aircraft type is a necessity.

8 Scripting of Specific Events

Table 4 presents a list of recommended best practices to be considered when scripting events. Each best practice is described in more detail following the table.

Table 4 - Best Practices for Scripting Specific Events

8	Scripting Specific Events
bp 2-24	Plan more scripted events than you need to ensure desired results
bp 2-25	Forewarn participants
bp 2-26	Do not try to script operational errors
bp 2-27	Consider scripting system-wide events
bp 2-28	Consider the use of dynamic scripting
bp 2-29	Do not try to recover dynamically

8.1 Scripted Scenarios

A Scripted Scenario is one that deliberately introduces a disturbance into the system. Such a disturbance is generally known in advance to the validation team but not the simulation participants. They are particularly useful for assessing the impact of 'unplanned' events in a simulation environment where, if you did not deliberately introduce them, you would probably not encounter these (often critical) situations.

***bp 2*-24. Plan more scripted events than you need to ensure desired results**

When accessing a certain operational concept, you may need certain events to happen to show the benefits of the concept. In this situation, you should plan more scripted events than needed to ensure desired results. For example, script nine potential conflicts to obtain at least four. However, adding more events should not compromise realism.

***bp 2-25*. Forewarn participants**

In real-time simulation, it is advisable to forewarn participants about what is expected in the simulation. However, any pre-warning should be kept to a minimum, since you still require controllers to work normally and not in a contrived manner.

***bp 2-26*. Do not try to script operational errors.**

It is useful to study the occurrence and impact of operational errors. There may be a temptation to script certain events to induce such errors. However, this practice is not recommended. Errors such as these should simply be allowed to happen in order not to bias any results or even controller opinions of the concept.

Should such errors occur be careful how you report it. The report should not be seen to take a 'telling tales' approach. Perhaps rather than referring to Operational Errors, introduce some adaptation of 'performance scoring'. This should enable the same discussion but appears less negative on the participants. This is necessary because in any simulation you may get an unfair proportion of Operational Errors through controller learning or controllers experimenting with the system.

***bp 2-27*. Consider scripting system-wide events.**

As well as using scripts that relate to single events in the system being simulated, introduce system wide scripts such as airport closure, system failures, etc. Obviously, this can be a major body of work and may only be done once or twice in a simulation. However, stakeholders are increasingly concerned about the affects of such system-wide events.

***bp 2-28*. Consider the use of dynamic scripting.**

Predetermining scripts may not always be easy, especially in new systems or if you are unsure of the events beforehand. The dynamic nature of the system means that events can unfold rather differently than planned. A way to adjust when scripts don't go as planned, is to use dynamic scripts. For example, instead of picking a particular aircraft, wait and see how the situation unfolds and pick another aircraft. This dynamic nature also means it can be difficult to accurately reproduce events.

***bp 2-29*. Do not try to recover dynamically**

Whilst the use of dynamic scripting can be useful, trying to recover dynamically is not always recommended, especially in experiments where performance measurements and comparisons are being made. Any attempt to recover in such a way will greatly influence any data being recorded and thus statistical analysis will be useless.

In the early stages of concept development, the participants may be allowed to recover dynamically. This may help in the development of roles and procedures should these not already be defined.

9 Validation Data Repository

9.1 How the VDR Can Help

The Validation Data Repository (VDR) is a means of storing data, but it should also establish a common reference framework with which validation activities can and should be mapped. With regard to ‘scenarios,’ this has already begun with the introduction of a separate "Concepts and Scenarios" view. While the Concepts view is not as clearly defined yet, the Scenario sections are becoming quite mature.

At present, the VDR provides a structure for scenario information giving:

- Summary information about a scenario
- Descriptions of relevant operational concepts under test
- Defining characteristics of the scenario (e.g., geographic area, airspace design, traffic samples and profiles, system configurations, task allocations, procedures)
- Hyperlinks to detailed source reference documents

In terms of the re-use of information, the VDR enables mapping of scenarios to exercises, or in other words, provides a means of referencing scenarios. The VDR can also map a scenario to more than exercise. This means standard scenarios can be defined and linked to many exercises.

9.2 How the VDR Treats Scenarios

The VDR provides a structure that defines information in a consistent manner to enable scenario comparisons, but also provides sufficient flexibility to allow researchers to make comparisons based on their needs. Figure 3 depicts the role of scenarios in the VDR.

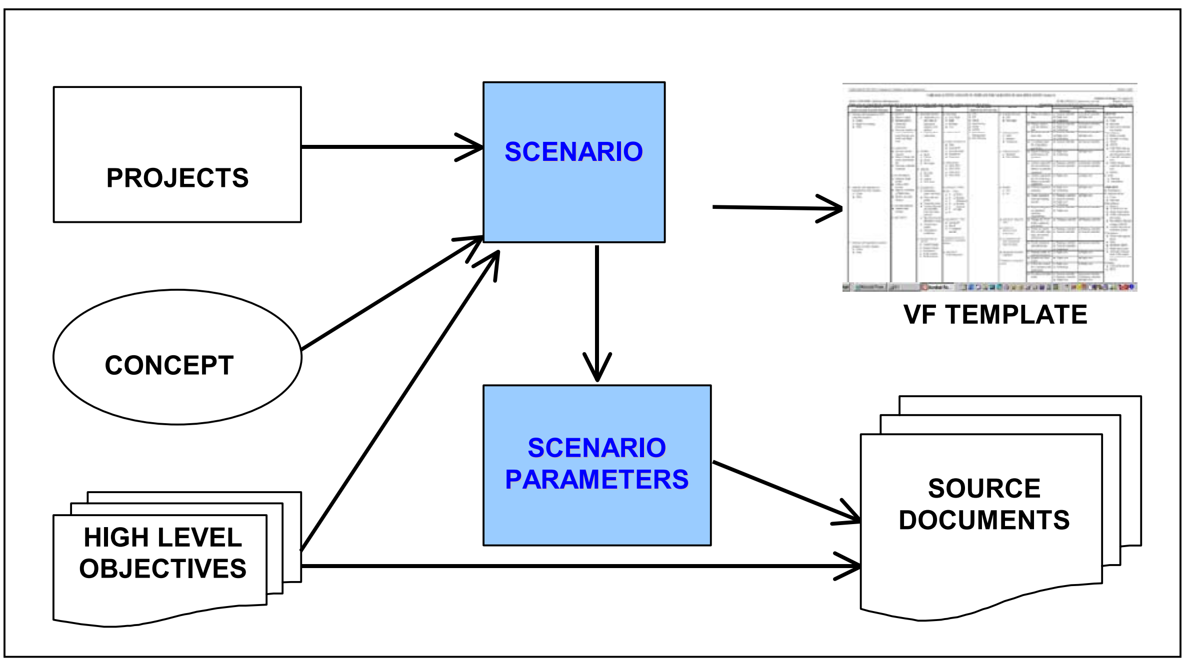


Figure 2 - The role of scenarios in the VDR.

10 Steps in Scenario Design

The following outline reviews the basic steps to be carried out when designing scenarios. The steps are iterative in nature, and as such should be continually applied as the concept matures.

1. Clearly define simulation objectives
   - Consider concept level of maturity

   - Consider operational understanding
   - Obtain stakeholder input

2. Identify and prioritize appropriate metrics/ evaluation criteria

3. Determine relevant scenario characteristics
   - Bear in mind that scenarios should cover a variety of dimensions: normal to abnormal, simple to complex, low to high taskload, and procedural tasks versus problem solving
   - Consider including a wide range of scenarios to allow for greater generalization of results
   - Consider a range of scenario characteristics, including the following:
     - Domain
     - Scope
     - Duration
     - Fidelity
     - Focal points
     - Experimental validity
     - Human performance requirements
   - Obtain stakeholder input

4. Determine assumptions and limitations

5. Identify data management and analysis requirements

6. Prepare detailed reports
   - Document assumptions and limitations and their implications on the results
   - Obtain stakeholder input

11 Summary of Best Practices in Scenario Design

The collaborative Action Plan 5/Action Plan 9 Workshop brought together a range of practitioners from the U.S. and Europe with experiences in fast-time and real-time simulation to discuss Scenario Development for purposes of concept validation. The following is a summary of those discussions and the best practices that emerged related to scenario design.

The practitioners' expertise and backgrounds varied; therefore, a common definition of scenario had to first be developed. To satisfy both fast-time and real-time practitioners, a definition of scenario was created for each technique. Scenario uses and scenario objectives were presented, providing a common ground for topic discussions.

Coordination of fast and real-time simulation was a topic that generated much discussion. It was clear that few practitioners had coordinated fast-time and real-time activities. It was also clear that many practitioners of fast-time and real-time simulation were inexperienced in each others' areas of expertise. Increased communication between fast-time and real-time practitioners became one of the first best practices. Standardizing scenarios and clearly stating and documenting design assumptions and the implications of such assumptions for both real-time and fast-time was recommended to allow for data sharing and re-use between the two techniques. Through topic and informal discussion, it was clear that the issue of coordination between fast-time and real-time activities should be addressed further.

Although there are quite a number of factors to consider, the practitioners composed the following 12 best practices for scenario development:

BP-1. Allow the research questions to drive scenario development. It is crucial to build scenarios to address the research questions and objectives.

BP-2. Map out and prioritize all potential scenario characteristics. The researcher should develop a scheme to map out and prioritize potential scenario characteristics to aid in the development process.

BP-3. Identify and maintain common scenario characteristics for comparison. One of the biggest challenges is to create scenarios that are experimentally comparable, but that are different enough to present a "new" problem for the participants.

BP-4. Ensure that the test environment can support your scenarios and be prepared to make compromises. If the scenario changes, ensure that the simulation environment is capable of supporting the scenario requirements.

BP-5. Start with baseline derived from current information (including traffic). It is difficult to obtain necessary data from past information, procedures, and tools, and traffic schedules are constantly changing.

BP-6. Consider options such as generic airspace. Using site-specific airspace may induce constraints on the sample of subjects who participate in the study.

BP-7. Seek subject matter expertise during the development of your scenarios. Subject matter experts can help build and validate the scenarios, point out weaknesses, and add credibility and fidelity to any simulation effort.

BP-8. Obtain involvement from all stakeholders (e.g., sponsors, airlines, unions) during scenario design. This is key to gaining buy-in of the simulation results. This best practice does not only apply to scenario development, but also to the overall simulation effort.

BP-9. Define and maintain necessary levels of realism. When developing baseline scenarios, researchers usually are interested in maintaining high realism to emulate current operations as closely as possible. However, for future concepts, scenarios may need to be modified to reflect projected environments and operations to satisfy experimental objectives.

BP-10. Strike a balance in the order of scenario presentation. Researchers should keep in mind that controllers are accustomed to a certain 'order' in their work. Over-randomizing the scenarios is unnatural. On the other hand, presenting scenarios in too similar of an order can lead to over-familiarization.

BP-11. Scenarios should carefully isolate key technologies that require large capital investments in order to be clear on potential marginal gains. The cost/benefit analysis should show the benefits of these key technologies.

BP-12. A generic list of 'minimum' scenarios should be formed. The use of this list will enable some cross-referencing among different concepts or simulations.

Building traffic takes a considerable amount of effort in scenario development. As such, the practitioners identified key considerations related to this activity. For example, with respect to traffic volume in real-time simulation, starting or ending scenarios slowly is not usually efficient and traffic peaks and troughs may be relevant to research. In the case of aircraft performance validation, unrealistic performance can be detected by an experienced controller. The practitioners also addressed building traffic to represent growth based on airport and city pair projections, translating yearly forecasts into daily and hourly operations, and trying to identify the appropriate aircraft types and mixes.

Scripting an event is deliberately introducing a disturbance into the system. In real-time simulation, this would be accomplished by planning a specific type of event and introducing it during the simulation run to induce some kind of human response. In fast-time simulation, a scripted event is usually introduced at a specific time during a simulation run. The best practices resulting from the workshop that relate to scripted events mostly deal with real-time simulation. Best practices include the following: Planning more scripted events than needed ensures desired results. Forewarning participants as to what to expect is a common practice, but one should keep this to a minimum. Intentional scripting of operational errors should not occur, operational errors should happen naturally in the simulation. Consideration should be taken to script system wide events.

Practitioners should keep in mind that the VDR provides a means of storing data, including scenario information, and it establishes a common reference framework with which validation activities can and should be mapped. The VDR provides a concept and scenario view. While the concept view is not clearly defined, the scenario view is quite mature.

Based on feedback surveys, most practitioners felt that the workshop was useful and that other workshops should be developed in the future. Metrics and measures, reporting of results, and more specific scenario development topics (separate fast and real-time) were some of the suggestions for future workshops. Bringing together practitioners from different backgrounds to discuss "best practices" was a challenge, but overall the process was a successful one.